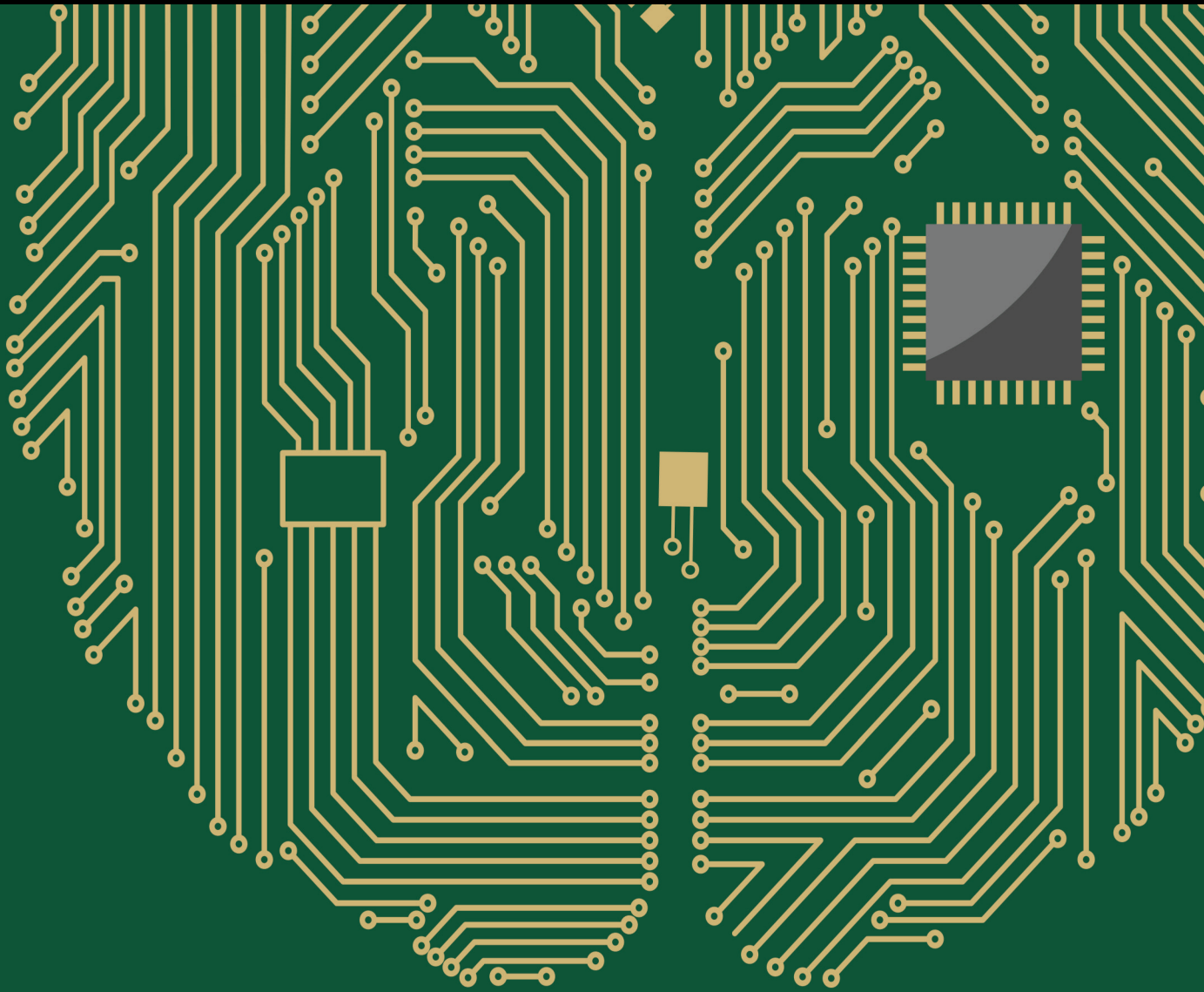


Mental Illness Detection and Analysis on Social Media

Lead Guest Editor: Deepika Koundal

Guest Editors: Alexander Gelbukh, Syed Ahmad Chan Bukhari, and Gabriella Casalino





Mental Illness Detection and Analysis on Social Media

Computational Intelligence and Neuroscience

Mental Illness Detection and Analysis on Social Media

Lead Guest Editor: Deepika Koundal

Guest Editors: Alexander Gelbukh, Syed Ahmad
Chan Bukhari, and Gabriella Casalino




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 Razmjoo, 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: A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9878345, Volume 2023 (2023)

Retracted: Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9890201, Volume 2023 (2023)

Retracted: A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9872487, Volume 2023 (2023)

Retracted: Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9846879, Volume 2023 (2023)

Retracted: Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9839672, Volume 2023 (2023)

Retracted: Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9825080, Volume 2023 (2023)

Retracted: Psychological Analysis for Depression Detection from Social Networking Sites

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9796187, Volume 2023 (2023)

Retracted: Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9792821, Volume 2023 (2023)

Retracted: Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9785326, Volume 2023 (2023)

Retracted: Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9780642, Volume 2023 (2023)

Retracted: Identification and Classification of Depressed Mental State for End-User over Social Media

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9756137, Volume 2023 (2023)

Retracted: Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9754359, Volume 2023 (2023)

Retracted: Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9896136, Volume 2023 (2023)

Retracted: Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9893015, Volume 2023 (2023)

Retracted: Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9871294, Volume 2023 (2023)

Retracted: Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9869258, Volume 2023 (2023)

Retracted: Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9851365, Volume 2023 (2023)

Retracted: Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9834938, Volume 2023 (2023)

Retracted: A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9798269, Volume 2023 (2023)

Contents

Retracted: Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9798169, Volume 2023 (2023)

Retracted: Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9790823, Volume 2023 (2023)

Retracted: Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9781063, Volume 2023 (2023)

Retracted: Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9762865, Volume 2023 (2023)

Retracted: Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9756078, Volume 2023 (2023)

Retracted: Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9860875, Volume 2023 (2023)

Retracted: Human-Computer Interaction Using Manual Hand Gestures in Real Time

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9832754, Volume 2023 (2023)

Retracted: Evolving Long Short-Term Memory Network-Based Text Classification

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9827632, Volume 2023 (2023)

Retracted: Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9878074, Volume 2023 (2023)

Retracted: Secure Complex Systems: A Dynamic Model in the Synchronization

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9874218, Volume 2023 (2023)

Retracted: Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9871283, Volume 2023 (2023)

Retracted: A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9856274, Volume 2023 (2023)

Retracted: Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9815652, Volume 2023 (2023)

Retracted: Learning Enhanced Feature Responses for Visual Object Tracking

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9815496, Volume 2023 (2023)

Retracted: Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9795101, Volume 2023 (2023)

Retracted: Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9839162, Volume 2023 (2023)

Retracted: Computational Intelligence Approaches in Developing Cyberattack Detection System

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9837830, Volume 2023 (2023)

Retracted: DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9832183, Volume 2023 (2023)

Retracted: A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9828479, Volume 2023 (2023)

Retracted: Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9817176, Volume 2023 (2023)

Contents

Retracted: Stress-Relieving Video Game and Its Effects: A POMS Case Study

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9810102, Volume 2023 (2023)

Retracted: A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9804521, Volume 2023 (2023)

Retracted: 5G Massive MIMO Signal Detection Algorithm Based on Deep Learning

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9795427, Volume 2023 (2023)

Retracted: Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9794125, Volume 2023 (2023)

Retracted: An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9793473, Volume 2023 (2023)

Retracted: Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9764674, Volume 2023 (2023)

Retracted: A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9763069, Volume 2023 (2023)

Retracted: Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9757924, Volume 2023 (2023)

Retracted: Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9753840, Volume 2023 (2023)

Retracted: Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques

Computational Intelligence and Neuroscience

Retraction (1 page), Article ID 9893215, Volume 2023 (2023)

Retracted: Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9892510, Volume 2023 (2023)

Retracted: Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9867256, Volume 2023 (2023)

Retracted: TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9857835, Volume 2023 (2023)

Retracted: Applying Dynamic Systems to Social Media by Using Controlling Stability

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9854161, Volume 2023 (2023)





Retracted: Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpening of Remote Sensing Images

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9846743, Volume 2023 (2023)

Hybrid Recommender System for Mental Illness Detection in Social Media Using Deep Learning Techniques

Sayed Sayeed Ahmad , Rashmi Rani, Ihab Wattar , Meghna Sharma, Sanjiv Sharma, Rajit Nair , and Basant Tiwari 
Research Article (14 pages), Article ID 8110588, Volume 2023 (2023)

Multi-Ideology, Multiclass Online Extremism Dataset, and Its Evaluation Using Machine Learning

Mayur Gaikwad , Swati Ahirrao , Shraddha Phansalkar , Ketan Kotecha , and Shalli Rani
Research Article (33 pages), Article ID 4563145, Volume 2023 (2023)

Retracted: An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding



Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9850820, Volume 2023 (2023)

Retracted: Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty

Computational Intelligence and Neuroscience
Retraction (1 page), Article ID 9792521, Volume 2022 (2022)

Contents

[Retracted] Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning

Tanzila Saba, Amjad Rehman Khan, Ibrahim Abunadi , Saeed Ali Bahaj, Haider Ali , and Maryam Alruwaythi

Research Article (9 pages), Article ID 8622022, Volume 2022 (2022)

[Retracted] Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury

Wanliang Yang , Xin Pan , Yibo Wang, and Wenhao Chen




Research Article (7 pages), Article ID 6861749, Volume 2022 (2022)

[Retracted] Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems

Alka Londhe , P. V. R. D. Prasada Rao , Shrikant Upadhyay , and Rituraj Jain 


Research Article (14 pages), Article ID 8579640, Volume 2022 (2022)

[Retracted] Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks

Jatla Srikanth, Avula Damodaram, Yuvaraja Teekaraman , Ramya Kuppusamy , and Amruth Ramesh Thelkar 







Research Article (11 pages), Article ID 8898100, Volume 2022 (2022)

[Retracted] Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning

T. Veeramakali, A. Shobanadevi , Nihar Ranjan Nayak, Sumit Kumar, Sunita Singhal, and Manoharan Subramanian 




Research Article (8 pages), Article ID 4690936, Volume 2022 (2022)

[Retracted] Identification and Classification of Depressed Mental State for End-User over Social Media

Akhilesh Kumar , Anuradha Thakare , Manisha Bhende , Amit Kumar Sinha , Arnold C. Alguno , and Yekula Prasanna Kumar 

Research Article (10 pages), Article ID 8755922, Volume 2022 (2022)

[Retracted] Stress-Relieving Video Game and Its Effects: A POMS Case Study

Abdullah Ajmal, Hamza Aldabbas , Rashid Amin , Sundas Ibrar, Bader Alouffi , and Mehdi Gheisari 









Research Article (11 pages), Article ID 4239536, Volume 2022 (2022)

[Retracted] Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services

Abdelaziz Elfaki , Amani M AlQarni, Amal A. AlGhamdi, Malak A. AlShammari, Farheen Nasir, and Rana Alabdulqader






Research Article (6 pages), Article ID 6530019, Volume 2022 (2022)

[Retracted] Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers

Meng Ji , Wenxiu Xie , Mengdan Zhao , Xiaobo Qian , Chi-Yin Chow , Kam-Yiu Lam , Jun Yan , and Tianyong Hao 




Research Article (15 pages), Article ID 6722321, Volume 2022 (2022)

[Retracted] Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms

Sudeshna Chakraborty , Hussain Falih Mahdi , Mohammed Hasan Ali Al-Abyadh , Kumud Pant , Aditi Sharma , and Fardin Ahmadi 





Research Article (10 pages), Article ID 5731532, Volume 2022 (2022)

[Retracted] A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things

Zard Ali Khan, Sheneela Naz, Rahim khan , Jason Teo , Abdullah Ghani, and Mohammed Amin Almaiah 

Research Article (14 pages), Article ID 5112375, Volume 2022 (2022)

[Retracted] Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model

Asif Hasan , Tripti Sharma , Azizuddin Khan , and Mohammed Hasan Ali Al-Abyadh 







Research Article (8 pages), Article ID 8153791, Volume 2022 (2022)

[Retracted] Psychological Analysis for Depression Detection from Social Networking Sites

Sonam Gupta , Lipika Goel , Arjun Singh , Ajay Prasad , and Mohammad Aman Ullah 









Research Article (14 pages), Article ID 4395358, Volume 2022 (2022)

[Retracted] Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology

Abhinav Juneja , Hamza Turabieh , Hemant Upadhyay , Zelalem Kiros Bitsue , Vinh Truong Hoang , and Kiet Tran Trung 



Research Article (8 pages), Article ID 4602072, Volume 2022 (2022)

[Retracted] Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder

G. Murugesan , Tousief Irshad Ahmed , Mohammad Shabaz , Jyoti Bhola , Batyrkhan Omarov , R. Swaminathan , F. Sammy , and Sharmin Akter Sumi 

Research Article (10 pages), Article ID 8555489, Volume 2022 (2022)









[Retracted] A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment

Junjie Cen  and Yongbo Li 

Research Article (8 pages), Article ID 1354233, Volume 2022 (2022)

Contents



[Retracted] Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection

Mustafa Kamal , A. Raghuvira Pratap , Mohd Naved , Abu Sarwar Zamani , P. Nancy , Mahyudin Ritonga , Surendra Kumar Shukla , and F. Sammy 
Research Article (8 pages), Article ID 5261942, Volume 2022 (2022)

[Retracted] Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model

Syed Nasrullah  and Asadullah Jalali 
Research Article (6 pages), Article ID 9404242, Volume 2022 (2022)







[Retracted] Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview

Muhammad Faisal Manzoor , Abid Hussain, Diana Tazeddinova, Aizhan Abylgazinova, and Bin Xu 
Review Article (16 pages), Article ID 6531483, Volume 2022 (2022)





[Retracted] Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm

Jun Zhou  and Mei Yang 
Research Article (8 pages), Article ID 3975853, Volume 2022 (2022)



[Retracted] Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms

Shakir Khan , V. Saravanan , Gnanaprakasam C. N , T. Jaya Lakshmi , Nabamita Deb , and Nashwan Adnan Othman 
Research Article (8 pages), Article ID 9985933, Volume 2022 (2022)



[Retracted] Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System

Qingjiang Li , Xuejiao Chai , Chunqing Zhang , Xinjia Wang , and Wenhui Ma 
Research Article (10 pages), Article ID 8936103, Volume 2022 (2022)




[Retracted] Computational Intelligence Approaches in Developing Cyberattack Detection System

Mohammed Saeed Alzahrani  and Fawaz Waselallah Alsaade 
Research Article (16 pages), Article ID 4705325, Volume 2022 (2022)


[Retracted] A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1

Hulin Zhao, Xuetao Dong, Tianxiang Huang , and Xueji Li 
Research Article (8 pages), Article ID 9979200, Volume 2022 (2022)



[Retracted] Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies

Vinit Kumar Gunjan , Y. Vijayalata , Susmitha Valli , Sumit Kumar , M. O. Mohamed , and V. Saravanan 
Research Article (10 pages), Article ID 3604113, Volume 2022 (2022)




[Retracted] Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty

Xue Zhang, Yingmo Shen, Shuo Yang, Baoshan Wang, and Jie Chen 
Research Article (6 pages), Article ID 1054299, Volume 2022 (2022)






[Retracted] Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness

Ning An  and Jing Chuo 
Research Article (7 pages), Article ID 3440437, Volume 2022 (2022)







[Retracted] A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts

Gayathri Devi S, Subramaniaswamy Vairavasundaram, Yuvaraja Teekaraman , Ramya Kuppusamy , and Arun Radhakrishnan 
Research Article (15 pages), Article ID 3432330, Volume 2022 (2022)




[Retracted] Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media

Suneet Gupta , Sumit Kumar , Sunil L. Bangare , Shibili Nuhmani , Arnold C. Alguno , and Issah Abubakari Samori 
Research Article (9 pages), Article ID 3490860, Volume 2022 (2022)








[Retracted] A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model

Kamil Zeberga , Muhammad Attique , Babar Shah , Farman Ali , Yalew Zelalem Jembre , and Tae-Sun Chung 
Research Article (18 pages), Article ID 7893775, Volume 2022 (2022)






[Retracted] Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers

Anjali Goswami, Muddada Murali Krishna, Jayavani Vankara, Syam Machinathu Parambil Gangadharan , Chandra Shekhar Yadav, Manoj Kumar , and Mohammad Monirujjaman Khan 
Research Article (18 pages), Article ID 9194031, Volume 2022 (2022)

[Retracted] A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System


N. Arunkumar , V. Pandimurugan , M. S. Hema , H. Azath , S. Hariharasitaraman , M. Thilagaraj , and Petchinathan Govindan 
Research Article (11 pages), Article ID 9441357, Volume 2022 (2022)

[Retracted] An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease




Ahmad Ali AlZubi , Shailendra Tiwari , Kuldeep Walia , Jazem Mutared Alanazi , Firas Ibrahim AlZobi , and Rohit Verma
Research Article (9 pages), Article ID 2933015, Volume 2022 (2022)

Contents



[Retracted] Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Fawaz Waselallah Alsaade  and Mohammed Saeed Alzahrani 
Research Article (10 pages), Article ID 8709145, Volume 2022 (2022)

[Retracted] 5G Massive MIMO Signal Detection Algorithm Based on Deep Learning

Lichao Yan , Yi Wang , and Ning Zheng 
Research Article (9 pages), Article ID 9999951, Volume 2022 (2022)









[Retracted] Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System

B. Sumathy, Anand Kumar, D. Sungeetha, Arshad Hashmi, Ankur Saxena , Piyush Kumar Shukla, and Stephen Jeswinde Nuagah 
Research Article (10 pages), Article ID 5906797, Volume 2022 (2022)


[Retracted] Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model

M Thilagaraj , N. Arunkumar , and Petchinathan Govindan 
Research Article (12 pages), Article ID 6785707, Volume 2022 (2022)



[Retracted] Evolving Long Short-Term Memory Network-Based Text Classification

Arjun Singh , Shashi Kant Dargar , Amit Gupta , Ashish Kumar , Atul Kumar Srivastava ,
Mitali Srivastava , Pradeep Kumar Tiwari , and Mohammad Aman Ullah 
Research Article (11 pages), Article ID 4725639, Volume 2022 (2022)











[Retracted] Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques

Qiqian Li, Yali Liu, Esra Sipahi Döngül, Yufen Yang, Xiaoyuan Ruan, and Wegayehu Enbeyle 
Research Article (14 pages), Article ID 2290644, Volume 2022 (2022)

[Retracted] An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding

Ahmad Alsharif, Karan Aggarwal, Sonia , Deepika Koundal, Hashem Alyami, and Darine Ameyed 
Research Article (8 pages), Article ID 8467349, Volume 2022 (2022)

[Retracted] Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain

Divakar Yadav , Naman Lalit , Riya Kaushik , Yogendra Singh , Mohit , Dinesh , Arun Kr. Yadav , Kishor V. Bhadane , Adarsh Kumar , and Baseem Khan 
Research Article (14 pages), Article ID 3411881, Volume 2022 (2022)

[Retracted] Learning Enhanced Feature Responses for Visual Object Tracking

Runqing Zhang , Chunxiao Fan , and Yue Ming
Research Article (15 pages), Article ID 1241687, Volume 2022 (2022)

[Retracted] Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection

N. Krishnaraj , B. Sivakumar, Ramya Kuppusamy , Yuvaraja Teekaraman , and Amruth Ramesh Thelkar 



Research Article (13 pages), Article ID 8501738, Volume 2022 (2022)

[Retracted] Applying Dynamic Systems to Social Media by Using Controlling Stability

Abdulsattar Abdullah Hamad, M. Lellis Thivagar , Malik Bader Alazzam , Fawaz Alassery , Fahima Hajje , and Ali A. Shihab







Research Article (7 pages), Article ID 4569879, Volume 2022 (2022)

[Retracted] TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis

Li Lin , Suqing Yang, Qingsong Bai, and Yuepeng An 

Research Article (7 pages), Article ID 4268681, Volume 2022 (2022)

[Retracted] DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning

Atef Zaguia , Deepak Pandey , Sandeep Painuly , Saurabh Kumar Pal , Vivek Kumar Garg , and Neelam Goel 

Research Article (11 pages), Article ID 8393498, Volume 2022 (2022)

[Retracted] Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility

Malik Bader Alazzam , Ahmad Tawfig Al-Radaideh , Nasser Binsaiif , Ahmed S. AlGhamdi , and Md Adnan Rahman 




Research Article (5 pages), Article ID 1422963, Volume 2022 (2022)

[Retracted] Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App

Malik Bader Alazzam , Hoda Mansour , Fawaz Alassery , and Ahmed Almulihi 





Research Article (7 pages), Article ID 5759184, Volume 2021 (2021)

[Retracted] A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization

Ahmed S. Al-Obeidi, Saad Fawzi Al-Azzawi, Abdulsattar Abdullah Hamad , M. Lellis Thivagar , Zelalem Meraf , and Sultan Ahmad 

Research Article (11 pages), Article ID 3081345, Volume 2021 (2021)

[Retracted] Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives

Khalid Twarish Alhamazani , Jalawi Alshudukhi , Saud Aljaloud , and Solomon Abebaw 

Research Article (8 pages), Article ID 3941978, Volume 2021 (2021)









Contents

[Retracted] Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP

Raghavendra Rao Althar , Debabrata Samanta , Manjit Kaur , Abeer Ali Alnuaim , Nouf Aljaffan, and Mohammad Aman Ullah 

Research Article (19 pages), Article ID 8522839, Volume 2021 (2021)

[Retracted] Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpening of Remote Sensing Images

Ahmed I. Iskanderani , Ibrahim M. Mehedi , Abdulah Jeza Aljohani , Mohammad Shorfuzzaman , Farzana Akhter, Thangam Palaniswamy , Shaikh Abdul Latif , Abdul Latif , and Rahtul Jannat 



Research Article (11 pages), Article ID 7615106, Volume 2021 (2021)

[Retracted] Secure Complex Systems: A Dynamic Model in the Synchronization

Abdulsattar Abdullah Hamad , M. Lellis Thivagar , Jalawi Alshudukhi, Talal Saad Alharbi , Saud Aljaloud , Khalid Twarish Alhamazani , and Zelalem Meraf 



Research Article (6 pages), Article ID 9719413, Volume 2021 (2021)

[Retracted] Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements

Malik Bader Alazzam , Hoda Mansour, Mohamed M. Hammam, Said Alsheikh, Ali Bakir, Saeed Alghamdi, and Ahmed S. AlGhamdi 







Research Article (6 pages), Article ID 1094054, Volume 2021 (2021)

[Retracted] Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy

Jamal Alshorman , Yulong Wang, Guixiong Huang, Tracy Boakye Serebour, and Xiaodong Guo 

Research Article (6 pages), Article ID 8560668, Volume 2021 (2021)

[Retracted] Human-Computer Interaction Using Manual Hand Gestures in Real Time

Mohammad Alsaffar , Abdullah Alshammari , Gharbi Alshammari , Tariq S Almurayziq , Saud Aljaloud , Dhahi Alshammari, and Assaye Belay 

Research Article (5 pages), Article ID 6972192, Volume 2021 (2021)

Retraction

Retracted: A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment

Computational Intelligence and Neuroscience

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 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] J. Cen and Y. Li, "A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1354233, 8 pages, 2022.

Retraction

Retracted: Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] W. Yang, X. Pan, Y. Wang, and W. Chen, "Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6861749, 7 pages, 2022.

Retraction

Retracted: A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] K. Zeberga, M. Attique, B. Shah, F. Ali, Y. Z. Jembre, and T. Chung, "A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7893775, 18 pages, 2022.

Retraction

Retracted: Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] A. Goswami, M. M. Krishna, J. Vankara et al., "Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9194031, 18 pages, 2022.

Retraction

Retracted: Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] M. Ji, W. Xie, M. Zhao et al., "Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6722321, 15 pages, 2022.

Retraction

Retracted: Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] A. Juneja, H. Turabieh, H. Upadhyay, Z. K. Bitsue, V. T. Hoang, and K. T. Trung, "Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4602072, 8 pages, 2022.

Retraction

Retracted: Psychological Analysis for Depression Detection from Social Networking Sites

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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. Gupta, L. Goel, A. Singh, A. Prasad, and M. A. Ullah, "Psychological Analysis for Depression Detection from Social Networking Sites," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4395358, 14 pages, 2022.

Retraction

Retracted: Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Chakraborty, H. F. Mahdi, M. H. A. Al-Abyadh, K. Pant, A. Sharma, and F. Ahmadi, "Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5731532, 10 pages, 2022.

Retraction

Retracted: Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] J. Srikanth, A. Damodaram, Y. Teekaraman, R. Kuppasamy, and A. R. Thelkar, "Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8898100, 11 pages, 2022.

Retraction

Retracted: Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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. Nasrullah and A. Jalali, "Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9404242, 6 pages, 2022.

Retraction

Retracted: Identification and Classification of Depressed Mental State for End-User over Social Media

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] A. Kumar, A. Thakare, M. Bhende, A. K. Sinha, A. C. Alguno, and Y. P. Kumar, "Identification and Classification of Depressed Mental State for End-User over Social Media," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8755922, 10 pages, 2022.

Retraction

Retracted: Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] T. Saba, A. R. Khan, I. Abunadi, S. A. Bahaj, H. Ali, and M. Alruwaythi, "Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8622022, 9 pages, 2022.

Retraction

Retracted: Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] J. Alshorman, Y. Wang, G. Huang, T. B. Serebour, and X. Guo, "Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8560668, 6 pages, 2021.

Retraction

Retracted: Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] F. W. Alsaade and M. S. Alzahrani, "Classification and Detection of autism Spectrum Disorder Based on Deep Learning Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8709145, 10 pages, 2022.

Retraction

Retracted: Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] G. Murugesan, T. I. Ahmed, M. Shabaz et al., "Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8555489, 10 pages, 2022.

Retraction

Retracted: Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] K. Twarish Alhamazani, J. Alshudukhi, S. Aljaloud, and S. Abebaw, "Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3941978, 8 pages, 2021.

Retraction

Retracted: Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] N. Krishnaraj, B. Sivakumar, R. Kuppasamy, Y. Teekaraman, and A. R. Thelkar, "Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8501738, 13 pages, 2022.

Retraction

Retracted: Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] N. An and J. Chuo, "Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3440437, 7 pages, 2022.

Retraction

Retracted: A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] H. Zhao, X. Dong, T. Huang, and X. Li, "A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9979200, 8 pages, 2022.

Retraction

Retracted: Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] A. Elfaki, A. M. AlQarni, A. A. AlGhamdi, M. A. AlShammari, F. Nasir, and R. Alabdulqader, "Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6530019, 6 pages, 2022.

Retraction

Retracted: Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] B. Sumathy, A. Kumar, D. Sungeetha et al., "Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5906797, 10 pages, 2022.

Retraction

Retracted: Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] A. Hasan, T. Sharma, A. Khan, and M. Hasan Ali Al-Abyadh, "Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8153791, 8 pages, 2022.

Retraction

Retracted: Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] V. K. Gunjan, Y. Vijayalata, S. Valli, S. Kumar, M. O. Mohamed, and V. Saravanan, "Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3604113, 10 pages, 2022.

Retraction

Retracted: Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulih, "Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5759184, 7 pages, 2021.

Retraction

Retracted: Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview

Computational Intelligence and Neuroscience

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 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] M. F. Manzoor, A. Hussain, D. Tazeddinova, A. Abylgazinova, and B. Xu, "Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6531483, 16 pages, 2022.

Retraction

Retracted: Human-Computer Interaction Using Manual Hand Gestures in Real Time

Computational Intelligence and Neuroscience

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 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] M. Alsaffar, A. Alshammari, G. Alshammari et al., "Human-Computer Interaction Using Manual Hand Gestures in Real Time," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6972192, 5 pages, 2021.

Retraction

Retracted: Evolving Long Short-Term Memory Network-Based Text Classification

Computational Intelligence and Neuroscience

Received 13 September 2023; Accepted 13 September 2023; Published 14 September 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] A. Singh, S. K. Dargar, A. Gupta et al., "Evolving Long Short-Term Memory Network-Based Text Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4725639, 11 pages, 2022.

Retraction

Retracted: Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System

Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 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] Q. Li, X. Chai, C. Zhang, X. Wang, and W. Ma, "Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8936103, 10 pages, 2022.

Retraction

Retracted: Secure Complex Systems: A Dynamic Model in the Synchronization

Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 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] A. A. Hamad, M. L. Thivagar, J. Alshudukhi et al., "Secure Complex Systems: A Dynamic Model in the Synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9719413, 6 pages, 2021.

Retraction

Retracted: Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] D. Yadav, N. Lalit, R. Kaushik et al., "Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3411881, 14 pages, 2022.

Retraction

Retracted: A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] G. Devi S, S. Vairavasundaram, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3432330, 15 pages, 2022.

Retraction

Retracted: Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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. Khan, V. Saravanan, C. N. Gnanaprakasam, T. J. Lakshmi, N. Deb, and N. A. Othman, "Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9985933, 8 pages, 2022.

Retraction

Retracted: Learning Enhanced Feature Responses for Visual Object Tracking

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] R. Zhang, C. Fan, and Y. Ming, "Learning Enhanced Feature Responses for Visual Object Tracking," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1241687, 15 pages, 2022.

Retraction

Retracted: Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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. Gupta, S. Kumar, S. L. Bangare, S. Nuhmani, A. C. Alguno, and I. A. Samori, "Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3490860, 9 pages, 2022.

Retraction

Retracted: Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Bader Alazzam, H. Mansour, M. M. Hammam et al., "Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1094054, 6 pages, 2021.

Retraction

Retracted: Computational Intelligence Approaches in Developing Cyberattack Detection System

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. S. Alzahrani and F. W. Alsaade, "Computational Intelligence Approaches in Developing Cyberattack Detection System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4705325, 16 pages, 2022.

Retraction

Retracted: DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Zaguia, D. Pandey, S. Painuly, S. K. Pal, V. K. Garg, and N. Goel, "DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8393498, 11 pages, 2022.

Retraction

Retracted: A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. S. Al-Obeidi, S. Fawzi Al-Azzawi, A. Abdullah Hamad, M. L. Thivagar, Z. Meraf, and S. Ahmad, "A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3081345, 11 pages, 2021.

Retraction

Retracted: Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Kamal, A. R. Pratap, M. Naved et al., "Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5261942, 8 pages, 2022.

Retraction

Retracted: Stress-Relieving Video Game and Its Effects: A POMS Case Study

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] A. Ajmal, H. Aldabbas, R. Amin, S. Ibrar, B. Alouffi, and M. Gheisari, "Stress-Relieving Video Game and Its Effects: A POMS Case Study," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4239536, 11 pages, 2022.

Retraction

Retracted: A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] Z. A. Khan, S. Naz, R. Khan, J. Teo, A. Ghani, and M. A. Almaiah, "A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5112375, 14 pages, 2022.

Retraction

Retracted: 5G Massive MIMO Signal Detection Algorithm Based on Deep Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] L. Yan, Y. Wang, and N. Zheng, "5G Massive MIMO Signal Detection Algorithm Based on Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9999951, 9 pages, 2022.

Retraction

Retracted: Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Londhe, P. V. R. D. P. Rao, S. Upadhyay, and R. Jain, "Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8579640, 14 pages, 2022.

Retraction

Retracted: An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. A. AlZubi, S. Tiwari, K. Walia, J. M. Alanazi, F. I. AlZobi, and R. Verma, "An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2933015, 9 pages, 2022.

Retraction

Retracted: Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] J. Zhou and M. Yang, "Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3975853, 8 pages, 2022.

Retraction

Retracted: A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] N. Arunkumar, V. Pandimurugan, M. S. Hema et al., "A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9441357, 11 pages, 2022.

Retraction

Retracted: Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Thilagaraj, N. Arunkumar, and P. Govindan, "Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6785707, 12 pages, 2022.

Retraction

Retracted: Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] T. Veeramakali, A. Shobanadevi, N. R. Nayak, S. Kumar, S. Singhal, and M. Subramanian, "Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4690936, 8 pages, 2022.

Retraction

Retracted: Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] Q. Li, Y. Liu, E. S. Sipahi Döngül, Y. Yang, X. Ruan, and W. Enbeyle, "Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2290644, 14 pages, 2022.

Retraction

Retracted: Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. B. Alazzam, A. T. Al-Radaideh, N. Binsaif, A. S. AlGhamdi, and M. A. Rahman, "Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1422963, 5 pages, 2022.

Retraction

Retracted: Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] R. R. Althar, D. Samanta, M. Kaur, A. A. Alnuaim, N. Aljaffan, and M. Aman Ullah, "Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8522839, 19 pages, 2021.

Retraction

Retracted: TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] L. Lin, S. Yang, Q. Bai, and Y. An, "TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4268681, 7 pages, 2022.

Retraction

Retracted: Applying Dynamic Systems to Social Media by Using Controlling Stability

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajje, and A. A. Shihab, "Applying Dynamic Systems to Social Media by Using Controlling Stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

Retraction

Retracted: Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpener of Remote Sensing Images

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. I. Iskanderani, I. M. Mehedi, A. J. Aljohani et al., "Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpener of Remote Sensing Images," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7615106, 11 pages, 2021.

Research Article

Hybrid Recommender System for Mental Illness Detection in Social Media Using Deep Learning Techniques

Sayed Sayeed Ahmad ¹, Rashmi Rani, ¹ Ihab Wattar ², Meghna Sharma,³ Sanjiv Sharma,⁴ Rajit Nair ⁵, and Basant Tiwari ⁶

¹College of Engineering and Computing, Al Ghurair University, Dubai, UAE

²Electrical Engineering and Computer Science Department, Cleveland State University, Cleveland, USA

³Department of Computer Science and Engineering, The NorthCap University, Gurugram, India

⁴Department of Computer Science and Engineering, KIET Group of Institutions, Delhi-NCR, Meerut Road, India

⁵School of Computing Science & Engineering, VIT Bhopal University, Bhopal-Indore Highway Kothrikalan, Bhopal, MP, India

⁶Department of Computer Science, Ethiopia Hawassa University, Awasa, Ethiopia

Correspondence should be addressed to Basant Tiwari; basanttiw@hu.edu.et

Received 1 November 2021; Revised 19 January 2022; Accepted 19 April 2022; Published 8 July 2023

Academic Editor: Sheng Du

Copyright © 2023 Sayed Sayeed Ahmad 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.

Recommender systems are chiefly renowned for their applicability in e-commerce sites and social media. For system optimization, this work introduces a method of behaviour pattern mining to analyze the person's mental stability. With the utilization of the sequential pattern mining algorithm, efficient extraction of frequent patterns from the database is achieved. A candidate sub-sequence generation-and-test method is adopted in conventional sequential mining algorithms like the Generalized Sequential Pattern Algorithm (GSP). However, since this approach will yield a huge candidate set, it is not ideal when a large amount of data is involved from the social media analysis. Since the data is composed of numerous features, all of which may not have any relation with one another, the utilization of feature selection helps remove unrelated features from the data with minimal information loss. In this work, Frequent Pattern (FP) mining operations will employ the Systolic tree. The systolic tree-based reconfigurable architecture will offer various benefits such as high throughput as well as cost-effective performance. The database's frequently occurring item sets can be found by using the FP mining algorithms. Numerous research areas related to machine learning and data mining are fascinated by feature selection since it will enable the classifiers to be swift, more accurate, and cost-effective. Over the last ten years or so, there have been significant technological advancements in heuristic techniques. These techniques are beneficial because they improve the search procedure's efficiency, albeit at the potential sacrifice of completeness claims. A new recommender system for mental illness detection was based on features selected using River Formation Dynamics (RFD), Particle Swarm Optimization (PSO), and hybrid RFD-PSO algorithm is proposed in this paper. The experiments use the depressive patient datasets for evaluation, and the results demonstrate the improved performance of the proposed technique.

1. Introduction

Because of the deluge of information available on the Internet, people have turned to a variety of strategies to help them make a variety of choices, including who to go out with, which phone to purchase, and where to spend their vacation. User-centric recommender systems provide excellent suggestions to users while engaging with massive amounts of information. Using social media,

these systems may provide suggestions on everything from music to books to news to depressed patients to Web sites to even more complicated recommendations in social media for financial services, electrical gadgets, and so on. The majority of recommendation algorithms are based on a variety of filtering approaches, including Collaborative Filtering (CF) and Content-Based Filtering (CBF) (CBF). Many studies have been carried out in the area of recommender systems over the last 10 years in order to

develop unique algorithms that would improve the accuracy of suggestion [1].

The following points will define the essential operational as well as technological objectives of a recommender system in detail: One of the most important characteristics of a recommendation system is its effort to discover the most relevant items for each query. While this is the fundamental purpose of a recommender system, it also has a number of secondary objectives. 2. Newness: Human beings are always on the lookout for novel experiences. As a result, the Recommender system will not be successful if it just recommends products that are outdated yet still popular among the users. 3. Surprising yourself: This objective is all about being startled. To distinguish between serendipity and novelty, we may use the following example to illustrate our point: Consider the following scenario: you are using a depressed patient recommender system. If the system has suggested a new depressed patient from your chosen genre, it has met the aim of providing a fresh experience. However, if the algorithm has suggested a popular depressed patient, even if it is from a different genre than the one you previously liked viewing, it has achieved the serendipity objective. 4. Diverse suggestions: Recommender systems will often provide the user with a selection of things to choose from. As a result, ensuring that the objects are as diverse since possible is an important aim, as it increases the likelihood that a user will choose at least one of them [2].

The classification of recommender systems will often comprise the following categories: CF, CBF, Demographic-based, and Hybrid recommender systems. CF is founded on the assumption that persons who have provided permission would continue to do so in the future. Similar users are recognised based on previous information of the user's activity, and things will be recommended based on the behaviour of similar users, as well as the behaviour of the user. The specs of things are taken into consideration by the CBF Recommender systems, and the suggested items are quite comparable to the items that the user previously liked. When it comes to social media, the difficulty with content-based filtering is that it cannot provide good suggestions if the material does not have sufficient information to differentiate between the things. When searching for comparable users to provide their suggestions in social media, demographic-based recommender systems will take use of the demographic data of the users (for example, age, gender, and occupation) to make their recommendations. Via order to make suggestions in social media, Hybrid Recommender Systems will integrate two or more of the strategies described above. In terms of result potential, the CF approaches are the most promising of all of the system types discussed above. In spite of this, these methodologies have a number of shortcomings including sparsity, a slow start, a lack of tailored suggestions in social media, and the inability to generate context-aware recommendations in social media [3].

It is the process of obtaining meaningful patterns or information from large datasets that is referred to as "data mining." The mining of useful as well as practicable patterns from large databases is critical for a variety of data mining

activities, such as pattern mining of the frequently occurring itemset in large transactional databases. Due to the amount of time necessary for completing numerous database scans and providing additional candidate itemsets for the larger dataset, it has significant drawbacks, particularly with large datasets. In order to tackle these challenges, a new growth algorithm known as FP growth must be developed. By constructing the prefix-tree without taking into account the outcomes, this approach will reduce the number of steps required to complete the task. Although it has some advantages, it has two significant drawbacks: it will consider all of the items as being comparable, and it will display every item in the transaction database in binary (0/1) form, which means that it will be either current or deficient [4].

Data mining uses feature selection procedures to automatically identify the features in a dataset that are relevant for the specified prediction model while minimising the risk of overfitting. The viability of the feature selection process is due to its capacity to remove redundant or unnecessary qualities that either make no contribution to the predictive model's accuracy or end up reducing the accuracy of the predictive model, respectively. The following three points will serve as the goals for the feature selection process: (1) to improve the predictors' ability to forecast, (2) to provide predictors that are both quicker and more cost-effective, and (3) to provide a better knowledge of the data creation technique [5].

A smart strategy for feature selection will remove characteristics that are of little or no use in terms of adding information to the database. Three general metrics are used in the feature selection process. Filters are the first form of measure, and they will apply a statistical measure to each characteristic in order to award a score to it. The feature selection approach will be used as a preprocessing step for these metrics, and it will be independent of the learning process [6]. When using wrappers as a second form of measure, the task will be modelled as a search problem, and the learning system will be used as a black box for scoring the feature subsets. The third sort of measure is embedded techniques, which will be used to carry out the selection process while the training procedure is being carried out. When it comes to recommender systems, the kind of feature selection is determined by the aim of making the suggestion.

The optimization of the feature selection is accomplished via the use of metaheuristic approaches since the feature selection is an NP-hard issue. Swarm intelligence techniques [7] are built on a collection of basic entities that will interact with one another depending on the knowledge available to them in their immediate environment. The goal of these interactions is to work together to find a suitable solution to a specific issue that has been identified. Different swarm intelligence metaheuristics [8] for discrete combinatorial optimization issues (such as RFD or Ant Colony Optimization (ACO)) as well as problems of continuous domain optimization (such as PSO, Artificial Bee Colony (ABC)) have been proposed in different research articles.

In a nutshell, the RFD is a water-based metaheuristic that will reproduce the geological processes that would result in the formation of the river. The RFD is particularly well-

suitable for NP-hard issues involving the construction of a specific tree type since the two aforementioned inclinations may be readily bent towards either of the two directions via the use of parameterization. It is possible to handle a wide range of conventional NP-hard optimization issues using RFD applications. Furthermore, the RFD has been used to solve industrial challenges such as network routing, optimization in electrical power systems, and VLSI design, to name a few. It is somewhat noteworthy that the RFD is believed to be a derivative-oriented form of the ACO, at least in broad terms. In the ACO, entities (ants) have a tendency to gravitate toward nodes with specific higher values than others (for example, the pheromone trail). As a result of this, the RFD shows that the drops tend to move towards nodes where there is a greater difference between the values (altitudes) at the origin and destination nodes than at other nodes (steeper slopes will have a bigger flow).

Firstly, it has been found that the informal nature of tweets is crucial for the classification of feelings. Based on the tweets, the mental illness of the person has been classified. Therefore, to categorise Indian language tweets is proposed a combination of grammar rules based on adjectives and negations. This type of categorization is unique and has a good way of explanation.

This work proposes a hybrid RFD-PSO algorithm for recommender system-based pattern mining was proposed. The rest of the paper presents the related works in literature, different techniques used in the work, experimental results, and conclusion.

2. Related Works

Cai et al. [9] proposed a rating-based many-objective hybrid recommendation method that could concurrently optimize the recommendation's coverage, novelty, diversity, recall, and accuracy. In addition, there were proposals of a novel strategy for generation-based fitness evaluation as well as a strategy for partition-based knowledge mining. These strategies would boost the Many-Objective Evolutionary Algorithms (MaOEOs) for performance improvement of the model's generated recommendations in social media. Eventually, upon comparison with the existing conventional MaOEOs, the experimental outcomes were able to demonstrate that the proposed algorithm could offer recommendations in social media having novel as well as more number of items in terms of the users' accuracy and diversity.

Alhijawi and Kilani [10] had presented a Genetic Algorithm (GA)-based recommender system (BLIGA), which was dependent on the historical rating and semantic information. Rather than assessing the items before the formation of the recommendation list, this research's key contribution involved the assessment of the potential recommendation lists. In the BLIGA, there was the utilization of the GA for identifying the most relevant items for the user. Hence, every individual was a representation of the candidate recommendation list. The BLIGA has employed three distinct fitness functions to hierarchically assess the individuals. A comparison of the recommendation results was done between the BLIGA and other CF methods. It was

evident from the results that the BLIGA was much superior and was able to accomplish highly accurate predictions regardless of the number of K-neighbors.

Alhijawi et al. [11] presented three distinct novel GA-based Recommender systems: GARS+, GARS++, and HGARS to address the issue of offering users item recommendations in social media. As a combination of GARS+ and GARS++, HGARS was the genetic-based recommender system's enhanced version which had worked without being a hybrid model. The proposed recommender system employed GA in its search for the optimal similarity function, which in turn, was dependent on a linear combination of values as well as weights. Using experimentations, the authors were able to confirm that HGARS was able to accomplish improvements of 16.1% inaccuracy, 17.2% in the recommendation quality, and 40% in performance.

Rakshana Sri et al. [12] had devised a system that executed user cluster-based CF for venue recommendations in social media; wherein there was utilization of a bio-inspired Grey Wolf Optimization (GWO) algorithm for the cluster formation. With the clustering's utilization, there was the removal of the CF's shortcomings inaccuracy, sparsity as well as scalability. Moreover, the authors used the cosine similarity and the Pearson Correlation Coefficient (PCC) for identifying similar users. Performance evaluation was carried out using Trip Advisor and Yelp datasets to determine metrics such as accuracy, precision, recall as well as f-measure. The outcomes of the experimentation, as well as the evaluation, were able to show the efficiency of newly generated recommendations in social media and also had displayed user satisfaction.

Tohidi and Dadkhah [13] had introduced an approach for increasing the accuracy and boosting the performance of a CF Recommender system. This work had put forward a hybrid approach to boost the video CF Recommender system's performance based on the clustering and the evolutionary algorithm. The proposed approach was a combination of the k-means clustering algorithm with two metaheuristics: the Accelerated PSO (APSO), and the Forest Optimization Algorithm (FOA). This work's key objective involved increasing the user-based CF video Recommender system's recommendation accuracy. Evaluation, as well as computational outcomes on the Depressive patient dataset, had shown the proposed approach's superior performance over the other related methods.

El-Ashmawi et al. [14] had devised a novel algorithm for the detection of a feasible cluster set of similar users to boost the procedure of recommendation. Utilization of the genetic uniform crossover operator in the conventional Crow Search Algorithm (CSA) was able to increase the search's diversity as well as to aid the algorithm in avoiding capture in the local minima. There was the presentation of the top-N recommendations in social media based on the feasible cluster's members. The Jester dataset was used for the evaluation of the proposed algorithm's performance. It was indicated from the results that the proposed algorithm was able to attain superior results with regards to the mean absolute error, the root means square errors as well as the objective function's minimization.

Wang et al. [15] had examined a novel bacterial colony-based feature selection algorithm with an attribute learning strategy [16] for obtaining personalized product recommendations in social media. In specific terms, the features were weighted following their historic contributions to the individual-based as well as the group-based subsets. Furthermore, the feature candidates' occurrence frequency was recorded for improvement of the feature distribution's diversity and avoidance of overfitting. With regards to the weight-based feature indexes as well as the occurrence frequency records, performance enhancement of these feature subsets was achieved through the replacement of features that has repeatedly appeared within the same vector. The optimization's objective involved minimization of the classification error using the acceptable number of features. There was the utilization of the KNN as a learning method for cooperation with the proposed feature selection algorithm. Upon comparison with seven different feature selection methods, the proposed algorithm's superior performance was evident from its accomplishment of a higher rate of classification accuracy with the utilization of a smaller number of features. Table 1 has represented the comparison of the existing methodology with different methods.

3. Methodology

In the proposed hybrid recommender system for mental illness detection in social media, the features are extracted from the transactions, feature selection is applied, and a systolic tree is used for frequent pattern mining. Various dataset has been included and experiment is carried out using Depressive patient-Lens dataset that is used for evaluating the techniques. This section discusses the systolic tree, TF-IDF feature extraction method, RFD, PSO, and hybrid RFD-PSO-based feature selection methods.

In order to test both classifiers, the model includes a variety of processes, including the SVM classifier and the Nave Bayes classifier. It comprises of two datasets and seven primary operators, which are described below. In the first dataset, which is called the training dataset, there are 2073 sad posts and 2073 non-depressed posts that have been manually trained. In addition, it is divided into three columns: the first is a binominal sentiment (Depressed or Not-Depressed), the second is a depression category (in the event of depressed sentiment, one of the nine categories), and the third is the trained post (if applicable). The second dataset consists of the patient SNS posts, and it is different for each person in order to evaluate the model's prediction.

In the Select Characteristics operator, the user may specify whether or not certain attributes from the training dataset should be retained and which attributes should be eliminated from the training dataset. The second and third operators are the Nominal to Text operators, which transform the type of chosen nominal attributes to text and also map all values of the attributes to their corresponding string values. This operator is used in both the training dataset and the test dataset. The fourth and fifth processes are Process Documents, and they are used in both the training dataset

and the test set to create word vectors from string characteristics. They are composed of four operators and are utilized in both the training dataset and the test set.

Process Document operator has four operators: Tokenize, Filter Stop-words, Transform Cases, and Stem. Tokenize is the first of these operators. With the Tokenize operators, you may break down the text of a document into a series of tokens. A text is filtered for English stopwords using the Stop-words filter, which removes every token in the document that matches a stopword from the built-in stopword list in the RapidMiner.

The Transform Cases operation converts all of the characters in a document to lower-case lettering. The Porter stemming method is used by the Stem operator to stem English words, with the goal of reducing the length of the words until a minimal length is attained by the operator. The sixth operator is the Validation operator, which applies to the training dataset, which is divided into two sections: training and testing. The Validation operator has two parts: training and testing. The classifier operator is included in the training part, and we change the classifier model from SVM (Linear) to Nave Bayes Classifier (Kernel) for each patient we test in the training phase. The testing stage comprises of two operators: the Apply Model operator, which applies the trained model to the supervised dataset, and the Performance operator, which is used to evaluate the model's performance. In the seventh and final operator, we have the Apply model. This operator connects the test dataset with the training dataset in order to provide us with the final prediction result utilising one of the classifiers in the patients.

The accuracy of the classification is dependent on the training set that was used to train the classifier and to execute it. Rather than selecting simply apparent instances of a class, it is critical to choose sample training nodes that reflect edge cases that belong in or out of a class. As a result, it is a best practice to include as many different types of samples as feasible in the training set. This has been accomplished via the collection, organisation, and manual training of a supervised dataset. The postings for the dataset were gathered from three social media platforms: Facebook, LiveJournal, and Twitter. The dataset was manually trained to identify two types of sentiment: depressed and not depressed. In the case of depressed sentiment, we classified the depressed post into one of the nine depression symptoms defined by the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). To conclude, we have 6773 posts in the training dataset, 2073 of which are trained as depressed posts and 4700 of which are not trained as depressed posts.

3.1. Dataset. The following information is included inside the dataset: There are two folders, with one containing the data for the controls and the other containing the data for the condition group. We give a csv file with the actigraph data that has been gathered throughout time for each patient. Each of the following columns contains information: timestamps (one-minute intervals), date (date of

TABLE 1: Comparison of existing methodology.

Method	Description	References
Unified relevance model	It is a probabilistic item-to-user relevance framework that uses the parzen-window approach to estimate the density of relevant items. This strategy helps to alleviate the issue of data sparsity.	Si and Jin [17]
Hybrid CF model	Effective recommender systems are introduced, which make use of sequential mixture CF and joint mixture CF to achieve their results. It also incorporates sophisticated bayes belief theory.	Su et al. [18]
Fuzzy association rules and multilevel similarity (FARAMS)	It makes use of fuzzy association rule mining in order to expand the capabilities of the current methodologies. It was possible for FARAMS to complete the goal of producing higher qualitative forecasts.	Wang et al. [19]
Flexible mixture model (FMM)	The formation of user and item clusters might happen at the same time. It adds preference nodes in order to investigate a significant variance in rating among users who have similar preferences.	Leung et al. [20]
Maximum entropy approach	In order to lower the apriori likelihood of an item, it is clustered depending on the user's access route. This is beneficial in dealing with sparsity and dimensionality.	Pavlov and Pennock [21]

measurement), and activity (activity measurement from the actigraph watch). In addition, we supply the MADRS scores in the file `emphscores.csv`, which may be seen below. There are nine columns in this table: number (patient identifier), days (numbers of days of measurements), gender (1 or 2 for female or male), age (age in age groups), afftype (1: bipolar II, 2: unipolar depressive, 3: bipolar I), melanch (1: melancholia, 2: no melancholia), inpatient (1: inpatient, 2: outpatient), marriage (1: married or cohabiting, 2: single), and work (1: MADRS when measurement stopped).

3.2. Mental Illness Detection Based on Systolic Tree. The systolic tree structure is utilized for frequent pattern mining. In the VLSI terminology, it will refer to an assembly of pipelined Processing Elements (PEs) in a multidimensional tree pattern. Its configuration will store the candidate patterns' support counts in a pipelined manner. For a given transactional database, the relative positions of the systolic tree's elements must be similar to that of the FP tree. The transaction items, the Web page request sequence, will be updated into the systolic tree using operations like candidate item matching and count update [22] and the flow of the proposed Mental Illness Detection Based on Systolic Tree detection method. The sample dataset are collected from depressive patient lens from social media to analyze the person's mental illness from their recommender system.

The following PEs will constitute the structure of a systolic tree:

- (1) PE is under control. The root PE of the systolic tree will not contain any items. It is required that all data be entered via it. There is a link between one of its interfaces and the kid on the left side of the tree.
- (2) Physical education in general. The other PEs are referred to as "generic PEs" for the most part. There will be just one bidirectional interface on every generic PE, which will be connected to its parent. In contrast to the general PE with children, which will have a single interface that is connected to its leftmost kid, the general PE with siblings may have an

interface that is connected to its leftmost sibling. They may be used to create an item as well as increase the support count of the stored item.

- (3) Every PE will be assigned a level that corresponds to it. While the control PE is at level 0, the level of the general PE is determined by the distance between the general PE and the control PE. In a physical education class, all of the students will be at the same level. Every generic PE will have just a single parent who will have a direct relationship to the leftmost kid of the PE hierarchy. Because of their left siblings, the other children will be able to establish a secondary connection with their parents.
- (4) Among the PE's three operating modes are the WRITE mode, the SCAN mode, and the COUNT mode, all of which are described below. The WRITE mode is used to create a systolic tree and to control the flow of things within it. Counting the number of times a candidate itemset has been supported may be done in both the SCAN and the COUNT modes. Candidate itemset matching is the phrase used to describe this operation.

3.3. Term Frequency and Inverse Document Frequency (TF-IDF). The most popularly employed weighting metric is the Term Frequency and Inverse Document Frequency (TF-IDF) to quantify the relationship of words and instances. This measure will take into account the word or TF in the instance and also the word's uniqueness or how infrequent (IDF) it is in the whole corpus. Thus, the TF-IDF will allocate higher values to topic representative words while devaluing the common words. The TF-IDF has multiple variations [23]. Equation (1) will define the TF-IDF weighted value $w_{t,d}$ of the word t in the instance d as follows:

$$w_{t,d} = tf_{t,d} \times \log_{10} \left(\frac{N}{df_t} \right). \quad (1)$$

This equation $tf_{t,d}$ will denote the term frequency, N will denote the total instances in the corpus, and df will denote

the number of instances that have an occurrence of the word t .

COVID-19 may only occur once in a lifetime, but the experience of coping with such circumstances is still necessary. Although some nations have effectively controlled the pandemic, others have failed miserably in their attempts to deal with the problem as it has arisen. Because of the times we live in, it is extremely common for social media to play a significant part in our daily life. Social media is present everywhere, and everyone is either directly or indirectly linked to it. When faced with a pandemic, the government has implemented new measures (stay at home and social isolation), as well as placing limits on the mobility of individuals. It would have been preferable if social media networks had provided us with appropriate guidance in this dreadful scenario. Contrary to expectations, it has been discovered that individuals were engaged in the distribution of bogus drugs or fraudulent information through social media. As a result of the shutdown, millions of individuals were introduced to social media for the first time, allowing them to stay up to date. It would be preferable if accurate information could be disseminated and people could keep up to speed on the fatal epidemic that has engulfed the whole planet. It has produced a worrisome scenario among individuals as a result of the incorrect material about COVID-19 being circulated, which has resulted in mental disorders. Many people feel that utilising social media is really detrimental. The facts regarding coronavirus include that it is spread via the air and that it may remain on surfaces for many hours. It targets older adults with ease; it causes breathlessness; it causes death in a matter of days; it is incurable; and so on. It is making the rounds on social media at an unexpectedly rapid speed, causing widespread panic.

3.4. River Formation Dynamics (RFD) Algorithm. River Formation Dynamics (RFD) is a technique that uses Evolutionary Computation to model river formation. RFD may be thought of as a gradient-oriented variant of the ACO algorithm. This concept is based on the observation of how water makes rivers in the natural world. It changes the environment as water flows through a steep decreasing slope, eroding the ground underneath it, and depositing the sediments carried by the water when the water falls onto a flatter surface. The basic method and formulas for the dynamics of river have been developed.

Any route from the origin point to the target point has a gradient that, when considered the path as a whole (i.e., from the origin to the target), must be decreasing from the beginning of the RFD execution.

We have used this approach to tackle the issue of location management in order to reduce the overall cost of location management as much as possible.

The RFD algorithm's principle replicates the procedure of riverbed formation. A set of drops situated at a starting point will be subjected to gravitational forces that will attract these drops towards the Earth's center. Hence, these drops will undergo distribution all over the environment in search of the lowest point—the sea. This procedure will result in the

formation of various new riverbeds. Now this concept is used by the RFD for problems of graph theory. First, there will be the creation of an agent-drop set. Afterwards, these drops will travel on the edges between the nodes to discover the environment, searching for the best solution. They will utilize mechanisms of erosion as well as soil sedimentation, which are associated with variations in the altitudes that are allocated to every node. Upon the drops' movement across an environment, it will modify the measurement of the nodes along its route. The shift from one node to another will be done by the nodes' decreasing altitude, which in turn will offer numerous benefits, such as the avoidance of local cycles [24].

The following describes the RFD algorithm. There is an assignment of an amount of soil to every node. When the drops move, they either erode their paths or deposit the carried sediment (and hence, increase the nodes' altitudes). The probability of picking the next node is dependent on the gradient, which is in proportion to the difference between the heights of the node where the drop resides as well as its neighbor's height. The procedure will commence with a flat environment; that is, all the nodes will have equivalent altitudes, except for the zero equivalent goal node that will maintain this value throughout the entire procedure. To facilitate the environment's further exploration, the drops will be situated at the initial node. At every step, a group of drops will successively traverse the space, and later, will execute erosion on the nodes visited. Algorithm 1 represents the RFD algorithm's pseudocode.

Drops will move one till their arrival at the goal, or they have traveled the maximum set number of nodes. The total number of nodes in an environment will constitute the aforementioned maximum number of nodes. Equations (2) to (4) will express the probability $P_k(i, j)$ that a drop k which resides in node i would pick the next node j :

$$P_k(i, j) = \begin{cases} \frac{\text{gradient}(i, j)}{\text{total}}, & \text{for } j \in V_k(i), \\ \frac{\omega/|\text{gradient}(i, j)|}{\text{total}}, & \text{for } j \in U_k(i), \\ \frac{\delta}{\text{total}}, & \text{for } j \in F_k(i), \end{cases} \quad (2)$$

where

$$\text{gradient}(i, j) = \frac{\text{altitude}(i) - \text{altitude}(j)}{\text{distance}(i, j)}, \quad (3)$$

$$\text{total} = \left(\sum_{l \in V_k(i)} \text{gradient}(i, l) \right) + \left(\sum_{l \in U_k(i)} \frac{\omega}{|\text{gradient}(i, l)|} \right) + \left(\sum_{l \in F_k(i)} \delta \right). \quad (4)$$

$V_k(i)$ will denote a neighboring node-set that has a positive gradient (that is, node i 's altitude is higher than that of node j), $U_k(i)$ will denote a neighboring node-set that has

a negative gradient (that is, node j 's altitude is higher than that of node i), and $Fk(i)$ will denote neighbors having a flat gradient. ω and δ coefficients have fixed values.

Once all the drops have finished moving, there is the execution of a procedure of erosion on all the traveled paths through the reduction of the nodes' altitudes based on the gradient to the successive node. According to equation (5), the amount of erosion for each pair of nodes i and j will be dependent on the number of all used drops D , the number of all nodes in the graph N , as well as a specific erosion coefficient E .

$$\forall i, j \in \text{Path}_k, \text{altitude}(i) := \text{altitude}(i) - \frac{E}{(N-1).D} \cdot \text{gradient}(i, j). \quad (5)$$

Here, Path_k will denote the drop k 's traversed path.

Furthermore, when a drop stops, it will deposit a fraction of the carried sediment and also will end up evaporating for the remaining portion of the algorithm iteration. Since this will minimize the likelihood of transition towards blind alleys, this will result in weakening the bad paths.

Upon each iteration's completion, there is the addition of a specific as well as minimal sediment amount to all the nodes (line 8). This is for the avoidance of a situation in which all the altitudes would be close to zero since it would result in negligible gradients and ruination of all the formed paths. The below equation (6) will formulate the sediment to be added as follows:

$$\forall i \in G \wedge i \neq \text{goal}, \text{altitude}(i) := \text{altitude}(i) + \frac{\text{erosion Produced}}{N-1}. \quad (6)$$

In this equation, G will denote the node-set of the utilized graph, the goal will denote the goal node, and erosion produced will denote the sum of all the erosion produced in the current iteration, that is, $\sum_{\text{Path}_k} E / (N-1).D \cdot \text{gradient}(i, j), \forall k \in \text{drops}$.

Till arrival at the final condition, the algorithm iterates. This final condition may indicate all the drops which are moving along the same path. For the computation time's minimization, maximum iterations are defined, and also a condition to verify whether the earlier n loops made any improvements on the solution.

3.5. Particle Swarm Optimization (PSO) Algorithm. PSO algorithm's inspiration was derived from the intelligent [25] collective behavior of certain creatures like fish schools or bird flocks. Akin to other evolutionary algorithms, the evolution of a potential solution population in the PSO will undergo successive iterations. In comparison to other strategies of optimization, the PSO's key benefits are its implementational ease and the low number of parameters for adjustment. In the PSO, every potential solution to a problem of optimization is taken into account as a bird and is also referred to as a particle. The particle set, also termed a swarm, will be made to fly across the problem's D -dimensional search space. Each particle's position will undergo a

change which is based on the experiences of the particle itself as well as those of its neighbors [26].

Equation (7) will express the i th particle's position as below:

$$x_i = (x_{i1}, x_{i2}, \dots, x_{iD}). \quad (7)$$

Here, $x_{id} \in [l_d, u_d], d \in [1, D]$ while l_d and u_d will denote the lower and upper bounds of the search space's d th dimension. Akin to each particle's position, a vector is used to represent each particle's velocity. $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ will express the i th particle's velocity. During every time step, equations (8) and (9) will update each particle's position and velocity as below:

$$v_{ij}(t+1) = v_{ij}(t) + R_{1ij}c_1(P_{ij} - x_{ij}(t)) + R_{2ij}c_2(P_{gj} - x_{ij}(t)), \quad (8)$$

$$x_i(t+1) = x_i(t) + v_i(t+1). \quad (9)$$

These equations R_{1ij} and R_{2ij} will denote two distinct random values within the $[0, 1]$ range; c_1 and c_2 will denote acceleration constants; p_i will denote the particle's best previous position while P_g will denote the best previous position of all particles in the swarm (that is, the global best PSO).

A successful optimization algorithm is primarily dependent on the balance between the global search and the local search throughout a runner's course. For this goal's accomplishment, certain mechanisms are employed by a majority of all the evolutionary algorithms. Examples of balance controlling parameters are inclusive of the temperature parameter in Simulated Annealing and the normal mutation's step size in strategies of evolution. To strike a balance between the PSO's attributes of exploration as well as exploitation, Shi and Eberhart had proposed an inertia weight-based PSO wherein the update of each particle's velocity was following the below :

$$v_{ij}(t+1) = wv_{ij}(t) + R_{1ij}c_1(P_{ij} - x_{ij}(t)) + R_{2ij}c_2(P_{gj} - x_{ij}(t)). \quad (10)$$

While a global search is enabled by a huge inertia weight, a local search is enabled by a small inertia weight. The search ability's dynamic adjustment was achieved via dynamic alteration of the inertia weight. Numerous other researchers were also in agreement with this general statement related to the w 's impact on the PSO's search behavior.

3.6. Proposed RFD-PSO Algorithm. The standard RFD algorithm suffers from a few shortcomings that hinder its performance. The huge number of coefficients will make it extremely unintuitive to tune the algorithm to a specific case. In addition, the algorithm has a very low rate of convergence for environments with more complexity. The intelligence-based PSO has a lot of applicability in scientific research as well as engineering. It does not have any overlapping and mutation calculation. The particle's speed will search. At the time of development of various generations, only the most

- (1) Height of nodes \leftarrow initial height
- (2) Height of target node \leftarrow 0
- (3) while end conditions are not met do
- (4) Place all drops in starting node
- (5) Move all drops across the graph for a maximum number of steps
- (6) Analyse complete paths
- (7) Height of nodes on paths- = erosion based on path costs
- (8) Height of all nodes+ = small amount of sediment
- (9) end while Drops

ALGORITHM 1: RFD algorithm's pseudocode.

optimist particle will have the ability to transmit information towards the other particles. The search's speed will also be rapid. The PSO has a very simplistic calculation.

In comparison with various other developing calculations, it has occupied the bigger optimization capability and also can be easily completed. The PSO has adopted the real number code, and the solution directly determines it. The dimension's number will be equivalent to the solution's constant.

The hybridization's key objectives will be as follows: advancement of the individual basic algorithms' effectiveness, search space's expansion, enhancements in convergence, and local search. In addition, hybridization must have the ability to design effective, coherent, and flexible algorithms to manage multi-objective or continuous optimization problems. To enhance the river drops' quality and convergence in the basic RFD, a novel hybrid RFD-PSO algorithm has been introduced in this work. The evolution process in this proposed algorithm will involve the participation of all solutions in each drop. Moreover, we can achieve enhancements in the local search and the global search through the utilization of the PSO's particle velocity as well as position procedure, respectively. With the utilization of the PSO's concept in the RFD at the time of global information exchange as well as local deep search, we can accomplish enhancements in accuracy, rate of convergence, global exploration as well as local exploration (Algorithm 2).

Flowchart for the hybrid RFD-PSO algorithm can be seen in Figure 1 [28].

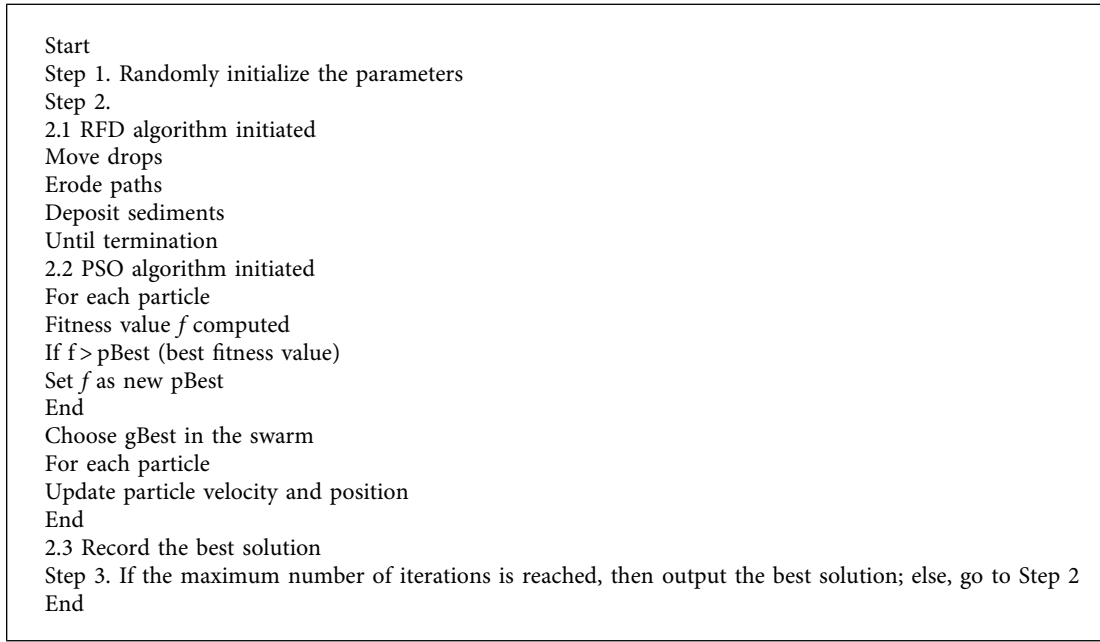
In order to cope with the TTNR issue, a route weighted graph strategy has been used to deal with the situation in question. As a tiled pattern, the routing area may be represented as a grid graph, in which each node (or vertex) represents a tile and each edge (or border between two adjacent tiles) represents the boundary between two adjacent tiles. Simply put, the grid is represented as a square matrix of size $n \times n$, where the size of the matrix equals the number of nodes on one of the grid's sides, multiplied by one hundred. The nodes of the grid are numbered sequentially, starting at the bottom left corner. One node will be designated as the source node, and the other will be assigned as the destination node—both nodes effectively representing the two terminals that will be routed with the least amount of distance. Figure 2 depicts a grid graph of the same size (6×6). It is allocated to the Source and Destination nodes, respectively, the two nets or pins that

are to be linked to one other. The ants, the drops, and the ant-drops, which are the Swarm agents in their respective algorithms, i.e., ACO, RFD, and Hybrid RFD-ACO, are each deployed in the graph in their own way. The ants, the drops, and the ant-drops are each deployed in their own way. At the start of each iteration, the agents are initialized at the Source node, and each agent attempts to discover a route with a high probability of success. The method lists all of the nodes that are accessible from a certain node (except the immediate previous visited node). According to greater likelihood, the next node is picked, i.e., the node with the higher probability value based on pheromones (in the case of ACO) or gradients (in the case of RFD) or both is chosen as the next node (in case of Hybrid RFD-ACO). Any one of the nodes is picked randomly using a random function if there is a tie between two or more nodes depending on the likelihood of the tie occurring. It is in this manner that the agents go from node to node in search of a route to the Destination Node.

When an ant walks from one node to another, a pre-determined quantity of pheromone is deposited along the path that the ant leaves behind. In the case of the RFD, the initial node is eroded, which means that the altitude value of the initial node decreases as a function of the slope of the gradient. This procedure is followed for each such transition in each cycle of transportation from source to destination. Pheromones are evaporated along all of the edges in the case of ACO, and sediment is deposited (altitude value is raised) over every node in the case of RFD over the whole graph once each cycle is completed. In addition, the best roads are reinforced by additional trails of pheromone deposition and soil erosion to make them even more effective. The procedure is repeated until the most efficient path is discovered. The node count and edge count along the route are determined using the algorithm's software, which may be found here. Using the information in this table, it is possible to compute and compare the costs of several courses, and therefore progress towards the convergence of the best routes.

4. Results and Discussion

Depressive patient, a historical dataset for the depressive patient recommender systems, is used to evaluate the algorithm and its quality. It consists of 100,000,029 anonymous ratings from about 6,040 users from 3,952 depressive



ALGORITHM 2: RFD-PSO algorithm's pseudocode [27].

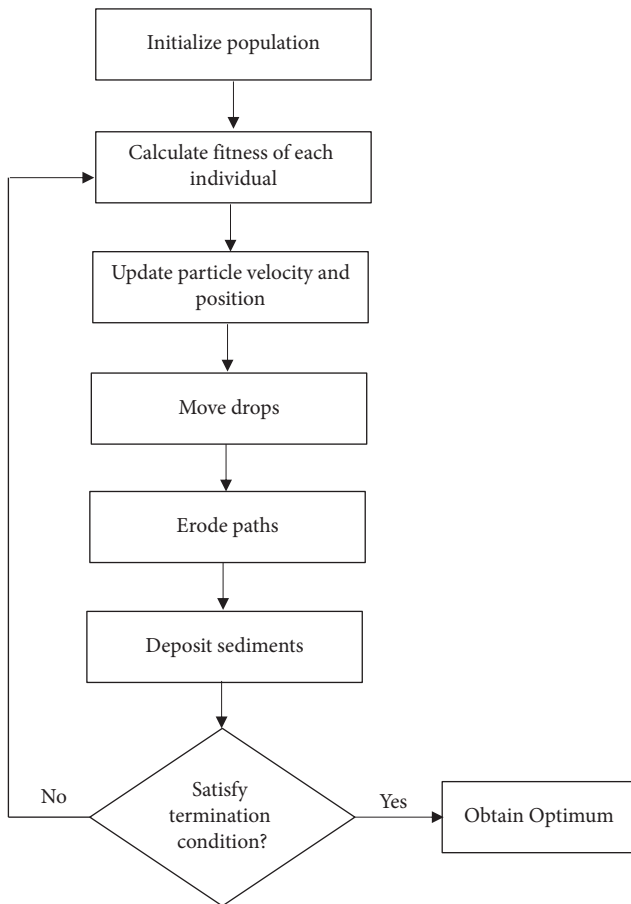


FIGURE 1: Flowchart for hybrid RFD-PSO algorithm.

patients. The Depressive patient datasets are primarily used to evaluate a collaborative recommender system for the depressive patient domain.

In this section, the RFD feature selection-without Frequent Pattern Mining, RFD feature selection-without Frequent Pattern Mining + CF, RFD feature selection-with Systolic Tree frequent pattern mining, RFD feature selection-with Systolic Tree frequent pattern mining + CF, RFD - PSO feature selection-without Frequent Pattern Mining, RFD - PSO feature selection-with Systolic Tree frequent pattern mining and RFD - PSO feature selection-with Systolic Tree frequent pattern mining + CF are used. The experiments were conducted with top $N = 2$ to 18 recommended items. Precision and recall results are shown in Tables 2 and 3 and Figures 2 and 3.

From Figure 3, it can be observed that the RFD - PSO feature selection-with Systolic Tree frequent pattern mining + CF has higher average precision by 9.76% for RFD feature selection-without Frequent Pattern Mining, by 8.07% for RFD feature selection-without Frequent Pattern Mining + CF, by 7.31% for RFD feature selection-with Systolic Tree frequent pattern mining, by 4.91% for RFD feature selection-with Systolic Tree frequent pattern mining + CF, by 5.06% for RFD - PSO feature selection-without Frequent Pattern Mining, by 3.29% for RFD - PSO feature selection-without Frequent Pattern Mining + CF and by 2.28% for RFD - PSO feature selection-with Systolic Tree frequent pattern mining when compared with various top- N recommended items, respectively.

From Figure 2, it can be observed that the RFD - PSO feature selection-with Systolic Tree frequent pattern

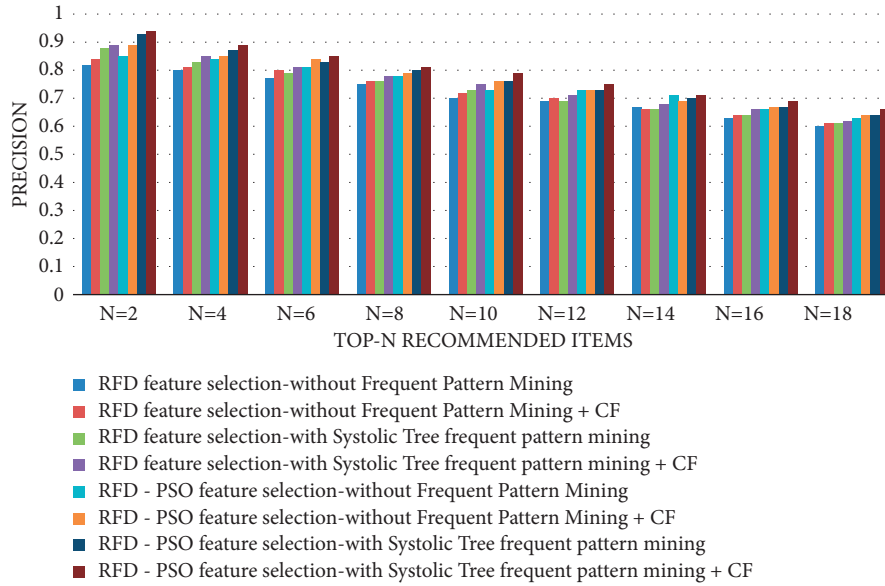


FIGURE 2: Precision for RFD - PSO feature selection - with systolic tree frequent pattern mining + CF.

TABLE 2: Precision for RFD - PSO feature selection - with systolic tree frequent pattern mining + CF.

Top-N recommended items	RFD feature selection-without frequent pattern mining	RFD feature selection-without frequent pattern mining + CF	RFD feature selection-with systolic tree frequent pattern mining	RFD feature selection-with systolic tree frequent pattern mining + CF	RFD - PSO feature selection-without frequent pattern mining	RFD - PSO feature selection-without frequent pattern mining + CF	RFD - PSO feature selection-with systolic tree frequent pattern mining	RFD - PSO feature selection-with systolic tree frequent pattern mining + CF
N=2	0.82	0.84	0.88	0.89	0.85	0.89	0.93	0.94
N=4	0.8	0.81	0.83	0.85	0.84	0.85	0.87	0.89
N=6	0.77	0.8	0.79	0.81	0.81	0.84	0.83	0.85
N=8	0.75	0.76	0.76	0.78	0.78	0.79	0.8	0.81
N=10	0.7	0.72	0.73	0.75	0.73	0.76	0.76	0.79
N=12	0.69	0.7	0.69	0.71	0.73	0.73	0.73	0.75
N=14	0.67	0.66	0.66	0.68	0.71	0.69	0.7	0.71
N=16	0.63	0.64	0.64	0.66	0.66	0.67	0.67	0.69
N=18	0.6	0.61	0.61	0.62	0.63	0.64	0.64	0.66

TABLE 3: Recall for RFD - PSO feature selection - with systolic tree frequent pattern mining + CF.

Top-N recommended items	RFD feature selection-without frequent pattern mining	RFD feature selection-without frequent pattern mining + CF	RFD feature selection-with systolic tree frequent pattern mining	RFD feature selection-with systolic tree frequent pattern mining + CF	RFD - PSO feature selection-without frequent pattern mining	RFD - PSO feature selection-without frequent pattern mining + CF	RFD - PSO feature selection-with systolic tree frequent pattern mining	RFD - PSO feature selection-with systolic tree frequent pattern mining + CF
N=2	0.23	0.26	0.31	0.41	0.24	0.27	0.33	0.43
N=4	0.26	0.32	0.37	0.47	0.27	0.34	0.39	0.5
N=6	0.31	0.36	0.42	0.55	0.32	0.38	0.44	0.58
N=8	0.36	0.41	0.48	0.63	0.38	0.43	0.5	0.66
N=10	0.42	0.47	0.57	0.74	0.44	0.49	0.59	0.78
N=12	0.49	0.56	0.65	0.85	0.52	0.59	0.69	0.9
N=14	0.57	0.64	0.76	1	0.6	0.68	0.8	1
N=16	0.65	0.75	0.88	1	0.69	0.79	0.93	1
N=18	0.76	0.86	1	1	0.8	0.9	1	1

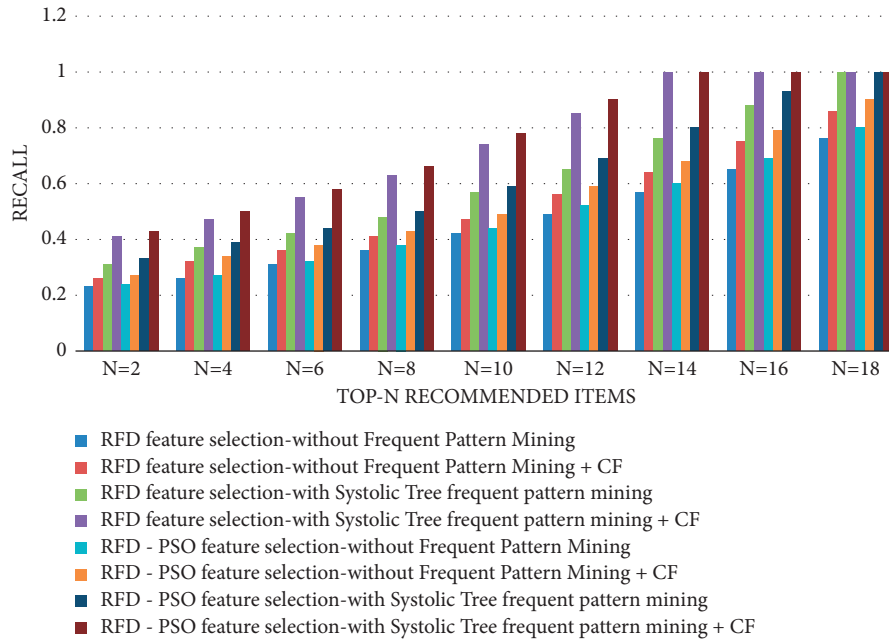


FIGURE 3: Recall for RFD - PSO feature selection - with systolic tree frequent pattern mining + CF.

TABLE 4: Performance metrics Comparison.

Research name	Accuracy (%)	Precision (%)	Recall (%)
SNS-based predictive model for depression [6]	77	78	85
Predicting depression via social media [16]	-70	70	61
Proposed system	63.3	100	57

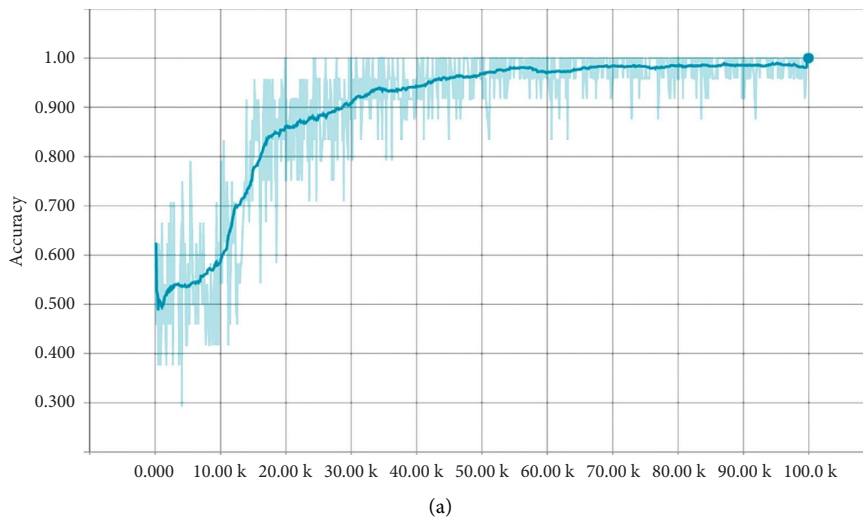


FIGURE 4: Continued.

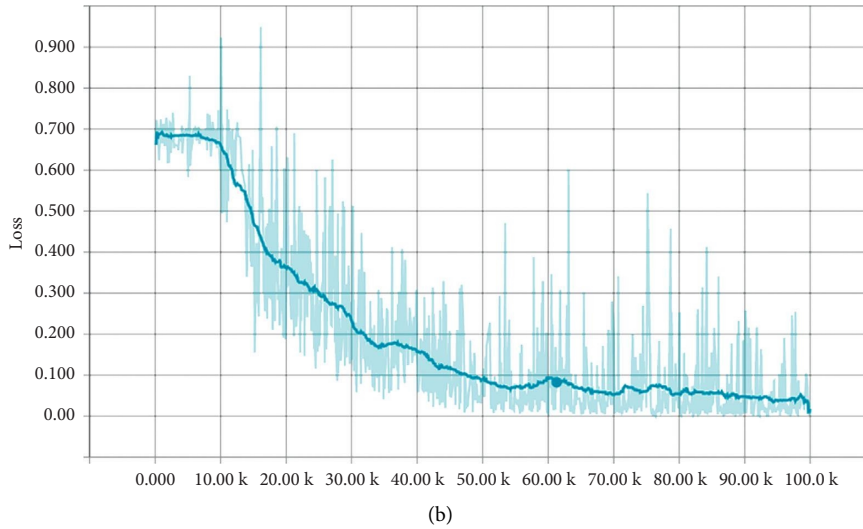


FIGURE 4: Performance metrics of proposed work.

mining + CF has higher average recall by 51.37% for RFD feature selection-without Frequent Pattern Mining, by 38.67% for RFD feature selection-without Frequent Pattern Mining + CF, by 22.94% for RFD feature selection-with Systolic Tree frequent pattern mining, by 2.96% for RFD feature selection-with Systolic Tree frequent pattern mining + CF, by 46.62% for RFD - PSO features selection-without Frequent Pattern Mining, by 33.78% for RFD - PSO feature selection-without Frequent Pattern Mining + CF and by 18.84% for RFD - PSO feature selection-with Systolic Tree frequent pattern mining when compared with various top-N recommended items, respectively. Table 4 represents the performance metrics comparison.

The accuracy of the classification is dependent on the training set that was used to train the classifier and to execute it. Rather than selecting simply apparent instances of a class, it is critical to choose sample training nodes that reflect edge cases that belong in or out of a class. As a result, it is a best practice to include as many different types of samples as feasible in the training set. This has been accomplished via the collection, organisation, and manual training of a supervised dataset. The postings for the dataset were gathered from three social media platforms: Facebook, LiveJournal, and Twitter. The dataset was manually trained to identify two types of sentiment: depressed and not depressed. In the case of depressed sentiment, we classified the depressed post into one of the nine depression symptoms defined by the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM-IV).

Figure 4 has shown the performance metrics of the proposed work through the parameters such as accuracy and the loss.

5. Conclusions

Using quality recommendations in social media, the Recommender Systems have been able to enhance the user experience and thus, effectively handle the information

overload issue. FP extraction has been done with the utilization of the techniques of association rule mining. Upon the preprocessed data's application with the TF-IDF feature extractor, every document will obtain a vectorized representation based on the TF-IDF scores on the terms within every document. With the utilization of the RFD optimization algorithm, there is the optimal path's computation under a specified constraint of time. As a swarm intelligence technique, the population-based PSO will execute the process of optimization to attain its fitness function optimization. In the hybrid RFD-PSO algorithm's proposal, a small constant updating strategy's introduction will boost the update capability of velocity, acceleration factor, and optimal individual location. There is the PSO strategy's utilization for optimizing the RFD's velocity as well as position. Results show that the RFD - PSO feature selection-with Systolic Tree frequent pattern mining + CF has higher average precision by 9.76% for RFD feature selection-without Frequent Pattern Mining, by 8.07% for RFD feature selection-without Frequent Pattern Mining + CF, by 7.31% for RFD feature selection-with Systolic Tree frequent pattern mining, by 4.91% for RFD feature selection-with Systolic Tree frequent pattern mining + CF, by 5.06% for RFD - PSO feature selection-without Frequent Pattern Mining, by 3.29% for RFD - PSO feature selection-without Frequent Pattern Mining + CF, and by 2.28% for RFD - PSO feature selection-with Systolic Tree frequent pattern mining when compared with various top-N recommended items, respectively [29].

6. Limitations

The basic concept of our research is to determine if there is a link between the actions of SNS users and mental health problems. We believe that social media activity might disclose the presence of mental disease in its early stages and the limitations are time consumption in training and testing. The psychiatrist will not be able to get all of the information

from the depressed patient if he or she uses typical questioning strategies. The SNS-based approach has the potential to address the difficulties associated with self-reporting. We may learn more about the depressed patient's natural behaviour and style of thinking by observing his or her social activities, and we can better categorize the different mental levels based on these observations. So, in the future work, the online application gathers user-generated content (UGC) from the patient's Twitter and/or Facebook accounts. Following that, it takes depressive responses about the patient from the user, with the answers being based on the BDI-II depression questionnaire [11]. Following that, it examines the UGC using a variety of text analysis APIs. Finally, it assigns the patient to one of four categories of depression (Minimal, Mild, Moderate, or Severe) based on their symptoms. Following that, we developed a predated depression model in RapidMiner, which was used to evaluate two classifiers (SVM and Nave Bayes Classifier) for depression. Using the same patients' data that has been supplied to the proposed web application and in accordance with a training dataset, 2073 depressed post and 2073 not-depressed post have been manually categorised using depressed post and not-depressed post. The performance of the three outcomes, namely, the sentiment results, the SVM results, and the Nave Bayes findings, has been computed.

Data Availability

The data that 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 to report regarding the present study.

References

- [1] M. Wasid and V. Kant, "A particle swarm approach to collaborative filtering based recommender systems through fuzzy features," *Procedia Computer Science*, vol. 54, pp. 440–448, 2015.
- [2] Y. Nemati and H. Khademolhosseini, "Devising a profit-aware recommender system using multi-objective GA," *Journal of Advances in Computer Research*, vol. 11, no. 3, pp. 109–120, 2020.
- [3] S. Yadav and S. Nagpal, "An improved collaborative filtering based recommender system using bat algorithm," *Procedia Computer Science*, vol. 132, pp. 1795–1803, 2018.
- [4] M. S. Vignesh, M. K. S. Banu, and M. K. M. Kumar, "Efficient algorithms systolic tree with abc based pattern mining algorithm for high utility itemsets from transactional databases," *International Journal of Computer Science and Mobile Computing*, vol. 3, no. 7, pp. 350–357, 2014.
- [5] A. Ragone, P. Tomeo, C. Magarelli et al., "Schema-summation in linked-data-based feature selection for recommender systems," in *Proceedings of the Symposium on Applied Computing*, pp. 330–335, Morocco, April, 2017.
- [6] G. Khambra and P. Shukla, "Novel machine learning applications on fly ash based concrete: an overview," *Materials Today Proceedings*, pp. 2214–7853, 2021.
- [7] P. K. Shukla, J. K. Sandhu, A. Ahirwar, D. Ghai, P. Maheshwary, and P. K. Shukla, "Multiobjective genetic algorithm and convolutional neural network based COVID-19 identification in chest X-ray images," *Mathematical Problems in Engineering*, vol. 1, p. 9, Article ID 7804540, 2021.
- [8] P. Rabanal, I. Rodríguez, and F. Rubio, "Towards applying river formation dynamics in continuous optimization problems," in *Proceedings of the International Work-Conference on Artificial Neural Networks*, pp. 823–832, Springer, Cham, June, 2019.
- [9] X. Cai, Z. Hu, and J. Chen, "A many-objective optimization recommendation algorithm based on knowledge mining," *Information Sciences*, vol. 537, pp. 148–161, 2020.
- [10] B. Alhijawi and Y. Kilani, "A collaborative filtering recommender system using genetic algorithm," *Information Processing & Management*, vol. 57, no. 6, Article ID 102310, 2020.
- [11] B. Alhijawi, Y. Kilani, and A. Alsarhan, "Improving recommendation quality and performance of genetic-based recommender system," *International Journal of Advanced Intelligence Paradigms*, vol. 15, no. 1, pp. 77–88, 2020.
- [12] S. Rakshana Sri, L. Ravi, V. Vijayakumar, X. Z. Gao, V. Subramaniaswamy, and N. Sivaramakrishnan, "An effective user clustering-based collaborative filtering recommender system with grey wolf optimisation," *International Journal of Bio-Inspired Computation*, vol. 16, no. 1, pp. 44–55, 2020.
- [13] N. Tohidi and C. Dadkhah, "Improving the performance of video collaborative filtering recommender systems using optimization algorithm," *International Journal of Nonlinear Analysis and Applications*, vol. 11, no. 1, pp. 283–295, 2020.
- [14] W. H. El-Ashmawi, A. F. Ali, and A. Slowik, "Hybrid crow search and uniform crossover algorithm-based clustering for top-N recommendation system," *Neural Computing & Applications*, vol. 33, no. 12, pp. 7145–7164, 2020.
- [15] H. Wang, B. Niu, and L. Tan, "Bacterial colony algorithm with adaptive attribute learning strategy for feature selection in classification of customers for personalized recommendation," *Neurocomputing*, vol. 452, pp. 747–755, 2021.
- [16] S. Pandit, P. K. Shukla, A. Tiwari, P. K. Shukla, M. Maheshwari, and R. Dubey, "Review of video compression techniques based on fractal transform function and swarm intelligence," *International Journal of Modern Physics B*, vol. 34, no. 08, Article ID 2050061, 2020.
- [17] L. Si and R. Jin, "Flexible mixture model for collaborative filtering," *20th International Conference on Machine Learning*, vol. 2, pp. 704–711, August 2003.
- [18] X. Su, R. Greiner, T. M. Khoshgoftaar, and X. Zhu, "Hybrid Collaborative Filtering Algorithms Using a Mixture of Experts," in *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 645–649, Fremont, CA, USA, November 2007.
- [19] J. Wang, A. P. de Vries, and M. J. T. Reinders, "Unified relevance models for rating prediction in collaborative filtering," *ACM Transactions on Information Systems*, vol. 26, no. 3, pp. 1–42, June 2008.
- [20] C. W. k. Leung, S. C. F. Chan, and F. L. Chung, "A collaborative filtering framework based on fuzzy association rules and multiple-level similarity," *Knowledge and Information Systems*, vol. 10, no. 3, pp. 357–381, 2006.
- [21] D. Y. Pavlov and D. M. Pennock, "A maximum entropy approach to collaborative filtering in dynamic, sparse, high-dimensional domains," *Neural Information Processing Systems*, pp. 1441–1448, 2002.

- [22] S. P. Malarvizhi and B. Sathiyabhama, "Enhanced reconfigurable weighted association rule mining for frequent patterns of web logs," *International Journal of Computing*, vol. 13, no. 2, pp. 97–105, 2014.
- [23] Y. Zhang, Y. Zhou, and J. Yao, "Feature extraction with TF-IDF and game-theoretic shadowed sets," in *Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp. 722–733, Springer, Cham, 2020, June.
- [24] G. Redlarski, M. Dabkowski, and A. Palkowski, "Generating optimal paths in dynamic environments using River Formation Dynamics algorithm," *Journal of Computational Science*, vol. 20, pp. 8–16, 2017.
- [25] V. Roy, P. K. Shukla, A. K. Gupta, V. Goel, P. K. Shukla, and S. Shukla, "Taxonomy on EEG artifacts removal methods, issues, and healthcare applications," *Journal of Organizational and End User Computing*, vol. 33, no. 1, pp. 19–46, 2021.
- [26] A. Nickabadi, M. M. Ebadzadeh, and R. Safabakhsh, "A novel particle swarm optimization algorithm with adaptive inertia weight," *Applied Soft Computing*, vol. 11, no. 4, pp. 3658–3670, 2011.
- [27] C. B. Kalayci and S. M. Gupta, "River formation dynamics approach for sequence-dependent disassembly line balancing problem," *Reverse supply chains: Issues and Analysis*, pp. 289–312, 2013.
- [28] J. S. Wang and J. D. Song, "A hybrid algorithm based on gravitational search and particle swarm optimization algorithm to solve function optimization problems," *Engineering Letters*, vol. 25, no. 1, 2017.
- [29] N. K. Rathore, N. K. Jain, P. K. Shukla, U. S. Rawat, and R. Dubey, "Image Forgery Detection Using Singular Value Decomposition with Some Attacks," *National Academy Science Letters*, vol. 44, pp. 331–338, 2021.

Research Article

Multi-Ideology, Multiclass Online Extremism Dataset, and Its Evaluation Using Machine Learning

Mayur Gaikwad ¹, Swati Ahirrao ¹, Shraddha Phansalkar ², Ketan Kotecha ³,
and Shalli Rani⁴

¹Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, MH 412115, India

²MIT Art, Design and Technology University, Pune, MH 412201, India

³Symbiosis Centre for Applied Artificial Intelligence, Symbiosis International (Deemed University), Pune, MH 412115, India

⁴Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab 140401, India

Correspondence should be addressed to Swati Ahirrao; swatia@sitpune.edu.in and Ketan Kotecha; head@scaai.siu.edu.in

Received 2 November 2021; Revised 8 August 2022; Accepted 24 November 2022; Published 1 March 2023

Academic Editor: Lorenzo Putzu

Copyright © 2023 Mayur Gaikwad 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.

Social media platforms play a key role in fostering the outreach of extremism by influencing the views, opinions, and perceptions of people. These platforms are increasingly exploited by extremist elements for spreading propaganda, radicalizing, and recruiting youth. Hence, research on extremism detection on social media platforms is essential to curb its influence and ill effects. A study of existing literature on extremism detection reveals that it is restricted to a specific ideology, binary classification with limited insights on extremism text, and manual data validation methods to check data quality. In existing research studies, researchers have used datasets limited to a single ideology. As a result, they face serious issues such as class imbalance, limited insights with class labels, and a lack of automated data validation methods. A major contribution of this work is a balanced extremism text dataset, versatile with multiple ideologies verified by robust data validation methods for classifying extremism text into popular extremism types such as *propaganda*, *radicalization*, and *recruitment*. The presented extremism text dataset is a generalization of multiple ideologies such as the standard ISIS dataset, GAB White Supremacist dataset, and recent Twitter tweets on ISIS and white supremacist ideology. The dataset is analyzed to extract features for the three focused classes in extremism with TF-IDF unigram, bigrams, and trigrams features. Additionally, pretrained word2vec features are used for semantic analysis. The extracted features in the proposed dataset are evaluated using machine learning classification algorithms such as *multinomial Naïve Bayes*, *support vector machine*, *random forest*, and *XGBoost* algorithms. The best results were achieved by support vector machine using the TF-IDF unigram model confirming 0.67 F1 score. The proposed multi-ideology and multiclass dataset shows comparable performance to the existing datasets limited to single ideology and binary labels.

1. Introduction

Social media have become an integral part of life in the current era. People share their thoughts, beliefs, and ideas over social media platforms. Social media platforms such as Twitter, Facebook, WhatsApp, and Instagram are popular mediums of expression among people. Over 474,000 messages are posted on Twitter, and 293,000 statuses are updated on Facebook [1].

Social media platform offers extensive outreach and hence become extremely influential. This makes the social media platform a perfect tool for the extremists to spread their propaganda, radicalization, and recruitment. The extremist groups share violent messages, images, and videos over social media. The extremist organizations such as the *Islamic State of Iraq and Syria* (ISIS) [2] and *Al Qaeda* [3] use social media platforms for the spread of extremism amongst the susceptible youth.

Similarly, far-right-wing organizations such as *Alt-Right* [4] and *Proud Boys* [5] also use social media platforms to radicalize and recruit the youth. *Bill S-894* [6] claims that 73% of the violent incidents in the USA after 11 September 2001 have links with far right-wing organizations.

In the recent Christchurch mosque attack [7], perpetrators were influenced by Oslo attackers manifesto [8], spread through online means. Perpetrators live-streamed the Christchurch mosque attack on Facebook [8]. Facebook blocked the initial spread of the attack video; however, some reuploads were left undetected [9].

Online extremism research is crucial to constrain the spread of harmful ideologies amongst the susceptible youth. It also helps the regulatory bodies to monitor and control the spread of extremism.

Online extremism is carried out in the following three ways: (1) spreading propaganda, (2) attracting youths through the recruitment messages, and (3) the radical change in the perception towards an individual or community.

Propaganda is “content, generally biased, which is exploited for the personal or the political cause” [10]. Misinformation used for political gains is also termed “propaganda.” Propaganda is usually used by dictatorial administrations such as Nazism in Germany and the former Soviet Union to brainwash people. Propaganda such as “America is dead! Long Live America” [11] is used to attract people.

Jihadist propaganda mainly related to ISIS can be found in their online magazines “Dabiq” and “Rumiyah” [12]. The magazines contain propaganda in the form of glorification of the caliphate and battlefield [13]. White supremacist propaganda used by some organizations follows methods such as pamphlets similar to ISIS [11].

Radicalization is a “change in behavior, attitude, and perception towards a person or a community” [14]. Miscreants use online radicalization to mislead people by quoting their beliefs that may be political or religious [15]. Both jihadists and white supremacists use current events, encourage weapons, and violent attacks as radicalization strategies [11]. Text such as “you do realize IS wants to destroy every single nation-state, Arab or Kurd or communist does not matter, that they come across?” [16], radicalizes people in the name of religion, organization, or nation.

Recruitment in the area of extremism is the “incitement of youths to sacrifice themselves and perform violent acts on behalf of the extremist organization [17].” Jihadist-ISIS recruiters glorify ISIS fighters’ death as martyrdom and exploit it as a recruitment tactic [18]. White supremacists use “feelings of inadequacy,” “anti-government themes,” and recently “coronavirus themes” to recruit disgruntled youth [11]. Extremists use posters with text such as “Join the Atomwaffen Division,” which directly calls for recruitment to the specific extremist organization [11].

Every type of extremist text and speech such as propaganda, radicalization, and recruitment has distinct features and effects. These are also explained in [19]. As social media reach is ever-expanding, extremist organizations use

these platforms to spread propaganda, radicalize people, and recruit them for violent acts. Thus, it is necessary to develop a tool for identifying propaganda, radicalization, and recruitment to restrict the spread of extremism on social media platforms [16]. The online extremism research faces the following challenges:

- (1) Lack of publicly available datasets of the extremism text
- (2) Lack of the ideology-independent and balanced datasets of the extremism text
- (3) Lack of automated data validation methods for checking the quality of data
- (4) Lack of accurate automated detection methods for the online extremism text
- (5) Limited work on extremism content classification into categories, such as radicalization, propaganda, and recruitment

The contribution of our work is as follows:

- (1) Construction of multi-ideology balanced and extremism text dataset collected from multiple sources such as StormFront Dataset [20], Gab dataset [21], ISIS Kaggle dataset [22], and Twitter
- (2) The application of statistical data validation methods for checking the quality of the proposed dataset
- (3) The development of an automated framework for the detection of online extremism text, which classifies the extremism content as radicalization, propaganda, and recruitment
- (4) Implementation of the proposed framework with AI techniques for efficient and accurate detection of online extremism
- (5) Comparative performance analysis of the proposed dataset Merged ISIS-White Supremacist (MIWS) with Merged ISIS dataset (MIS), Merged White Supremacist dataset (MWS)
- (6) Investigation of the best feature extraction technique and classifier for the proposed extremism text dataset

This research work targets two ideologies ISIS/jihadist and white supremacist. The reason behind selecting these two ideologies is based on various factors such as infamy [23], support of violence [2, 8], and the spread of ideology online and offline [24]. Twitter is one of the most popular social media platforms with an extensive reach. Multiple studies have proved that extremists prefer Twitter for spreading propaganda, radicalization, and recruitment [16, 25, 26]. So, StormFront [20] and Gab datasets [21] are referred to as hate speech datasets. Hate speech is defined as the “attack or use of discriminatory language with reference to a person or group” [27]. At the same time, extremism can be referred to as “ideas that are opposed to society’s core values which can be of various forms racial or religious supremacy or ideologies that deny basic human rights or democratic principles” [28]. There are multiple definitions of hate speech [29, 30] and similarly multiple definitions of

extremism [31, 32]. However, there is a significant similarity in the definitions and interpretations of hate speech and extremism overlaps. Organizations such as the EU already consider StormFront and Gab the primary platform for right-wing extremist views [33]. Therefore, StormFront and Gab datasets are considered extremists for this paper.

2. Related Work

Existing literature on extremism detection is analyzed by considering the employed datasets and the classifier techniques applied.

2.1. Datasets

2.1.1. Standard Dataset. In standard datasets, extremism text is collected, which is based on a specific ideology. The ISIS Kaggle dataset [22] was compiled by the Fifth Tribe organization to analyze the online spread of ISIS and to counteract them. The dataset contains 17,350 tweets from 112 *pro-ISIS* user accounts, collected after Paris attacks [34] in November 2015. The dataset contains 15,684 English-language tweets. This dataset includes username, location, number of followers, and timestamp of the tweet. It is used in multiple studies to detect and analyze ISIS supporters [35, 36]. The ISIS Kaggle dataset is unlabelled. Different researchers used various techniques to label the dataset. The main problem of the ISIS Kaggle dataset is that there are old accounts in the dataset, which Twitter may have suspended for discarding their hate speech policy.

The “*About ISIS Kaggle Dataset*” [37] acts as a counterpoise to the *ISIS Kaggle Dataset*. This dataset has around 122K tweets mentioning “isis,” “isil,” “daesh,” “islamic state,” “raqqa,” and “mosul.” The dataset is unlabelled, containing *pro-ISIS* accounts, as the data collected is based on keywords. Most of the accounts are unavailable or deleted in the *ISIS Kaggle dataset*.

In *ISIS Religious Text Kaggle dataset* [38], data is collected by Fifth Tribe. This dataset is compiled by scraping of fifteen and nine issues of Dabiq and Rumiyah magazines, respectively. The dataset contains a total of 2,685 texts. Standard datasets related to jihadism or ISIS ideology are unlabelled and contain suspended accounts.

There are very few standard datasets available in the literature on White supremism hate speech. de Gibert et al. [20] collected the extremist hate speech data from StormFront and the White supremacist website. de Gibert et al. compiles 10,568 posts and manually annotates them as *hate*, *nohate*, *relation*, and *skip*. The experts identified a total of 1,119 hate posts and 8,537 nohate posts. de Gibert et al. compare the characteristics of the StormFront dataset with the Hatebase dataset. The StormFront dataset has a major issue of class imbalance.

Kennedy [21] collected 27,000 posts from the Gab social network. Gab social network claims to preserve the freedom of speech and has become a haven for disseminating hate speech. The authors categorize posts into attack on human dignity (HD), call for violence (CV), and offensive/vulgar language (VO). The authors further classify HD and CV into

implicit, explicit, race/ethnicity, nationality, gender, religion, sexual orientation, ideology, political ideology, and mental/physical health. The authors considered three classes, HD, VO, and hate (a combination of HD and CV), for the classification.

The standard datasets in both ISIS and White supremacist ideology are very few. The accounts from which data is collected may have been inactive, suspended, or deleted by the user or the social media platforms. Therefore, the labels provided within datasets are inadequate to provide insights into extremism linguistics in both ideologies. Furthermore, there is a lack of data validation techniques to evaluate the standard datasets. Hence, many researchers prefer to collect extremism-related data from various sources, and manual annotation is performed due to these issues.

2.1.2. Custom Dataset. Similar to standard datasets, custom datasets are created to represent specific ideologies. Berger [25] in 2014 collected 20,000 ISIS-related accounts from Twitter. The author analyzed the location of supporters, languages spoken by the supporters, identification information of supporters, when the supporter accounts were created, the content of posts by ISIS supporters, and the methods used for the identification of propaganda and recruitment.

Chatfield et al. [16] collected 3,036 tweets from @shamiwitness, who was a known ISIS sympathizer. The tweets of @shamiwitness were manually annotated with propaganda, radicalization, and recruitment by the authors. The account of @shamiwitness is now suspended so that no further analysis can be performed. The authors rely on manual data validation methods with no statistical evidence.

Rowe and Saif [39] used the dataset provided by O’Callaghan et al. [40] as the SEED dataset. From the SEED dataset, the authors identified 154K users suspected of spreading ISIS propaganda. The authors collected 3,200 tweets from each user resulting in 104 million tweets. The authors found 43% of tweets in English, 41% in Arabic, and the rest in Spanish and Dutch. For validation of the dataset, the authors used *interrater agreement* using two annotators. In addition, the authors used a sample of 2,000 tweets for manual validation, and the agreement of annotators was between 0.4 and 0.6 Fleiss’ Kappa. The authors did not use any other statistical technique for data validation.

Kaati et al. [41] used 66 Twitter users as seeds obtained from Shumukh al-Islam Forum. The authors used hashtags such as #ISLAMICSTATE, #ILoveISIS, and #AllEyesOnISIS. Thus, a total of 27,253 English *pro-ISIS* tweets and 16,000 Arabic *pro-ISIS* tweets were collected. The authors did not provide any information on data validation.

Ashcroft et al. [42] used similar methods described by Kaati et al. [41] to collect a total of 7,500 tweets consisting of *pro-ISIS*, *anti-ISIS*, and random contexts. Unfortunately, most of the data were collected from older accounts, which may have been suspended.

Benigni et al. [43] used a two-step snowballing process to collect accounts related to ISIS. In the first step, the authors used five seed accounts to collect 1,345 unique accounts. The

authors collected 1,19,156 user accounts in the second step, which followed or related to 1,345 accounts of the previous step. Thus, the authors collected a total of 862M tweets by the end of step two. Unfortunately, due to the Twitter data-sharing policy, the tweets collected by the authors were not available to the public.

Abrar et al. [44] gathered 13,369 terrorism-supporting tweets, 16,506 terrorism-nonsupporting tweets, and 38,617 random tweets. However, the authors neither mentioned any seed accounts or terrorism-specific keywords used to gather tweets nor performed any data validation methods on the collected dataset.

Ahmad et al. [45] gathered ISIS-related tweets using keywords such as ISIS, bomb, and suicide. The authors also used manually identified seed words for identifying ISIS-related tweets. The authors conclude that 12,754 tweets were extremists and 8,432 were nonextremists. However, the research work lacks data validation on the collected data.

Asif et al. [46] used the Facebook pages of news agencies such as PTV news, Dawn, and Geo to gather extremist texts. A total of 19,497 posts were collected, from which 5,279 were labeled as moderate, 6,912 as highly extreme, 2,991 as low extreme, and 4,315 as neutral. The authors used survey-based validation, using 109 random people. However, the authors used only a sample of 25 posts which may not represent the whole data.

Gialampoukidis et al. [47] collected ISIS-related data by searching five keywords provided by law enforcement agencies and domain experts. So, this resulted in 9,528 tweets from 4,400 suspected ISIS-supporting users. Unfortunately, this dataset is unavailable due to the data-sharing policy of Twitter.

The researchers collected data for extreme right-wing, White supremacist ideology from different sources and locations. Jaki and De Smedt [48] collected 50,000 tweets from about 100 Twitter users suspected of supporting far-right ideology in Germany. The authors also collected 50,000 neutral tweets. The authors did not provide any details about data validation methods.

Berger [26] manually collected data from 41 Twitter users who supported the alt-right movement. By checking these accounts' followers, the author collected 27,895 user accounts suspected of supporting the alt-right movement. Berger also collected data from 33,766 neutral user accounts. The author used manual validation for the collected data. *Alt-Right Demographics dataset* is not available publicly due to Twitter data sharing policies. So, the reproducibility of results is not possible.

Some researchers also collected data from multiple ideologies. For example, De Smedt [49] used a multidomain perspective for extremism detection. The authors divided the text into jihadism (ISIS), extremism (far right-wing from Germany, Belgium, Netherlands, US, UK, and Canada), sexism, and racism. The authors collected 50,000 tweets for jihadism, 92,500 tweets for extremism, 10,000 tweets with 15,000 Facebook posts for racism, and 65,000 posts from Incels.me about sexism. The authors used *hate* and *safe* labels for extremism, jihadism, sexism, and racism domains. The authors also used left and right labels for the extremism domain. The authors also analyzed demographic profiling,

psychological profiling, sentiment analysis, and network analysis with detection. Unfortunately, De Smedt et al. do not provide access to the datasets due to strict Twitter policies on data sharing.

Similarly, Berger [23] compared two ideologies ISIS and Nazis, by collecting data from Twitter. First, to identify the users with White supremacist and Nazi sympathies, the author used 18 seed accounts. The author then collected around 200 tweets from a total of 25,406 followers of these 18 seed accounts. Then, for analysis, the authors used 4,000 highly relevant Nazi-sympathizing accounts. Finally, the author used a similar strategy to collect 4,000 ISIS sympathizing accounts from Twitter.

Heidarysafa et al. [50] compared the women-specific content of ISIS with women-specific Catholic preaching. The authors collected 20 articles from Dabiq and Rumiya targeting women and 132 articles from catholicwomensforum.org. The authors relied on manual validation but did not provide any statistical evidence.

Araque and Iglesias [51] used different datasets such as Pro-Neu, Pro-Anti, Magazines, SemEval2019 [52], and Davidson [53] to classify radicalization and hate speech using AffectiveSpace and SenticNet. The authors also used multiple features such as TF-IDF and similarity-based sentiment projection (SIMON) for prediction.

Mussiraliyeva et al. [54] collected religious extremist posts from VKontakte [55] social media platforms in the Kazakh language. The authors used different extremist keywords such as "kafir" and "kill" to identify extremist texts. The annotation of an extremist text is based on the appearance or absence of selected extremist keywords within the text.

From Table 1, it is observed that issues plaguing custom datasets are *data availability*, *result reproducibility*, *binary classification*, *data imbalance*, and *single ideology focus*. Data availability is an issue due to the policy of social media. So, in turn, this affects the reproducibility of the results for other researchers. Nearly all the researchers using the custom datasets use binary classification, which is inadequate for deeper analysis. The extremism data are less than non-extremist data. Thus, the class imbalance is inherent in the custom datasets. The biggest problem of both standard and custom datasets is that their focus is on a single ideology.

Thus, there is a need for a generic dataset of the extremism text, which accounts for multiple ideologies. Additionally, the dataset should help classify extremism text into popular types, that is, propaganda, radicalization, and recruitment. Thus, a generic dataset with multiple ideologies and a single-model multiclassification can efficiently detect online extremism text. These challenges are further explained in Section 3.

2.2. Challenges with Existing Online Extremism Datasets. There are various research gaps found in the dataset of online extremism text. The following challenges are observed in online extremism text datasets as illustrated in Figure 1:

2.2.1. Data Imbalance and Binary Classification. Data imbalance is a serious problem for online extremism datasets.

TABLE 1: Datasets.

Dataset	Dataset type	Source	Language	Data collection period	Labels and percentage in dataset	Validation methods
ISIS Kaggle dataset [22]	Standard	Twitter	English	2015	No labels	No validation
About ISIS Kaggle dataset [37]	Standard	Twitter	English	2016	No labels	No validation
ISIS Religious Text Kaggle dataset [38]	Standard	Twitter	English	2014–2017	Quran 38% Hadith 27% Other 35%	No validation
StormFront [20]	Standard	Website	English	2017	Hate 11.29% Nohate 86.09% Relation 1.69% Skip: 0.93%	Manual validation: Cohen's kappa and Fleiss' kappa
Gab Hate Corpus [21]	Standard	Gab	English	2016–2018	Assault on human dignity (HD) 8.5% Calls to violence (CV) 0.6% Vulgar or offensive language (VO) 6.3%	Manual validation: Fleiss' kappa and prevalence-adjusted and bias-adjusted Kappa (PABAK)
Berger and Morgan [25]	Custom	Twitter	Arabic, English, and French	2014	ISIS supporting accounts 93%	Manual validation
Chatfield et al. [16]	Custom	Twitter	English	2014	No labels	No validation
Rowe and Saif [39]	Custom	Twitter	English, Arabic, Dutch, and Spanish	Not specified	Pro-ISIS 0.4%	Manual validation: Fleiss' kappa
Kaati et al. [41]	Custom	Twitter	English and Arabic	2014	Pro-ISIS English 31.60% Random English 68.37% Pro-ISIS Arabic 26.22% Random Arabic 73.77%	No validation
Benigni et al. [43]	Custom	Twitter	Not specified	2014	ISIS OEC member 15.38% Terrorism supporting 24.25% Terrorism nonsupporting 29.94% Random 70.05%	Manual validation
Abrar et al. [44]	Custom	Twitter	English	2018	Extremist 60.02% Nonextremist 39.79%	No validation
Ahmad et al. [45]	Custom	Twitter	English and Arabic	Not specified	Moderate 27.07% High extreme 35.45% Low extreme 15.34% Neutral 22.13%	No validation
Asif et al. [46]	Custom	Facebook	English and Urdu	2018		Survey-based validation
Gialampoukidis et al. [47]	Custom	Twitter	English	Not specified	Not specified	No validation
Jaki and De Smedt [48]	Custom	Twitter	German and English	2017 F02D 2018	Hate 50% Safe 50%	No validation

TABLE 1: Continued.

Dataset	Dataset type	Source	Language	Data collection period	Labels and percentage in dataset	Validation methods
Berger [26]	Custom	Twitter	English, Spanish, and Dutch	Not specified	Not specified	Manual validation
De Smedt [49]	Custom	Twitter	English, Dutch, German, French, and Arabic	2014–2018	Jihadism 20% Extremism 40% Racism 10% Sexism 30%	No validation
Berger [23]	Custom	Twitter	Not specified	2016	Nazi supporting accounts ISIS supporting accounts 50%	Manual validation
Heidarysafa et al. [50]	Custom	Magazines and website	English	2017	Not specified	Manual validation
Araque and Iglesias [36]	Custom	Twitter and magazines	English	2018–2020	Extremist: 50% Nonextremist: 50% Hate: 58% Nonhate: 42%	No validation

StormFront dataset [20] and Gab dataset [21] are good examples of class imbalance. As extremism data is the fraction of the total data on social media, creating a balanced class dataset is challenging.

Another problem with the dataset is binary or at the most three-class classification of extremism data. Extremist-nonextremist, pro-ISIS-not Pro-ISIS, and hate-not hate are some of the available binary classes. The third class, if available, is either called “irrelevant” or “neutral.” Unfortunately, this classification does not provide analytical insights into the extremism text. Thus, limiting the understanding of extremist activities on social media. Moreover, the expressions of extremism are complex and change over time. Therefore, it is necessary to create the categories based on the context of extremist texts.

2.2.2. Language. The extremism in different ideologies is spread through different languages. Thus, the identification of the extremist text becomes more challenging. Most researchers use English as the global language. The extremist widely uses English to spread their ideology worldwide. Multiple studies by Jaki and De Smedt [48], and De Smedt [49], have addressed online extremism in Dutch and German languages. Rowe and Saif [39] collected dataset containing ISIS-related tweets in English, Arabic, Spanish, and Dutch languages, but limited their research studies to English and Arabic languages.

2.2.3. Outdated Dataset. Standard datasets such as ISIS Religious Text dataset [38] are old. This is because these datasets were obtained during the early days of ISIS. Another issue is the strict data-sharing policy of social media, which makes updating old datasets impossible. This strict data-sharing policy is also one reason for the fewer numbers of standard datasets.

2.2.4. Validation. Most researchers use manual validation with the interrater agreement. As it is impossible to validate an entire data manually, few random samples are used for data validation. Thus, bias is introduced unknowingly. The number of experts also affects the bias in data validation. Fewer experts may give good interrater agreement, but the bias persists. The use of multiple experts may lower the bias, but the *interrater agreement* may deteriorate [46].

2.2.5. Data Quality Assessment. In online extremism research, researchers often collect their own data [26, 35]. Due to the restriction of social media and other issues, previous custom datasets are not available publicly. So, the comparison of datasets is a huge issue in online extremism research. This also leads to another problem of comparison of results. As no study uses the same dataset, comparing results with different methods and techniques is difficult in online extremism detection research.

2.2.6. Suspended Accounts. Social media has a strict policy on violence and hate speech [29, 56]. Thus, many accounts



FIGURE 1: Challenges in dataset.

with such extreme ideologies get suspended immediately. So even after data collection, other researchers cannot reproduce the results due to the unavailability of suspended accounts.

This work aims to address *data quality challenges*, *data validation*, *data imbalance*, and *binary classification* in extremism datasets. The challenges about languages and suspended accounts do not fall into the scope of this work.

2.3. Classifiers. Network-based, machine learning-based, and deep learning-based techniques are popularly used in online extremism research [19].

2.3.1. Network/Graph-Based Techniques. Network/graph-based techniques are preliminarily used due to the following reasons:

- (i) To cluster extremists on social media
- (ii) To identify extremist communities on social media
- (iii) To perform data collection by identifying connections among the extremists

Since 2015, only few studies use the network/graph-based approach. Agarwal and Sureka [57] used the *breadth-first search* and *shark search algorithms* to find the extremists and their communities on YouTube. The authors used the class name relevant (extremist) and irrelevant (nonextremist). By using the shark search algorithm, the authors achieved an accuracy of 0.74 and an F1 score of 0.85.

Saif et al. [58] used closegraph to extract subgraphs of extremists on Twitter. The authors used these subgraphs as features for machine learning algorithms such as Naïve Bayes, maximum entropy, and SVM. In addition to subgraphs, the authors used unigram, sentiment, and semantic features. The authors

concluded that SVM performs the best with a precision, recall, and F1 score of 0.93 for pro-ISIS and anti-ISIS classes.

Petrovskiy and Chikunov [59] also used graph techniques to extract features such as node page rank, hub and authority measure, and betweenness centrality. These features are then used as input for algorithms such as logistic regression, random forest, and XGBoost. The XGBoost algorithm outperforms other algorithms with a ROC curve of 0.95 for train and 0.94 for test data.

Moussaoui et al. [60] used a possibilistic graph for extremist community detection. Features such as semantic similarity, structural similarity, and possibilistic similarity are extracted using a possibilistic graph-based approach. The authors used subgraphs as features input to machine learning algorithms. The authors used *Naïve Bayes*, *multinomial Naïve Bayes* (MNB), and *stochastic gradient decent* (SGD) classifiers for extremism detection. SGD achieved a precision of 0.81 and an accuracy of 0.86 for extremism detection.

Network/graph techniques are used mostly to identify communications and interconnections but suffer from multiple challenges:

- (i) It cannot work for disconnected nodes in the graph
- (ii) Semantic analysis of extremism text cannot be performed with network/graph techniques

Thus, to overcome the network/graph approach challenges, machine learning-based and deep learning-based methods are used for online extremism detection.

Machine learning-based approach is used for the classification of data into extremist, nonextremist, or neutral [46, 61] or the classification of data into extremist and antiextremist [39, 42].

2.3.2. Machine Learning-Based Techniques. In machine learning-based approach different classifiers such as MNB [46], *logistic regression* [65], SVM [46], *random forest* [68], and *XGBoost* [71] are used for online extremism detection.

Agarwal and Sureka [64] used k-nearest neighbor and libSVM to identify hate-oriented text from Twitter. The authors used the term frequency as the feature. The authors got an accuracy of 0.97, a precision of 0.78, and a recall of 0.83.

Asif et al. [46] used MNB and *support vector classifier* (SVC) to classify Facebook posts and comments as *moderate*, *high extreme*, *low extreme*, and *random*. SVC performs better for the classification than multinomial Naïve Bayes, giving an accuracy of 0.82.

Benigni et al. [43] proposed *iterative vertex clustering and classification* (IVCC) for extremism detection. The authors also used *k-means*, *Louvain grouping*, and *Newman method* for extremism detection. The authors classify Twitter users into ISIS members, nonmembers, and suspended. IVCC outperforms other classification methods with an accuracy of 0.96 and an F1 score of 0.93.

Araque and Iglesias [36] used feature engineering by creating emotion features (EmoFeat) and similarity-based feature extraction (SIMON) methods. The authors labeled the data as positive (extremist) and negative. The authors got the highest F1-score of 0.94 for EmoFeat and SIMON, with the dataset containing extremist and neutral tweets.

Ashcroft et al. [42] used a *stylometric*, *sentiment*, and *time-based* feature for online extremism detection. The authors classify data into radical and nonradical. The authors used SVM, Naïve Bayes, and AdaBoost. AdaBoost gave a precision of 0.88, specificity of 0.99, and sensitivity of 0.79, with all the features outperforming other algorithms.

Fernandez et al. [35] divided extremists into individual (*micro*) influence, group (*meso*) influence, and global (*macro*) influence based on their tweets. The authors used the *collaborative filtering* and *Naïve Bayes* classification method. The authors used precision as a performance metric. Using Naïve Bayes, the precision obtained for micro is 0.79, for meso is 0.69, and for macro is 0.90.

Mussiraliyeva et al. [62] divided Kazakh language posts from VKontakte [55] into extremist and nonextremist classes. The authors used different classifiers such as logistic regression, MNB, and SVM. The authors also used decision tree-based classifiers such as random forest and gradient boosting. From all these classifiers, gradient boosting with word2vec gave the best F1 score of 0.86.

Mussiraliyeva et al. [54] used multiple features such as linguistic inquiry and word count (LIWC), part-of-speech (POS), and TF-IDF. The authors used numerous machine learning algorithms such as SVM, k-nearest neighbors (KNN), decision tree, random forest, Naïve Bayes, and logistic regression. The KNN using the oversampling method with statistical and TF-IDF features gives an accuracy of 0.99 for religious extremism classification.

Araque and Iglesias [51] used a combination of multiple features such as AffectiveSpace, SenticNet, TF-IDF, and SIMON. The authors used machine learning algorithms such as logistic regression and linear SVM.

De Smedt et al. [67] identified extremist hate speech within English, Arabic, and French language tweets. The authors used character trigrams as features. The tweets were labeled as hate and safe. The authors used libSVM as the classifier. The F1 score for the English language was 79, for French was 80, and for Arabic was 84.

Ul Rehman et al. [63] used *religious words*, *radical words*, and *bad words* to detect online extremism. The authors used two classes, extremist and nonextremists. The authors preferred different algorithms such as *Naïve Bayes*, SVM, and *random forest* for the classification. The SVM with all the features outperforms other algorithms with an F1 score of 0.87.

Sharif [61] divided tweets into *pro-Taliban*, *pro-Afghan*, *neutral*, and *irrelevant*. The authors used unigrams, bigrams, and TF-IDF for feature extraction. The authors also used

principal component analysis (PCA) to reduce dimensions. The research work used *Naïve Bayes*, SVM, and *random forest*. SVM with TF-IDF and bigrams offers the best precision of 0.84. Table 2 provides a comparison of all these studies in brief.

2.3.3. Deep Learning-Based Techniques. Even if machine learning-based approaches are popular, they face some challenges such as the following:

- (i) They depend heavily on manual feature extraction or feature engineering
- (ii) Not suitable for large and unstructured datasets
- (iii) Context identification is a challenge

These issues of machine learning methods can be addressed by using the deep learning approach. In the deep learning-based approach, the researchers have tried CNN [45], gated recurrent unit (GRU) [45], LSTM [65], and BERT [65].

A deep learning-based approach is used due to the following reasons:

- (i) Automated feature extraction
- (ii) Pretrained models on a large corpus

Recently deep learning approaches are routinely used in online extremism detection due to automated feature extraction and large computing power.

Kaur et al. [72] classified data into *radical*, *nonradical*, and *irrelevant* classes. The authors used word2vec for features extraction. Multiple algorithms such as SVM, *maximum entropy*, and *random forest* were used. The authors primarily focused on the deep learning approach using LSTM. LSTM with word2vec gives the best precision of 85.96.

Ahmad et al. [45] used n-grams, TF-IDF, and bag-of-words (BoW) as feature extraction methods for online extremism detection. The authors used the CNN model, LSTM model, FastText with word embedding, and GRU. The LSTM with CNN model offers an accuracy of 0.92 and a precision of 0.90 outperforming other algorithms.

Alatawi et al. [65] used BERT to detect hate speech related to White supremism on Twitter. The work used pretrained networks such as *Google News Word Vectors*, GloVe trained on Wikipedia, and GloVe trained on Twitter. The authors also train the extremist data using word2vec, referring to it as White supremacist word2Vec (WSW2V). BERT with WSW2V outperformed other techniques with an F1 score of 0.79 and a precision of 0.80. The direct comparison between approaches in online extremism detection is a problem. This is due to the use of different datasets, most of which are custom and not publicly available.

Mussiraliyeva et al. [73] in a recent study used CNN and LSTM to classify extremist posts collected from VKontakte. The CNN and LSTM both provide an AUC of 0.99 for extremism classification in the Kazakh language. Table 3 compares the studies employing deep learning for extremism detection.

2.4. Proposed Architecture. This section proposes the architecture for constructing the dataset, which will be used to classify extremism text into propaganda class, radicalization class, and recruitment class, with discussions on data validation methods. The architecture is modularized into the following phases: data collection, data preprocessing, data annotation, and data validation which are shown in Figure 2.

2.4.1. Data Collection. The construction of the proposed dataset was performed by collecting data from popular standard extremist text datasets and recent extremist tweets collected from Twitter.

2.4.2. Standard Dataset. In this phase, three different datasets were chosen, namely, ISIS Kaggle dataset (~15,000), StormFront dataset by de Gibert et al. (~1100), and Gab Hate Corpus by Kennedy et al. (~8000). Initially, these datasets were divided according to ideology, ISIS dataset as jihadist, while StormFront and Gab datasets as White supremacist. All these three datasets together contain around 24,900 extremist tweets. StormFront and Gab have two unique labels as hate and nonhate labels, while ISIS contains only extremist tweets. In addition, the StormFront dataset accounted for the posts between the years 2002 and 2017, while no data collection timeline is given for Gab dataset. Twitter was the preferred social media platform for collecting extremist tweets as it is the first choice for the extremists to reach out to the target audience. In addition, it is popularly used in research work [48, 67] due to its easy accessibility and microblogging format.

2.4.3. Data Extraction from Twitter. As the standard dataset has its challenges such as outdated text, as mentioned in the previous section, we collected recent extremism tweets from Twitter from January 2021 to June 2021.

Twitter API allows the collection of real-time tweets with different parameters. Twitter API provides a choice to collect tweets based on specific terms or hashtags, tweets of a specific user, tweets from a specific geographical area, and tweets of a specific language. Twitter APIs also give additional information such as username, location, and @user mentions in the tweet. Different queries were formulated, and the final query was selected as

$$\text{Query}[] = \text{Search}\{\text{search}_{\text{term}}, \text{time}\}. \quad (1)$$

To collect ISIS extremism text, specific keywords such as “*murtadeen*,” “*munafiqeen*,” “*khawarij*,” “*tafkir*,” “*kuffar*,” and “*murtad*” were used. These are popularly used ISIS-related words obtained from works such as [16, 41]. In addition, the keywords such as “*white genocide*,” “*white lives matter*,” “*it’s okay to be white*,” and “*anti-white*” were used to collect White supremacist-related tweets. These White supremacist supporting keywords were obtained from [74–76].

A total of 2,000 ISIS supporters and 2,000 White supremacist supporting tweets were collected. All these collected tweets are in the English language. Figure 3 provides keywords used and the wordcloud of hashtags found for White supremacist and jihadist-ISIS supporting tweets.

TABLE 2: Popular machine learning techniques employed in online extremism detection.

Technique used for extremism detection	Study	Hyperparameter	Features	Performance metric	Remark
Naïve Bayes/multinomial Naïve Bayes	[35, 42, 46, 60–63]	Alpha = 0.01 [46]	n-grams, TF-IDF, and word2vec	Accuracy = 0.66 [46] and correctly classified instances = 89% [42]	Naïve Bayes or multinomial Naïve Bayes is used so as to build a model based on a probabilistic learning approach [46]
KNN	[64]	Distance = euclidean distance, $K = 100$ [64]	Term frequency	Precision = 0.48 and accuracy = 0.90 [64]	Distance-based approach for similarity in extremism text
Logistic regression	[36, 62, 65]	NA	Word2vec, fasttext, GloVe, and LjWC	F1 score = 99.77 [36] and accuracy = 0.70 [62]	Used for binary classification of extremism text
SVM	[36, 42, 44, 46, 49, 61–64, 66, 67]	Penalty = L1, tol = $1e-3$ [46]	n-grams, TF-IDF, word2vec, fasttext, GloVe, and PCA	Accuracy = 84 [67] and precision = 84 [49]	SVM segregates data using hyperplanes, so that classification is better. [46]
Random forest	[61, 63, 66, 68, 69]	Estimators = 100, Kfold = 5 [68], estimators = 100, max_depth = 50 [66]	n-grams, TF-IDF, word2vec, and GloVe	Accuracy = 100 [66] and F1-score = 0.93 [69]	Random forest is scalable and unaffected by outliers in extremism text dataset [66]
AdaBoost	[41, 42, 70]	Boosting iterations = 300 [41]	n-grams	Precision = 0.88 [70] and accuracy = 99.5 [42]	AdaBoost improves performance by combining weak classifiers
XGBoost	[59, 71]	Regularization = L2	Betweenness centrality and page rank	ROC-AUC curve = 0.95 [59]	XGBoost improves performance with faster learning

TABLE 3: Popular deep learning techniques employed in online extremism detection.

Technique used for extremism detection	Study	Hyperparameters	Features	Performance metric	Remark
CNN	[20, 73]	—	Embedding layer	Accuracy = 0.70 [20]	CNN 1D performs better for text classification [20]
LSTM	[20, 65, 73]	Layers = 4, units = 300, loss = binary cross entropy, optimizer = Adam, epochs = 10, batch size = 256 [65]	GloVe, Word2vec	F1 score = 0.7489 [65]	LSTMs are capable of handling long-term dependencies to learn efficiently from longer sentences
LSTM + CNN	[45]	LSTM: units = 100, max_features = 2000, and activation = "relu" CNN: kernel size = 2×2 [45]	Embedding layer	Accuracy = 92.68, and precision = 88.32 [45]	LSTM + CNN provides better results than only CNN and only LSTM
BERT	[65]	Layers = 24, units = 1024, learning rate = $2e-5$, epochs = 3, and batch size = 32 [65]	Word2vec, GloVe pre-trained on Twitter,	F1 score = 0.79605 [65]	BERT is used to understand the semantic context of words better and suited for the natural language queries

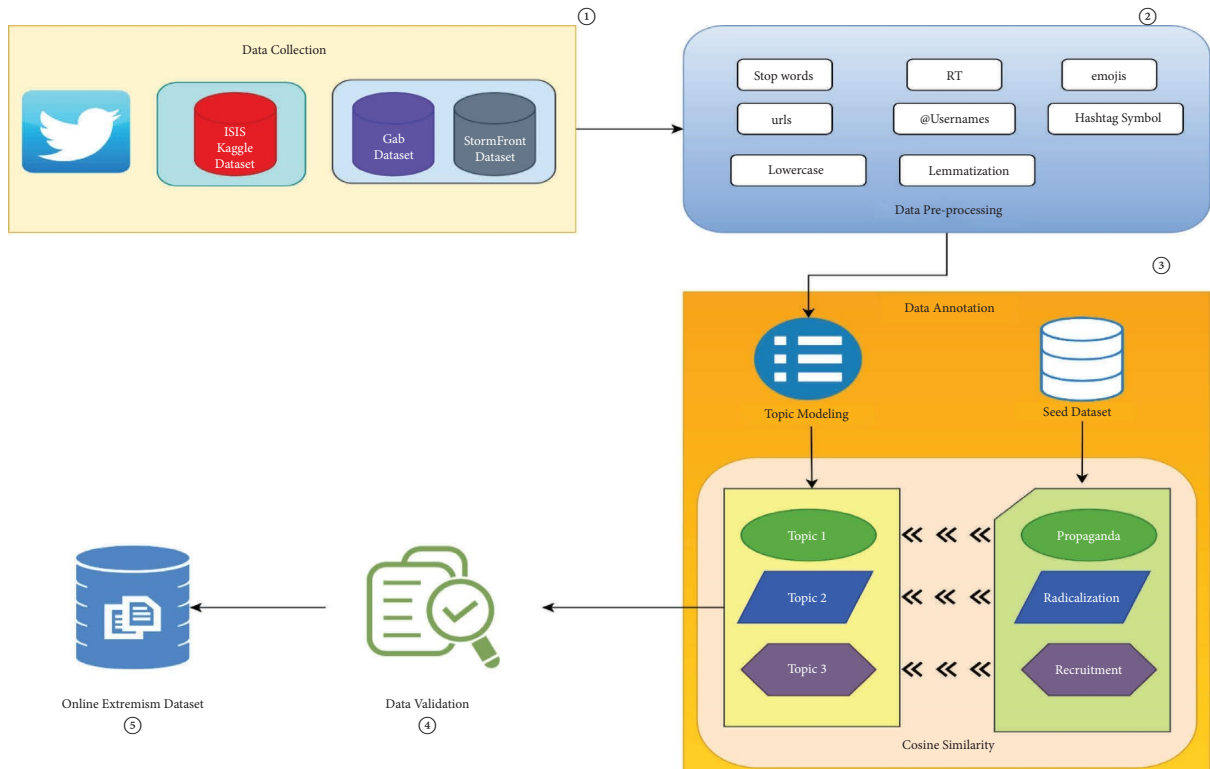


FIGURE 2: Proposed architecture.



FIGURE 3: Word cloud of hashtags for data collected for (a) White supremacists and (b) jihadist-ISIS with Twitter search terms.

2.4.4. SEED Dataset. One of these works aims to detect extremism and classify text into propaganda, radicalization, and recruitment. To achieve this, we collect examples of propaganda, radicalization, and recruitment from the existing literature. The collected examples are from both ideologies, jihadist-ISIS, and White supremacist.

Most of the examples from the literature [39, 46, 65] were manually annotated with fewer experts and are subject to bias. Hence, we extract examples from multiple resources [11, 16]. The assumption is that the seed example from different sources provided by different experts may reduce expert bias. A total of 100 examples were identified for jihadist-ISIS and 100 examples of White supremacists on propaganda, radicalization, and recruitment.

As the examples are taken from different research works, they have multiple keywords and different contexts associated with them, reducing the overall bias of the SEED dataset. In Table 4, a few examples are presented to show the tweets and posts considered propaganda, radicalization, and recruitment by respective studies.

2.4.5. Data Preprocessing. In this phase, data preprocessing is carried out in the following steps:

- (i) *Removing Stopwords.* Stopwords were removed at this step. Then, the words representing nouns, verbs, adverbs, and adjectives were selected. This ensured the inclusion of only relevant words in the final process

TABLE 4: Examples of propaganda, radicalization, and recruitment from the literature included in the SEED dataset.

Year	Study	Ideology	Propaganda	Radicalization	Recruitment
2015	Chatfield et al. [16]	Jihadist - ISIS	<p>“Coalition planes massacred these children in airstrikes on #Hit, #Anbar”</p> <p>“These PKK fellas are exceptional liars. After the city was almost fully abandoned by civilians, they now claim 55 thousand civilians there.”</p>	<p>“Himalaya888 you do realize IS wants to destroy every single nation-state, Arab or Kurd or communist does not matter, that they come across?”</p> <p>“Real warriors, I mean those who are not fags, dont need air support.”</p>	<p>“This is the time for muslim kurds in Turkey to show whether they can ever counter PKK.”</p> <p>“EbuRuana so if it is not right to make dua (the Islamic act of calling out to Allah) for a kafir (a disbeliever, someone who rejects Allah and who does not believe in Muhammad as the final messenger of Allah), what is the right thing to do ?”</p>
2001	Ray and Marsh [77]	White supremacist	<p>“Most victims of race crime, about 90 percent, are White. Almost 1 million White Americans were murdered, robbed, assaulted, or raped by black Americans in 1992, compared with about 132,000 blacks who were murdered, robbed, assaulted, or raped by Whites. Blacks are committing more than 50 times the number of violent racial crimes of Whites”</p>	<p>“He is someone who loves his race and his land more than he is concerned for his own welfare. He is neither an intellectual snob nor a street brawler. He is intelligent and educated enough to know what he is fighting for and brave enough to stand up for his beliefs when push comes to shove. He is someone who practices what he preaches, and who backs up his words with his deeds. He is definitely not someone who is long on talk and short on action.”</p>	<p>“(W)e need to take back America and the streets. I feel that Whites should be the predominant force in America. Whites can hardly walk down the street anywhere without being threatened. We must unite.”</p>
2020	Johnson [11]	White supremacist	<p>“America is dead! long live America!”</p>	<p>“Only right is right of lead”</p>	<p>“Join or die with rest”</p>

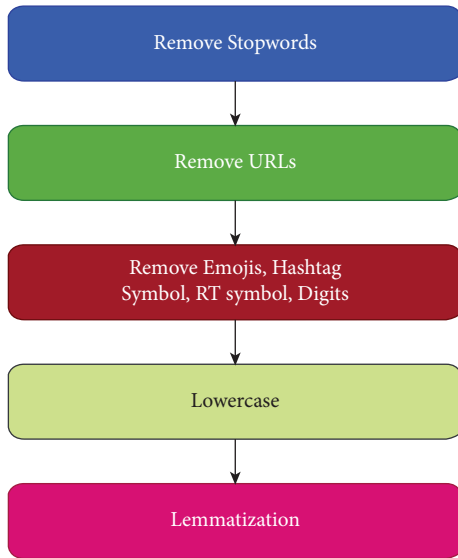


FIGURE 4: Data preprocessing.

- (ii) *Removal of URLs.* URLs were removed. Some studies do use URLs for further analysis. However, with standard datasets, many URLs are obsolete. Hence, the inclusion of URLs is not considered in this study
- (iii) *Removal of Emojis, Hashtag Symbols, Retweet Symbols (RT), And Digits.* Hashtag symbols and RT symbols are not the focus of this study. Numbers and digits may interfere with word analysis, hence are excluded
- (iv) *Removal of @username Mentions.* Due to constant communication between users, mention of usernames is fairly common. This may help algorithms to construct the pattern with usernames to build linkage
- (v) *Lowercase.* All words are converted to lowercase so that case of the alphabet does not affect the prediction results
- (vi) *Lemmatization.* Lemmatization of texts is also performed so that pronouns and the tense of words may not affect the final prediction

The preprocessing steps are illustrated in Figure 4.

2.5. Data Labelling

2.5.1. Topic Modelling. Topic modelling is a method to recognize, understand, and summarize a large collection of textual information. Topic modeling is a way to extract a group of words (topics) that accurately represent the collection of documents in a corpus. It is also a form of text mining in which word patterns in a corpus are identified.

2.6. Latent Dirichlet Allocation (LDA). LDA is a probabilistic topic modeling algorithm, which extracts topics from documents, and words in the document are collected by observing their probabilistic distribution.

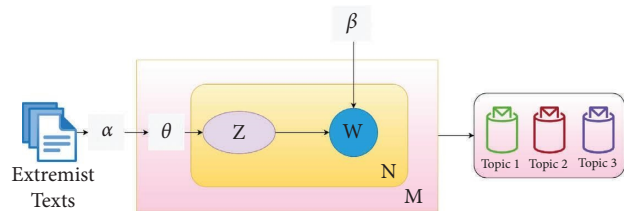


FIGURE 5: LDA model.

There are different techniques other than LDA to identify abstract information from a corpus. *Latent semantic analysis* (LSA) [78] and *probabilistic latent semantic indexing* (pLSI) [79] are some of them.

LDA focuses on topic identification and analysis, while LSA focuses on reducing matrix dimensions. LSA converges faster due to dimensionality reduction but at the expense of accuracy. pLSI uses a probabilistic model with dimensionality reduction and is faster with acceptable accuracy. Top2Vec is a recent development in finding topics within the documents. Top2Vec [80] has considerable advantages over LDA such as no need for stopword removal, stemming, or lemmatization. BERTopic [81] too has advantages such as deep learning and visualization. But both Top2Vec and BERTopic require a good amount of data which is a limitation of our study. In addition, LDA is preferred as we need a specific number of topics. Moreover, LDA is used in multiple studies for extremism detection, thus making LDA reliable for extremism detection research.

LDA assumes the mixture of the probabilistic distribution of topics over corpus and words over the topic. LDA works in the following ways as shown in Figure 5:

- (i) Assume there are k topics over the entire corpus
- (ii) Distribute k topics across document M which is per-document topic distribution also denoted as α . The topic distribution for document M is denoted as θ
- (iii) Calculate z which is the topic of n^{th} word in document M , while N is the number of words in the given document
- (iv) Calculate the probability of word w which belongs to a particular topic based on the following:
 - (a) Unique topics in document M .
 - (b) The frequency of the word w that has been assigned to a particular topic across all documents is also denoted as β .

For this study, it is needed to identify different topics within the extremism corpus. Later, these topics are compared for the labeling of extremist texts. So, LDA is used to extract topics from the extremism corpus due to its advantages as mentioned above and as described in Figure 5.

2.6.1. Cosine Similarity. Cosine similarity computes the similarity between vectors. It calculates the cosine of the angle between vectors and determines whether vectors point in the same direction. In NLP, cosine similarity is commonly used to measure the similarity between the extracted features. Cosine similarity takes a total length of vectors; for example, considers TF-IDF vectors, thus considering repetitions of the word [82].

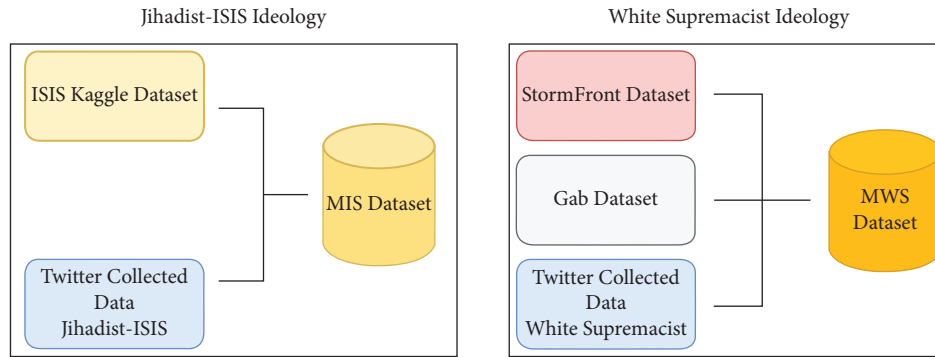


FIGURE 6: Datasets and their combinations.

This property is used to identify unique words for a particular class in this work. So, cosine similarity is considered for assigning labels from SEED datasets to primary datasets.

In this work, data labeling is designed to be a four-step process and the steps are described as follows:

(1) *Step 1.* In the first step, datasets are merged according to ideology. The ISIS Kaggle dataset was merged with recent tweets of jihadist-ISIS collected from Twitter, referred to as the Merged ISIS dataset (MIS). Similarly, StormFront dataset, Gab dataset, and White supremacist tweets collected from Twitter merged to form Merged White Supremacist dataset (MWS). This process is shown in Figure 6. Only the text or tweet data is selected from these standard datasets, everything else is discarded. To preserve the distinct characteristics of ideology, we adopt the strategy to identify individual clusters within the ideological datasets. To identify these clusters, the topic modelling approach was chosen [83]. For feature extraction, TF-IDF is used. TF-IDF calculates important words in the corpus concerning documents. However, even if TF-IDF presents important words, it lacks in identifying context. So, to extract topics from the primary dataset, latent dirichlet allocation (LDA) [83] is used. This work aims to classify text into three classes: *propaganda*, *radicalization*, and *recruitment*; three topics are extracted from the MIS and MWS datasets. To achieve this, GridSearchCV [84] is applied to the LDA model with hyperparameters such as $n_topics = [3-5]$, $learning_rate = [0.999, 0.99999]$, $cv = 10$, and $batch = \text{“online.”}$ Using these hyperparameters, the model with the best results gives n_topics of 3 with distinct words per topic.

(2) *Step 2.* In the second step, we extract a single topic for propaganda, radicalization, and recruitment examples for each SEED dataset of jihadist-ISIS and White supremacist ideology using LDA. This results in a single topic with respective important words in propaganda, radicalization, and recruitment. Figures 7(a)–7(c) show the word clouds of three topics obtained from the MIS dataset. Similarly, Figures 8(a)–8(c) show word clouds of the three topics obtained from the MWS dataset. These word clouds are based on the topic score obtained using LDA for MIS and MWS datasets, as shown in Figures 9 and 10.

(3) *Step 3.* To label text in the IS dataset and the WS dataset, cosine similarity [85] between the topics of individual MIS and MWS datasets, with the topic of propaganda, radicalization, and

recruitment from SEED dataset, is calculated. This results in similarity matrix. When similarity is maximum for topic and label, the respective label, propaganda, radicalization, and recruitment, is assigned to a particular topic. Thus, documents in IS and WS datasets with the topics labeled are propaganda, radicalization, and recruitment. Figure 11 shows the complete process of data labeling. The calculated cosine similarity between seed labels and identified topics is small. There are different reasons for low cosine similarity, such as few seed examples, and not enough significant features in SEED dataset. This low cosine similarities are accepted as two different datasets i.e., SEED dataset and tweet + website dataset are compared.

This research work aims to develop an ideology independent extremism detection model. So, to achieve this aim, two datasets MIS and MWS datasets, are merged. This is carried out by retaining tweets or posts, topics, ideology, and labels from both datasets. This merged dataset will be henceforth referred to as **Merged ISIS-White Supremacist dataset (MIWS)**. As seen in Table 5, for the MIS dataset, topic 0 is labeled as propaganda, topic 1 as radicalization, and topic 2 as recruitment, as significant cosine similarity was found with the respective classes in the SEED ISIS dataset. On the other hand, in the MWS dataset, topic 0, topic 1, and topic 2 are labeled as radicalization, recruitment, and propaganda as a significant similarity score was found with respective classes of the SEED White Supremacist dataset. Figures 12(a)–12(c) can provide important words in the MIWS dataset for propaganda, radicalization, and recruitment.

2.7. Data Validation (MIWS). In this Section, we discuss the statistical tests, which will be employed for the data quality assessment. We employed three statistical techniques that are *cosine similarity*, *Wilcoxon signed-rank test*, and *chi-square test*.

2.7.1. Cosine Similarity. Cosine similarity can be used to compare the similarity between samples. Propaganda, radicalization, and recruitment are compared based on words and their TF-IDF score. The cosine function was applied to a pair of classes. These pairs are described in Table 6. Thus, each class is represented by distinct unique words, and they influence each class differently. Figure 11 shows cosine



FIGURE 7: Word clouds for (a) topic 0, (b) topic 1, and (c) topic 2 in MIS dataset.

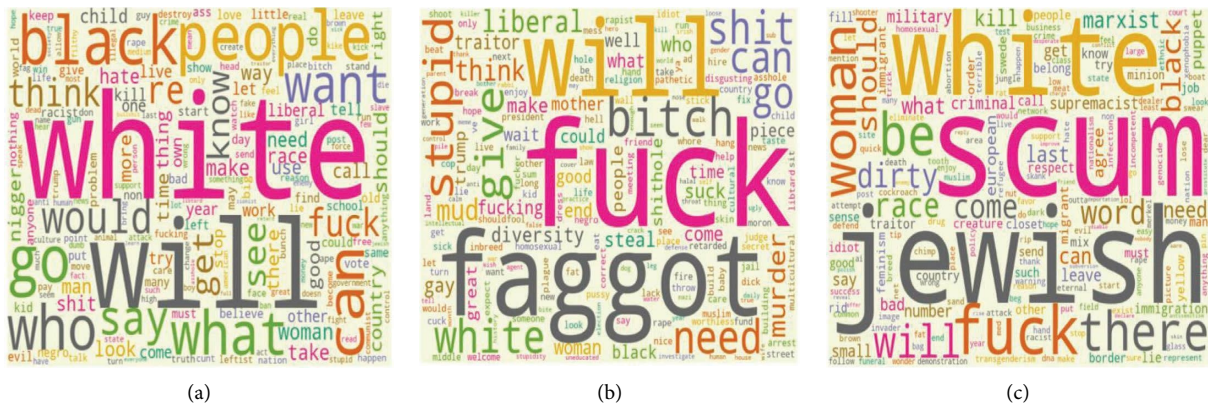


FIGURE 8: Word cloud for (a) topic 0, (b) topic 1, and (c) topic 2 in MWS dataset.

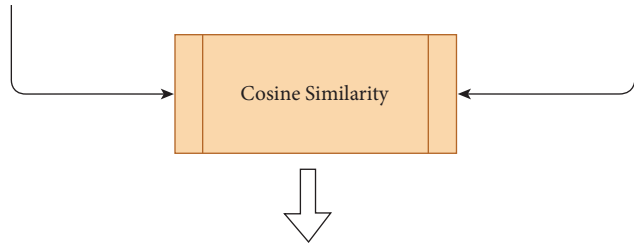
	abandon	abcnew	abdicate	abduct	abduction	abdul	abet	abhor	abhorrent	abide	...	zogworld	zombie	zone	zoo	zoom
Topic0	3.273194	0.889721	0.639427	1.058252	0.953454	0.713969	1.309879	0.930022	1.249693	1.599947	...	0.699101	2.715960	4.784419	3.142445	0.594423
Topic1	0.460962	0.381595	0.366446	0.364414	0.366880	0.364989	0.533811	0.362390	0.377678	0.365584	...	0.359935	0.375668	0.363517	0.373992	0.366417
Topic2	0.361743	0.384098	0.366361	0.371622	0.387797	0.365526	0.363269	0.359132	0.363632	0.361046	...	0.357538	0.376902	0.390312	0.413396	0.361322

FIGURE 9: LDA ranking of jihadist-ISIS words for three topic.

	abandon	abdicate	abduct	abet	abhor	abide	ability	ability	abject	able	...	yuk	ywnru	zealot	zionist	zogworld	zombie
Topic0	1.587991	0.866239	0.633916	0.425799	0.593244	0.648842	0.965337	0.595928	0.605659	3.343095	...	0.794940	0.596163	0.710336	7.424973	0.441100	0.452584
Topic1	0.409428	0.414233	0.418989	0.736551	0.406314	0.923721	0.407375	0.401482	0.404555	1.258568	...	0.437793	0.421771	0.400469	0.588412	0.413378	0.696941
Topic2	0.413395	0.419147	0.456127	0.407315	0.429035	0.422535	0.433495	0.423666	0.418037	1.174490	...	0.433062	0.439783	0.464315	0.675779	0.752007	0.785375

FIGURE 10: LDA ranking of White supremacist words for three topics.

LDA Topics of Examples from Seed Dataset								LDA Topics from IS Dataset/WS Dataset							
	ability	abstract	acceptable	accurate	acquire	act	action		abandon	abcnew	abdicate	abduct	abduction	abdul	abet
Radicalization	1.093294	1.093294	1.295743	1.093294	1.132283	1.380403	1.341120	Topic0	3.273194	0.889721	0.639427	1.058252	0.953454	0.713969	1.309879
Recruitment	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.326761	Topic1	0.460962	0.381595	0.366446	0.364414	0.366880	0.364989	0.533811
Propaganda	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	1.098419	Topic2	0.361743	0.384098	0.366361	0.371622	0.387797	0.365526	0.363269



	Propaganda	Radicalization	Recruitment
Topic 0	0.1954	0.1345	0.1895
Topic 1	0.1267	0.1498	0.2089
Topic 2	0.1657	0.1786	0.1421

FIGURE 11: Data labeling.

TABLE 5: Examples from the Merged dataset.

Sr. no	Tweet	LDA topic	Ideology	Label
1	“Did not i tell you that JN only make takfir on those who spill their holy blood? SRF, hazm?”	0	Jihadist-ISIS	Propaganda
2	“Mujahideen from Burma capture 3 pigs of the Buddhist army”	1	Jihadist-ISIS	Radicalization
3	“Our prophet, has ordered us to fight you till you worship allah alone or give jizya”	2	Jihadist-ISIS	Recruitment
4	“They act as if negros with aids should have the right to reproduce when they will die off and have no way to take care of the kid I cannot believe that they are doing this by the millions its crazy”	0	White supremacist	Radicalization
5	“You are white and you are better than them and the next time they harass you and someone else form a group of buddies, go up to the principal office.”	1	White supremacist	Recruitment
6	“It is not right unless it is white.”	2	White supremacist	Propaganda

similarity between different datasets, while in Table 6, similarities are seen within classes of the same dataset. Thus, even if values in Table 6 look significant, there is not enough similarity within the dataset given the N1 and N2 sizes.

2.7.2. *Wilcoxon Signed-Rank Test.* Wilcoxon signed-rank test [86] is a nonparametric test. It can determine whether the two samples are collected from the population of the same distribution. Wilcoxon signed-rank test is also used to compare two closely related samples and perfectly matched samples.

In this paper, Wilcoxon signed-rank test is used to prove whether the selected random samples belonged to a particular class, i.e., propaganda, radicalization, or recruitment. Figure 13 shows detailed experiments performed to calculate the Wilcoxon signed-rank test. CountVectorizer [87] was applied for feature extraction to the corpus of each class separately. CountVectorizer returns the matrix with the count of tokens. This was performed so that higher count words from each corpus may get priority. TfidfVectorizer [88] was also considered for this experiment but leads to a dimensional mismatch for Wilcoxon signed-rank test. The

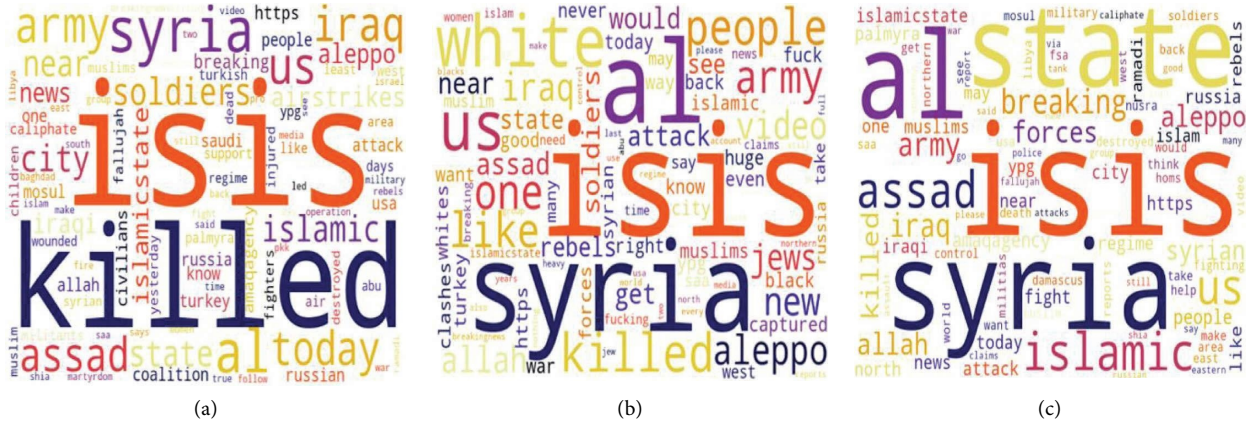


FIGURE 12: Word clouds for (a) propaganda, (b) radicalization, and (c) recruitment class from MIWS.

TfidfVectorizer also produces p values >0.05 when dimensions are matched.

To perform these experiments, a null hypothesis is required, which is as follows:

H0-medians of word count of classes are equal. Therefore, there is no significant difference between classes

H1-medians of word count of classes are not equal. Therefore, there is a significant difference between the classes

Wilcoxon signed-rank test compares examples based on two test statistics. First, W test statistics which is the sum of ranks with differences below or above zero. The second is the p value which is the confirmation against the null hypothesis. Together, W and p value determine the validity of the null hypothesis.

To calculate W , the following procedure is performed:

Let N be the sample size, and for pairs, let $x_{1,i}$, and $x_{2,i}$ denote the measurements.

- (i) Calculate $|x_{2,i}-x_{1,i}|$ and $\text{sgn}(x_{2,i}-x_{1,i})$, where sgn is the sign function that returns the sign of a real number
- (ii) Exclude the pair with $|x_{2,i}-x_{1,i}|=0$, and the new sample will be N_r
- (iii) Order the remaining pair in an ascending order with a difference of $|x_{2,i}-x_{1,i}|$
- (iv) Rank the pairs with the smallest nonzero difference as 1. Let R_i denote the rank
- (v) The test statistic W is calculated as

$$W = \sum_{i=1}^{N_r} [\text{sgn}(x_{2,i} - x_{1,i}) \cdot R_i]. \quad (2)$$

The p value is considered as the evidence against the null hypothesis. The null hypothesis is rejected if the p value is <0.05 . This threshold of 0.05 or 5% is considered a level of significance. The count for each word representing classes is calculated.

As the classification is a multiclass classification, the tests are divided into different cases which are as follows:

- (i) Case 1: here, the propaganda class and recruitment class are compared using CountVectorizer of n number of words from both classes
- (ii) Case 2: here, radicalization class and propaganda class are compared using CountVectorizer of n number of words from both classes
- (iii) Case 3: here, recruitment class and radicalization class are compared using CountVectorizer of n number of words from both classes

Table 7 shows cases, their samples, test statistics, hypothesis, and inference. The Wilcoxon signed-rank test provides test statistic “ W ” which is used to calculate the p value from the reference table [86].

2.7.3. Chi-Square Test. The chi-square test is a popular statistical test used to evaluate the relationship between two variables [89]. Most of the time, the chi-square test is applied to test the dependence of the occurrence of the term and the occurrence of the class. Moreover, it is commonly used as a feature selection method. For example, the following formula is used to calculate the rank of terms that appear in the corpus:

$$\chi^2(D, t, c) = \sum_{e_i \in \{0,1\}} \sum_{e_c \in \{0,1\}} \frac{(N_{e_i e_c} - E_{e_i e_c})^2}{E_{e_i e_c}}. \quad (3)$$

Here, e_t and e_c are binary variables in the contingency table, t is the term, c is the class, D is the corpus, N is the observed frequency, and E is the expected frequency. The term t and class c are said to be dependent if χ^2 is high. Thus, making term t an important feature that causes term t to indicate class c .

Table 8 shows important words within ISIS SEED, WS SEED, and MIWS datasets obtained by applying the chi-square test. Each dataset has a few repeated words. This can be attributed to different ideologies, sources, and dataset sizes.

TABLE 6: Cosine similarity for classes.

Case	Classes	No. of examples (N1)	No. of examples (N2)	Cosine similarity	Inference
Case 1	Recruitment and propaganda	9,669 Recruitment	10,046 Propaganda	0.2857	Words representing classes are sufficiently dissimilar
Case 2	Radicalization and propaganda	13,937 Radicalization	10,046 Propaganda	0.29436	Words representing classes are sufficiently dissimilar
Case 3	Radicalization and recruitment	13,937 Radicalization	9,669 Recruitment	0.2993	Words representing classes are sufficiently dissimilar

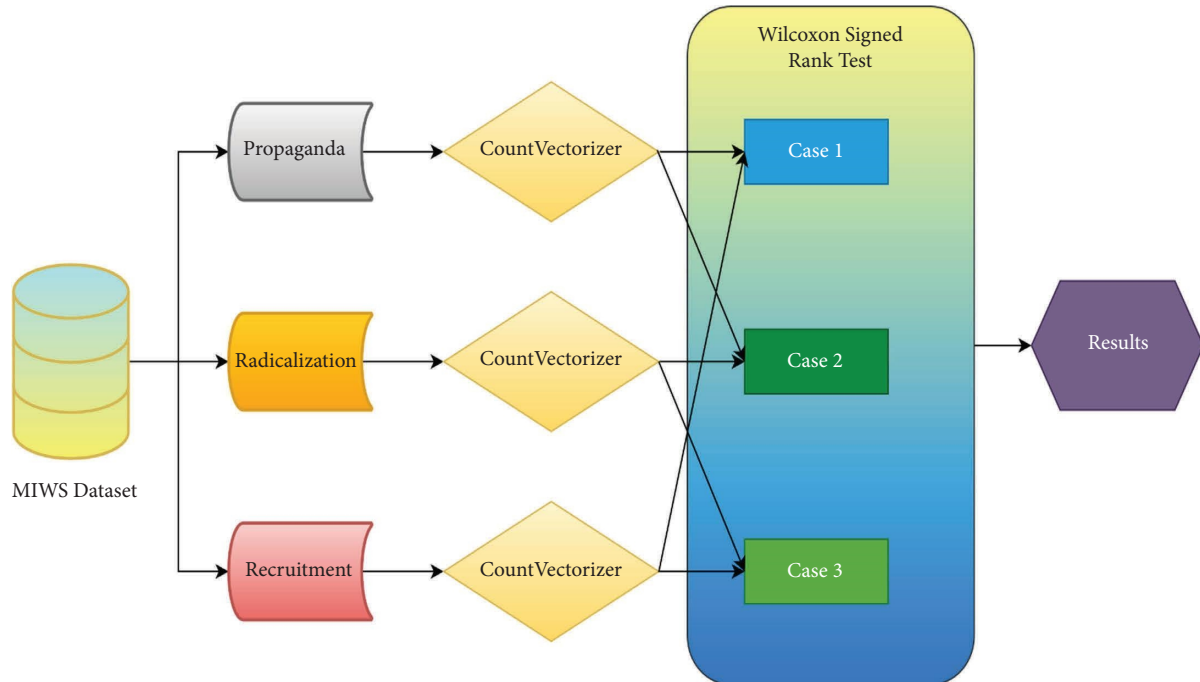


FIGURE 13: Complete process for performing Wilcoxon signed-rank test.

2.8. Inferences. As seen from Table 6, cosine similarity proves that the obtained classes, namely, propaganda, radicalization, and recruitment are significantly different. The Wilcoxon signed-rank test also shows significant differences between the classes, so they have distinct features to make them unique. The chi-square test in Table 8 shows distinct word features to depict propaganda, radicalization, and recruitment.

Thus, it can be inferred that the newly formed classes propaganda, radicalization, and recruitment stand unique with statistical validation methods.

2.9. Dataset Evaluation

2.9.1. Experimental Setup. Experiments were carried out on the HP Workstation Z8 G4 machine. It is equipped with a Xeon processor of 3 GHz, 128 GB of RAM, and Nvidia Quadro P400 GPU with 2 GB memory. In addition, some experiments were carried out on Nvidia DGX-Server with 4 Nvidia Tesla V-100 GPUs with 32 GB memory. Due to the limited capability of these systems, Google Colab was used. All the results in Table 9 are obtained on Google Colab.

2.9.2. Size of Datasets. The size of datasets are provided in Table 10.

2.9.3. Analyzing Imbalance in Datasets. The balance and imbalance in datasets are shown in Table 11.

2.10. Feature Extraction Techniques. To create word vectors, different feature extraction techniques are used in online extremism. In this work, the following feature extraction techniques are used:

2.10.1. Unigram with TF-IDF. As seen in Table 9, the TF-IDF is used as the feature extraction technique. TF-IDF gives important words in the document based on its weightage in corpus [90]. Thus, TF-IDF was chosen, as it shows the word importance and is also used in many studies. Unigrams are considered to identify and elevate the importance of unique words representing the particular class, propaganda, radicalization, or recruitment.

2.10.2. Bigrams and Trigrams with TF-IDF. Bigrams and trigrams features are used with TF-IDF for more complex analysis. These features provide the combination of words that affect the classification of the documents.

2.10.3. Word2Vec. Word2vec uses a neural network to learn word embeddings or word vectors from the given corpus. Word2vec is used to gather more dimensional features to classify extremism text into propaganda, radicalization, and recruitment. The word2vec model pretrained on Google News with 300 dimensions was used for feature extraction in this work. Figure 14 shows word vectors and their positions concerning each other using t-sne. Euclidean distance is used as a metric to calculate the distance between features. Thus, the lesser the Euclidean distance the more frequently the words appear together in a group. In Figure 14 it can be seen extremism influencing words are close to each other. Words such as "islamic state," "dead," "Afghanistan," "wounded," and "targeted" form a group. It can be also observed "bomb," "raqqa," "destruction," "gaza," "terror," "attack," and "battle" indicates the focus of groups on a particular location. The words such as "white," "muslims," "muslim," and "black" stood out from other keywords indicating their usage in different contexts. Thus, word2vec can be effectively used for online extremism detection.

TABLE 7: Cases, test statistics, and inferences.

Case	n	W	p value	Hypothesis testing	Inference
Case 1	19 and 325	46291754.5	$1.023e-18$	Reject null hypothesis	Classes propaganda and recruitment differ considerably from each other
Case 2	19 and 325	51169402	$6.181e-64$	Reject null hypothesis	Classes propaganda and radicalization differ considerably from each other
Case 3	19 and 325	61339810.5	$2.441e-16$	Reject null hypothesis	Classes recruitment and radicalization differ considerably from each other

TABLE 8: Important words obtained by using chi-square.

Sr no	Propaganda			Radicalization			Recruitment		
	ISIS SEED	WS SEED	MIWS	ISIS SEED	WS SEED	MIWS	ISIS SEED	WS SEED	MIWS
1	Arab	Backwards	Aspect	Matter	Race	Killed	Bless	Adapt	Coalition
2	Massacre	Alien	Islamic	Call	Attack	Airstrikes	Counter	Accept	Martyrdom
3	People	Based	Beasts	Nation	Purity	State	Pkk	Stand	Behead
4	Assadis	Aboriginals	Allahu	Destroy	Scripture	Communist	Behead	Student	Adapt
5	America	Auschwitz	Apostate	Bullet	Call	Caliph	Believe	School	Munafiq
6	Apostate	Beasts	Assadis	Communist	Chosen	Amaqagency	Achieve	Friend	Believe
7	Possible	Aspect	Black	Single	Body	Race	Protect	Apologist	Iraqi
8	Babylon	Base	Israeli	Caliph	Believe	Nation	Place	Many	Accept
9	Back	Armed	Arab	State	Resist	Assault	Pledge	Antifa	Join
10	Photographer	Other	Operation	Realize	Revelation	Saudi	Point	White	Full

TABLE 9: Algorithms, features, and performance.

Sr no	Algorithm	Features	MIS			MWS			MIWS		
			Precision	Recall	F1score	Precision	Recall	F1 score	Precision	Recall	F1 score
1	MNB	TF-IDF	0.67	0.67	0.67	0.73	0.76	0.74	0.61	0.61	0.61
		TF-IDF + bigrams	0.69	0.68	0.69	0.73	0.76	0.74	0.62	0.62	0.62
		TF-IDF + trigrams	0.64	0.64	0.64	0.72	0.75	0.73	0.60	0.60	0.60
		Word2vec	0.49	0.38	0.32	0.61	0.76	0.66	0.46	0.46	0.46
2	SVM	TF-IDF	0.71	0.70	0.70	0.74	0.77	0.70	0.69	0.68	0.68
		TF-IDF + bigrams	0.74	0.47	0.37	0.58	0.76	0.66	0.75	0.41	0.32
		TF-IDF + trigrams	0.75	0.45	0.35	0.58	0.76	0.66	0.76	0.38	0.27
		Word2vec	0.51	0.50	0.50	0.70	0.64	0.66	0.54	0.53	0.53
3	Random forest	TF-IDF	0.62	0.61	0.60	0.66	0.73	0.68	0.63	0.64	0.63
		TF-IDF + bigrams	0.69	0.69	0.68	0.58	0.76	0.66	0.61	0.61	0.60
		TF-IDF + trigrams	0.61	0.59	0.58	0.58	0.76	0.66	0.59	0.57	0.53
		Word2vec	0.56	0.56	0.56	0.72	0.77	0.69	0.53	0.53	0.52
4	XGBoost	TF-IDF	0.60	0.59	0.59	0.70	0.76	0.70	0.61	0.62	0.62
		TF-IDF + bigrams	0.58	0.58	0.58	0.69	0.76	0.70	0.60	0.60	0.59
		TF-IDF + trigrams	0.58	0.58	0.57	0.69	0.76	0.70	0.59	0.60	0.59
		Word2vec	0.59	0.59	0.59	0.75	0.78	0.73	0.60	0.60	0.60

TABLE 10: Size of datasets.

Datasets	Source	Ideology	Total tweets/posts
Twitter tweets	Twitter	White supremacist	2,000
Twitter tweets	Twitter	Jihadist-ISIS	2,000
ISIS Kaggle dataset [22]	Twitter	Jihadist-ISIS	~15,000
StormFront [20] + Gab dataset [21]	StormFront and Gab	White supremacist	~9000 (only hate class)

TABLE 11: Balance and imbalance in datasets.

Datasets	Number of classes	Class 1	Class 2	Class 3
ISIS Kaggle dataset	1	15,438	—	—
StormFront dataset	2	1180 (hate)	8,537 (nohate)	—
Gab dataset	2	8,327 (hate)	~25,000 (other)	—
MIS	3	7,214 (propaganda)	5,103 (radicalization)	51,21 (recruitment)
MWS	3	5,131 (propaganda)	3,214 (radicalization)	3,162 (recruitment)
MIWS	3	12,345 (propaganda)	8,317 (radicalization)	8,283 (recruitment)

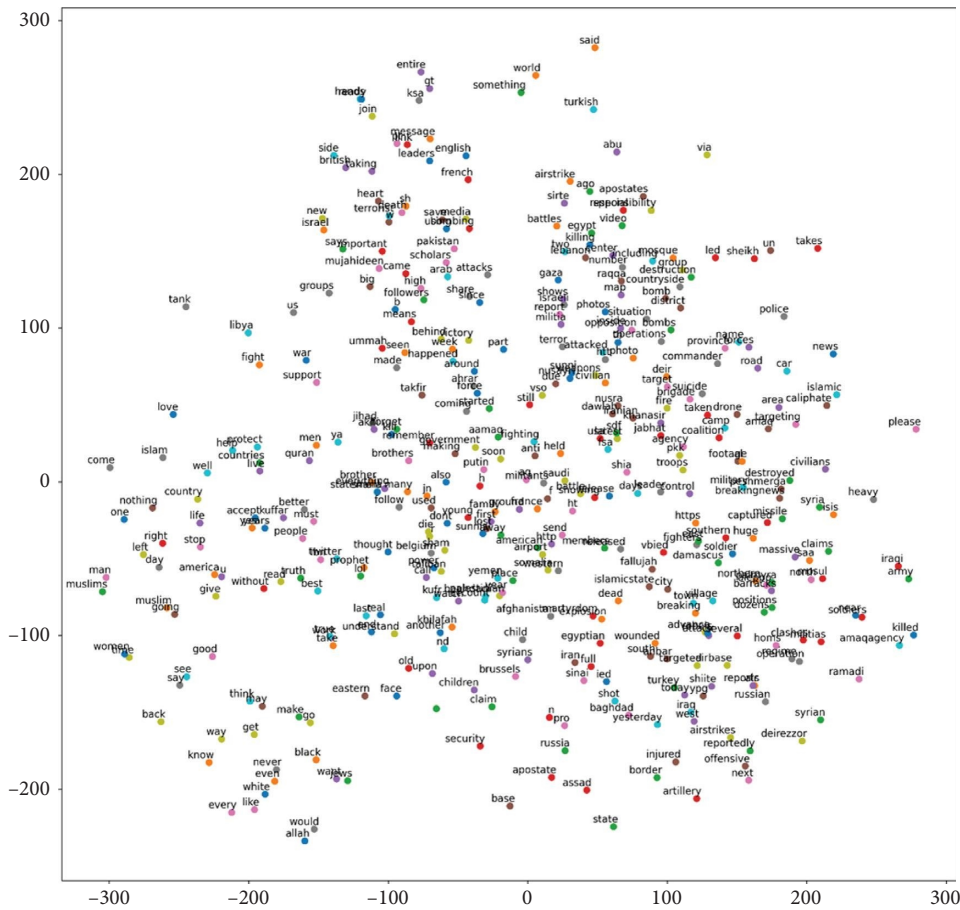


FIGURE 14: Word vectors using t-SNE and word2vec on MIWS dataset.

Word2vec is used in combination with classifiers mentioned in the next section. Word2vec is fine-tuned to a window size of 15, a minimum count of 10 words, and with ten iterations to provide the best possible performance metrics.

2.11. *Classifiers.* To classify and predict, this work uses the following ML algorithms:

2.11.1. *Multinomial Naïve Bayes.* MNB works on the probabilistic principle. Naïve Bayes assumes that there exists a conditional independence between every pair of features. In addition to this MNB, also assumes that distribution for all pair is multinomial distribution. This assumption of multinomial distribution works well in the case of word counts in the document. Thus, classifying text data based on the probabilistic appearance of a word within the document helps to get a baseline for performance metrics.

2.11.2. *Support Vector Machine.* In online extremism detection, SVM can separate important words of a particular group or class by defining the exact separation line. This separation line is referred to as a hyperplane. SVM creates support vectors that are at the optimal distance from the hyperplane. This ensures the words of a particular group are at a significant distance from

words of another group. So, one can get fairly accurate performance metrics due to this property of SVM.

2.11.3. *Random Forest.* Random forest uses multiple decision trees to classify data. Every decision tree consists of decision nodes, root nodes, and leaf nodes. Thus, every decision tree in random forest is trained on a subsample of the dataset. Thus, each tree is ensured to be built upon the best subset of features. It takes the majority output of the decision trees to arrive at the classification. This reduces overfitting, thus making random forest a good choice for the extremism text classification.

2.11.4. *XGBoost.* XGBoost uses gradient boosting for the classification. In XGBoost, gradient boosting is achieved by pruning trees backward that exceed the maximum depth of tree criteria, thus, increasing the speed of the algorithm by employing the depth-first technique. XGBoost can also work with a small amount of data. XGBoost also supports out-of-core computing, that is, it can handle data more than disk space and memory. Another advantage of XGBoost is, it provides parallelization, thus making the classification process faster.

Figure 15, provides details about the ML pipeline for the best-fit model. In this pipeline, the MIWS dataset with

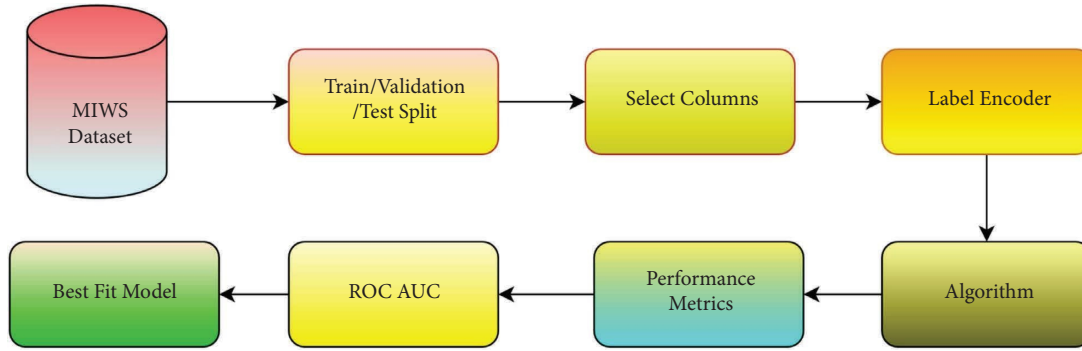


FIGURE 15: ML pipeline.

TABLE 12: Hyperparameters used for fine-tuning ML algorithms.

Algorithm	CV	Learning rate/gamma/alpha	C (max depth)	Kernel (max_features)	n_estimators
MNB	10	0.1, 0.5, and 1	—	—	—
SVM	10	$1e-3$, $1e-4$, and $1e-5$	1, 2, and 3	linear, poly, and rbf	—
Random forest	10	—	300, 350 , and 400	auto and sqrt	80 , 90, and 100
XGBoost	10	0.1 , 0.01, and 0.001	4, 5 , and 6	—	500, 550 , and 600

Bold values are the optimal hyperparameters for the respective algorithm.

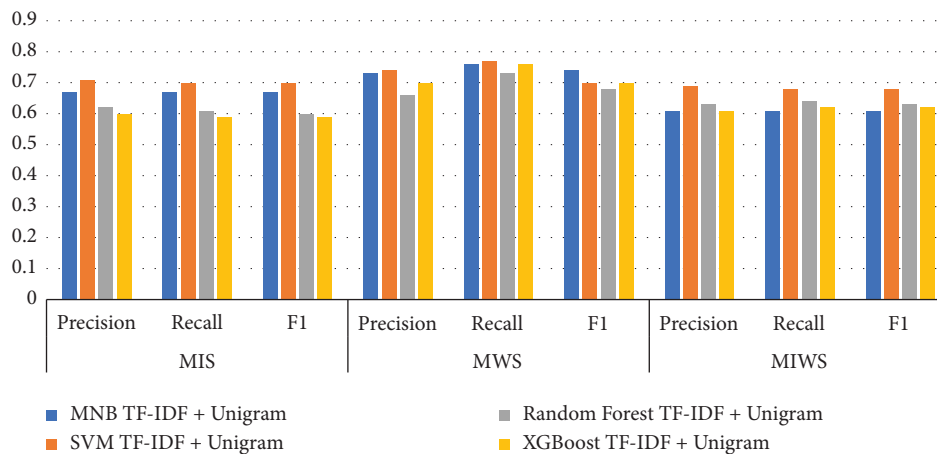


FIGURE 16: Performance metrics of algorithms on all datasets with TF-IDF unigram.

preprocessed data is taken as input. Table 10 shows the count of tweets while Table 11 describes data imbalance in datasets used in this study. Then, train/validation/test split is performed on selected data. Different split ratios are used such as 60:20:20, 70:15:15, 80:10:10, and 90:05:05. The better results were obtained for the 90:05:05 split. Particular columns such as preprocessed text and labels are selected for classification. As the labels are in string format, a label encoder is used to convert labels into the numerical format. Multiple ML algorithms as mentioned before are provided with GridSearchCV. The hyperparameters used for ML algorithms are shown in Table 12. The ML algorithms are scored on basis of performance metrics such as precision, recall, and F1 score. The ROC-AUC curve is also created for

the visualising the performance of algorithms. On the basis of performance metrics and the ROC-AUC curve, the best-fit model is selected. A total of 64 experiments were conducted to get consistent results. The final models for every algorithm provided stable results as shown in Table 12. The bold values in Table 12 indicate the best results due to these hyperparameter values.

3. Results and Discussion

Multiple machine learning classifiers are used to assess and measure the classification performance of extremism data into *propaganda*, *radicalization*, and *recruitment*. The algorithms used are MNB, SVM, random forest, and XGBoost.

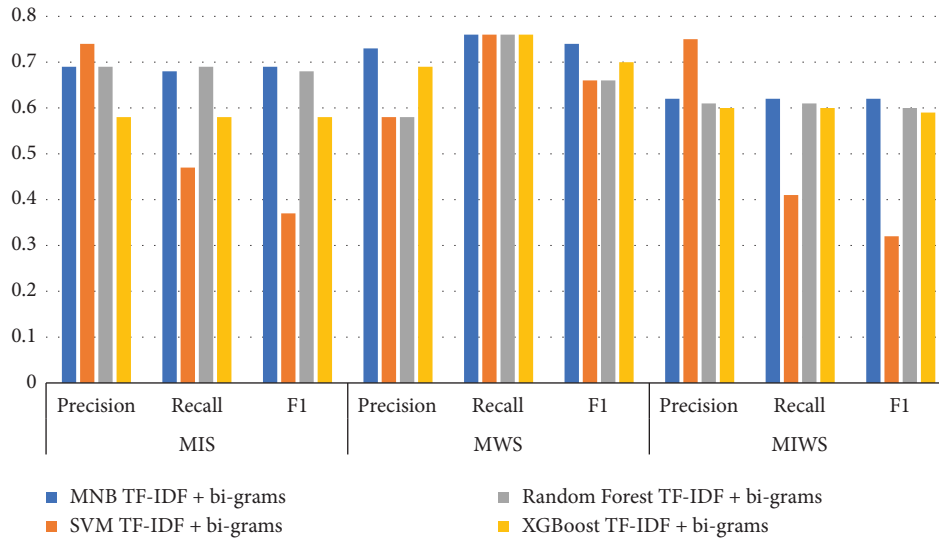


FIGURE 17: Performance metrics of algorithms on all datasets with TF-IDF bigrams.

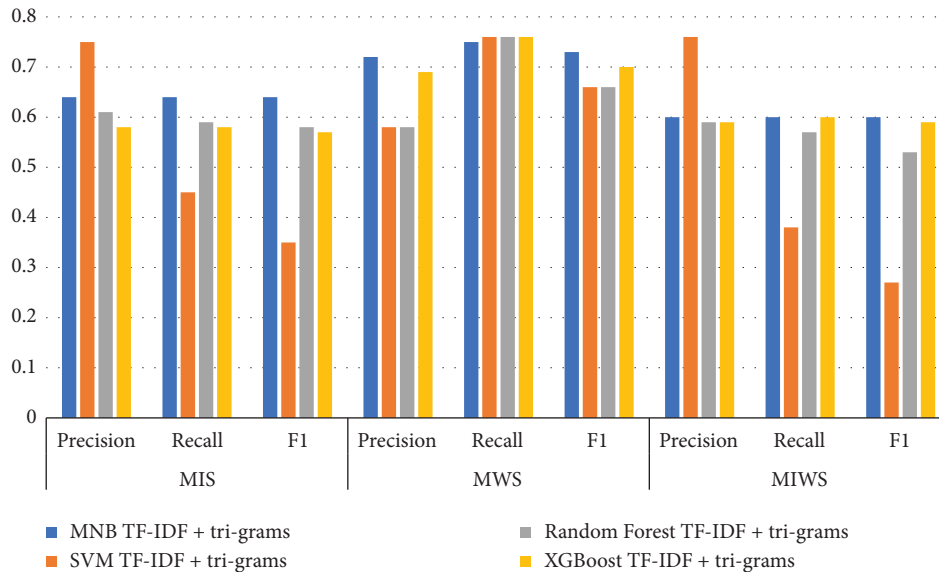


FIGURE 18: Performance metrics of algorithms on all datasets with TF-IDF trigrams.

These machine learning classifiers are chosen as they have been popularly used in online extremism detection research [36, 62].

3.1. Comparison of TF-IDF Unigram Results. Figures 16–19 shows the comparative performance of four feature extraction techniques with classifiers. It can be observed from the figures that TF-IDF unigram outperforms other feature extraction techniques, as unigram extracts the unique words that characterize the class. On the other hand, bigrams and trigrams offer comparatively low performance compared to unigrams for the frequent combinations of words in the multi-ideology MIWS dataset.

Word2vec with XGBoost offers comparable performance for the MIWS dataset, as it is pretrained on Google News data, as Google News may have accounted for extremism text. XGBoost with word2vec gives an F1 score of 0.60. It is also observed that word2vec can achieve better performance with more training epochs.

3.2. ROC-AUC (Unigram) for All Classifiers for MIWS. Receiver operating characteristics (ROC) is the graph that shows the performance of classification models at all classification thresholds [91]. Area under curve represents that the total two-dimensional are underneath ROC curve [92]. Figures 20 and 21 show the relative performance of chosen

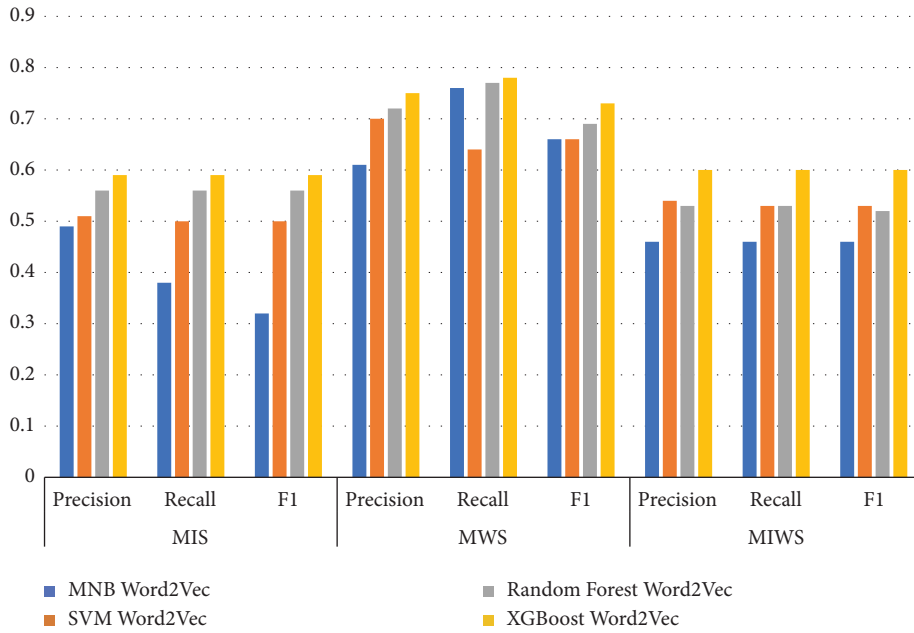


FIGURE 19: Performance metrics of algorithms on all datasets with word2vec.

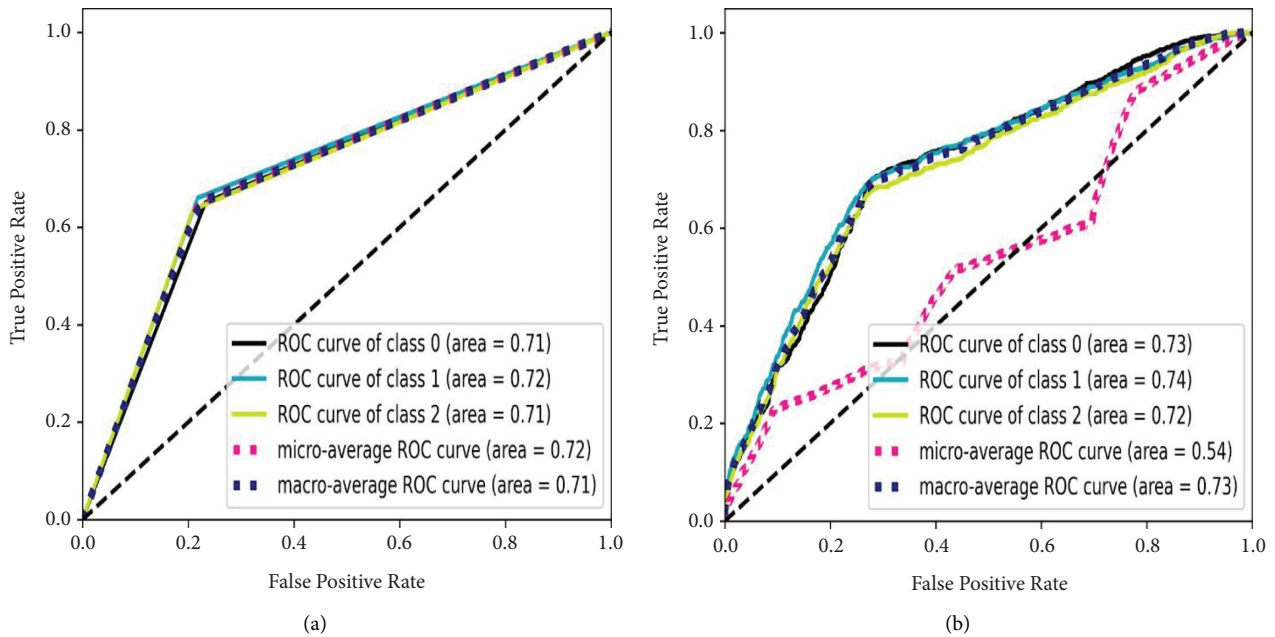


FIGURE 20: ROC curve for MIWS with TF-IDF unigram using (a) MNB and (b) SVM.

classifiers with the same feature extraction techniques, TF-IDF and unigram.

ROC-AUC are chosen for finding the relative performance of classifiers as they are focused on true positive values for multiclass propaganda (Class 0), radicalization (Class 1), and recruitment (Class 2).

It is observed that the performance of all classifiers on the MIWS dataset is satisfactory, with an AUC of around 0.70 for MNB and SVM. For random forest and XGBoost, the AUC is around 0.65. Thus, it can be said that SVM with TF-IDF unigram outperforms other classifiers.

Furthermore, SVM performs better due to marginalizing classes based on the unique words present in the MIWS dataset.

3.3. Multiclass Classification (Labelwise Precision and Recall and F1 Score with Support). A total of 64 experiments were conducted to obtain consistent results across algorithms and features combined with different random states. Four experiments for each combination of algorithm and feature were carried out. Table 13 provides the standard deviation

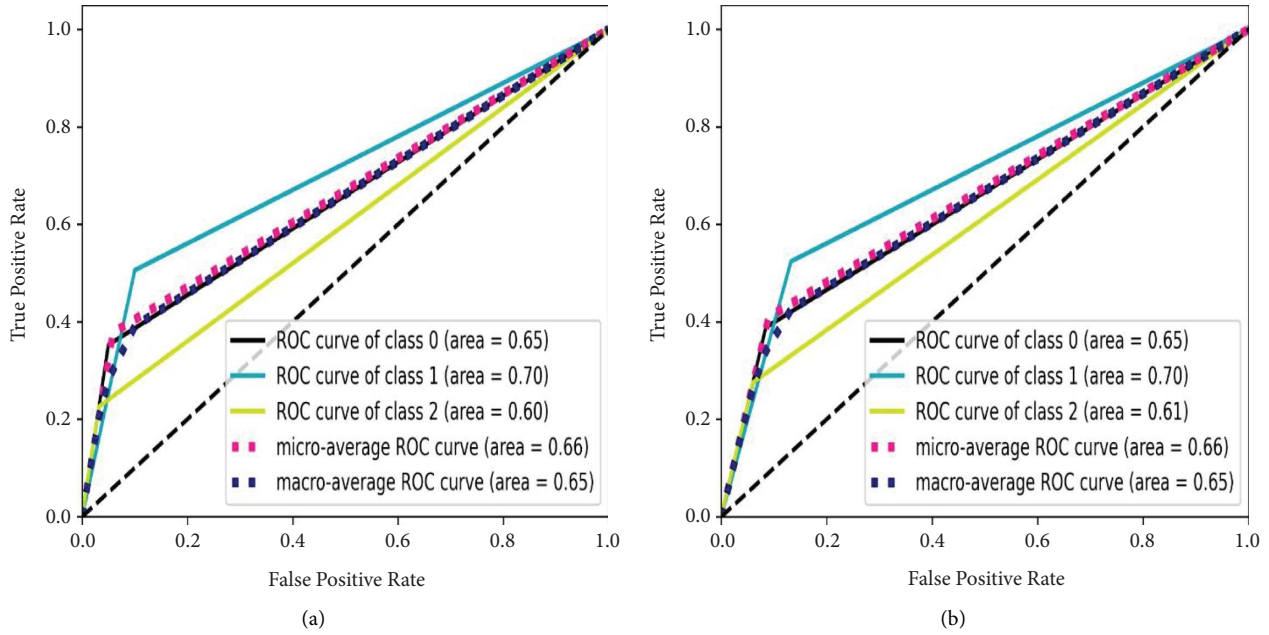


FIGURE 21: ROC curve for MIWS with TF-IDF unigram using (a) random forest and (b) XGBoost.

TABLE 13: Standard deviation of results for the MIWS dataset.

Algorithm	Features	MIWS		
		Precision	Recall	F1 score
MNB	TF-IDF	0.005774	0.011547	0.005774
	TF-IDF + bigrams	0.009574	0.005	0.008165
	TF-IDF + trigrams	0.01	0.01291	0.009574
	Word2vec	0.015	0.02	0.014142
SVM	TF-IDF	0.009574	0.01	0.005
	TF-IDF + bigrams	0.070711	0.025	0.017321
	TF-IDF + trigrams	0	0.01291	0.021602
	Word2vec	0.005	0.01	0.005
Random forest	TF-IDF	0.01	0.01	0.005
	TF-IDF + bigrams	0.005	0.005	0.005
	TF-IDF + trigrams	0.02	0.015	0.04
	Word2vec	0.005	0.005	0.005
XGBoost	TF-IDF	0	0.005	0.01
	TF-IDF + bigrams	0.009574	0.005	0.01893
	TF-IDF + trigrams	0.005	0.01	0.005
	Word2vec	0.008165	0.009574	0.008165

TABLE 14: Significant results by ranks.

Algorithm	Features	Rank
MNB	TF-IDF	4.5
MNB	TF-IDF + bigrams	2.5
SVM	TF-IDF	1
Random forest	TF-IDF	2.5
XGBoost	TF-IDF	4.5

of results on MIWS dataset. It can be observed that standard deviation is quite low. Thus, the results are stable. Table 14 provides rank for the algorithm with features based on results in Table 9. Freidman rank test

TABLE 15: Label wise performance metrics for MNB TF-IDF.

Algorithm and feature	Class	Precision	Recall	F1 score	Support
MNB and TF-IDF	Propaganda	0.59	0.63	0.61	1234
	Radicalization	0.65	0.64	0.65	831
	Recruitment	0.58	0.55	0.56	828

TABLE 16: Label wise performance metrics for SVM TF-IDF.

Algorithm and feature	Class	Precision	Recall	F1 score	Support
SVM and TF-IDF	Propaganda	0.60	0.79	0.68	1234
	Radicalization	0.75	0.69	0.72	831
	Recruitment	0.71	0.56	0.63	828

was performed to determine a rank-based significance for obtained results. The calculated p value by Freidman test was less than 0.05, that is, $1.7651e-8$. As seen in Tables 9 and 14, the ranks were calculated in descending order of results, so the lesser the rank, the more significant the results are. Therefore, SVM + TF-IDF results are significant and better than other algorithms and feature combinations.

Tables 15–18 give precision, recall, F1 score, and support for the TF-IDF unigram on the MIWS dataset for the chosen classifiers. It can be observed that SVM is the best classifier for propaganda, radicalization, and recruitment classes with an F1 score of 0.68, 0.72, and 0.63, respectively.

TABLE 17: Labelwise performance metrics for random forest TF-IDF.

Algorithm and feature	Class	Precision	Recall	F1 score	Support
Random forest and TF-IDF	Propaganda	0.60	0.66	0.63	1234
	Radicalization	0.65	0.76	0.70	831
	Recruitment	0.64	0.49	0.55	828

TABLE 18: Labelwise performance metrics for XGBoost TF-IDF.

Algorithm and feature	Class	Precision	Recall	F1 score	Support
XGBoost and TF-IDF	Propaganda	0.60	0.63	0.62	1234
	Radicalization	0.64	0.75	0.69	831
	Recruitment	0.59	0.46	0.51	828

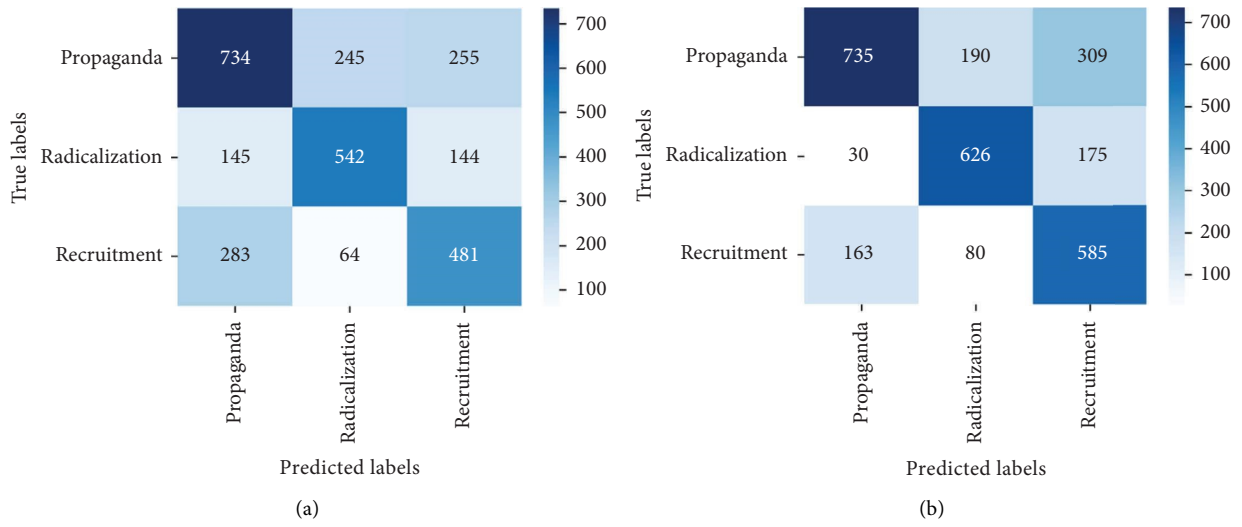


FIGURE 22: Confusion matrix of (a) MNB and (b) SVM.

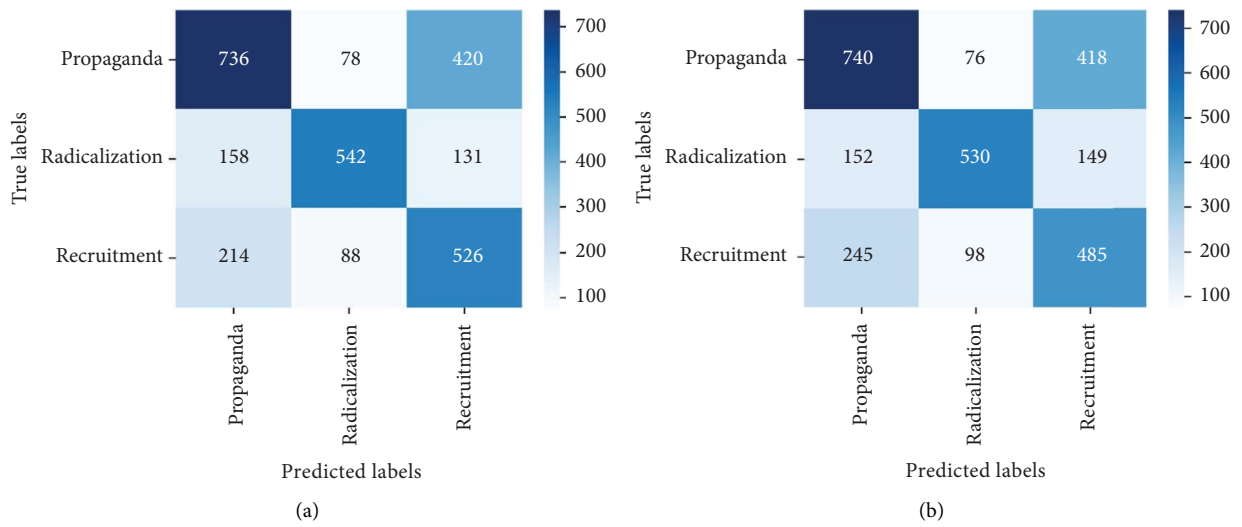


FIGURE 23: Confusion matrix of (a) random forest and (b) XGBoost.

3.4. MIS Dataset. As seen in Table 9, SVM with TF-IDF unigram provides better results than MNB, random forest, and XGBoost for MIS dataset. For the MIS dataset containing jihadist-ISIS ideology, SVM with TF-IDF provides a better F1 score of 0.70. MNB and random forest with TF-IDF bigrams show an F1score of 0.69 and 0.68, respectively. MNB gives an F1 score of 0.64 for TF-IDF trigrams, exceeding other classifiers for the same feature. Only XGBoost shows better results using word2vec for feature extraction with an F1 score of 0.59.

3.5. MWS Dataset. For the MWS dataset, XGBoost with word2vec outperforms all other features extraction and classifiers used. Table 9 shows that XGBoost with word2vec gives a precision, recall, and F1 score of 0.75, 0.78, and 0.73, respectively. This can be attributed to the unique words in the MWS dataset, which may frequently appear in Google News data. For TF-IDF unigram, bigram, and trigram, MNB outperforms other classifiers with an F1score of 0.74, 0.74, and 0.73.

3.6. MIWS Dataset. For the unigram features chosen, machine learning classifiers offer a better performance. MNB, SVM, random forest, and XGBoost give an F1 score of 0.61, 0.68, 0.63, and 0.62, respectively, for unigram features. SVM provides maximum performance if F1 scores are compared. This can be attributed to common unique words for MIS and MWS.

For bigram and trigram features, the performance of algorithms reduces drastically. This can be attributed to different words based on the ideologies that are merged in a single dataset. Thus, bigram and trigram may not be effective in identifying and analyzing multiple ideologies together. Word2vec gives better performance for XGBoost. The F1score obtained from XGBoost with word2vec is 0.60. Figures 22 and 23 show the confusion matrix obtained by applying MNB, SVM, RF, and XGBoost on the MIWS dataset.

3.7. Inferences and Discussion. As seen in Tables 9–18, the results are a bit low. This is due to the merging of two different ideologies as the aim is to develop a generalized and ideology-independent extremism detection model. Methods and techniques to improve the results are discussed in the Section, Future Work.

Table 9 shows the comparative performance of the classifiers on the different feature extraction methods. The MIWS dataset with ~17,000 ISIS and ~11,000 WS examples is a multi-ideology dataset. The extremist dataset was developed and validated with three statistical methods that proved that the dataset is robust with the unique features in the three classes. The performance of ML algorithms on these extracted features in the dataset also shows potential for applying DL classifiers.

3.8. Limitations. The size of the dataset is an important aspect of machine learning. However, the size of the SEED

dataset used in this work is limited, with fewer research articles. This is due to the lower availability of extremist text examples classified as propaganda, radicalization, and recruitment in the existing literature. Even with data imbalance, current data provides acceptable results, but balanced data is required to predict extremist text with precision.

The extremist text in the existing literature was manually labeled as propaganda, radicalization, and recruitment by experts. However, this labeling is limited by interrater agreement or expert opinion in the existing literature. Thus, the SEED dataset that is employed for topic modeling has the threat of expert bias. Hence the work relies on statistical validation techniques to verify the strength of the dataset. Furthermore, it is challenging both experimentally and ethically to quantify the bias of experts. Hence, at current stage of research it is not possible to compare the bias of both experts and the ML algorithm.

In this work, only three different topics or classes are considered for extremism classification text. Therefore, these topics were identified using simple LDA. The context-aware LDA [93] or context-aware topic modeling could be used to extract multiple different topics within extremism text.

Rigorous statistical tests were essential for estimating the strengths of the topic clusters. This work employed cosine similarity, Wilcoxon signed-rank, and chi-square tests for data validation as they were popularly employed in the literature. However, more statistical tests can be additionally employed to ensure the quality of data.

In this work, only four feature extraction techniques and four machine learning classifiers are employed on the developed MIWS dataset. Therefore, the results are limited by the choice of these representative classifiers and feature extractors. The classification and feature extraction purpose was to realize the model that would accurately classify the dataset.

A variety of advanced feature extraction techniques such as pretrained vectors can be further evaluated for a better accuracy. Advanced classifiers and transformers can also be employed for achieving better accuracy.

4. Conclusion

This work focuses on constructing a multi-ideology and multiclass extremism text dataset with a comparative analysis of the performance of features extraction techniques and machine learning classifiers. Most extremism research studies focuses on a single ideology, with binary or tertiary classification such as extremist, nonextremist, and irrelevant classes. Consequently, there are limited insights from such works [19].

In this work, we develop a multi-ideology dataset with the most popular jihadist-ISIS and White supremacist ideologies. This dataset provides a broader view of extremism text with popular extremist ideologies brought together for better insights into data. The dataset also builds a multilabel extremist text dataset by classifying data as propaganda, radicalization, and recruitment.

The extremist text dataset was made contemporary by collecting extremist texts from different data sources

(Twitter, ISIS Kaggle, StormFront dataset, and Gab dataset). In addition, we created ideology-specific datasets, which are called MIS (jihadist-ISIS), MWS (White supremacist), and proposed MIWS (multi-ideology) datasets with data pre-processing techniques applied.

A SEED dataset was created using existing literature that provided us with labeled examples of propaganda, radicalization, and recruitment. Then, the labeled SEED dataset was used to group/cluster the MIS, MWS, and MIWS datasets into propaganda, radicalization, and recruitment by using the LDA technique and cosine similarity. The grouping/clustering was further validated using statistical techniques. In this work, three different statistical tests, such as cosine similarity, Wilcoxon signed-rank test, and chi-square test, validated data labeling. Thus, our work is free from expert bias resulting due to manual validation such as previous literature. The visualization of word vectors with t-sne is also performed to highlight the unique words in propaganda, radicalization, and recruitment classes from the MIWS dataset.

To assess the performance of datasets, multiple features such as TF-IDF (unigram, bigram, and trigram) and pretrained word2vec (Google News) are used. These features were provided as input to classifiers such as MNB, SVM, RF, and XGBoost. For the proposed MIWS dataset, TF-IDF unigram with SVM provides the highest precision of 0.69, recall of 0.68, and F1score of 0.68. Thus, the results obtained using ML algorithms can be considered as a baseline for future work consisting of deep learning techniques.

This work, pioneers in developing the multi-ideology extremism text, MIWS dataset can classify extremism data into multiclass such as propaganda, radicalization, and recruitment with robust statistical data validation techniques employed. Furthermore, this work investigates the best feature extraction technique and classifier for the proposed MIWS dataset, which guarantees better classification performance.

4.1. Future Work. The presented work is an important milestone in online extremism text detection research. This will open multiple avenues in the following research areas:

4.1.1. Versatility of Extremism Text Dataset. Our work proves that multi-ideology datasets create a broader view of extremism text with comparable classification performance over single-ideology datasets. In the future, the presented dataset can be made more versatile with other popular extremist ideologies and sources. Increasing the SEED dataset also may produce more significant results. Different techniques such as word mover's distance [94] can also be used to calculate and improve the similarity between labels and topics.

4.1.2. Feature Extraction Techniques. Context-aware topic modeling can be used to extract multiple different topics such as promoting violent acts and antisemitism. Popular feature extraction techniques such as pretrained vectors,

GLoVe [95], and FastText [96], can be employed to extract complex relationships among extremism data. These can further enhance the accuracy of extremism detection models.

4.1.3. Transfer Learning and Deep Learning Approaches.

This research work uses machine learning classifiers for evaluating the proposed dataset. Future works can use deep learning models such as LSTM and CNN, and pretrained networks such as FastText, BERT, or RoBERTa for a better semantic analysis of extremism data. This can help achieve a higher performance for the classification of extremism text into propaganda, radicalization, and recruitment.

4.1.4. Detection of Extremism Based on Geographical Context.

The geographical location of extremists and extremist organizations plays an important role in analyzing propaganda, radicalization, and recruitment on social media platforms. The researchers have used the tweet location to identify extremist affiliations. It is necessary to identify the targeted nations through the extremist text which will speculate the activities of extremists. So, the extraction of geographical locations can play a major role in providing insights into extremist propaganda, radicalization, and recruitment tactics.

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

The author would like to thank Symbiosis International University and Symbiosis Institute of Technology for providing the opportunity and resources to make this research study possible. This project was partially funded by the Research Support Grant of Symbiosis International University.

References

- [1] J. Schultz, "How much data is created on internet each day?," 2019, <https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/>.
- [2] S. Baele, K. Boyd, and T. Coan, *ISIS Propaganda*, Oxford University Press, Oxford, USA, 2020.
- [3] A. Dornbierer, "How al-qaeda recruits online," 2011, <https://thediplomat.com/2011/09/how-al-qaeda-recruits-online/>.
- [4] J. Daniels, "The algorithmic rise of the 'alt-right,'" *Contexts*, vol. 17, no. 1, pp. 60–65, 2018.
- [5] N. MacFarquhar, A. Feuer, M. Baker, and S. Frenkel, "The proud boys, who trade in political violence, get a boost from trump - the New York times," 2020, <https://www.nytimes.com/2020/09/30/us/proud-boys-trump.html>.

- [6] Congress Gov, "Domestic terrorism prevention act of 2019," 2019, <https://www.congress.gov/bill/116th-congress/senate-bill/894>.
- [7] J. Martinson, "Christchurch attacks: the media's rush to be first causes its own kind of harm," 2019, <https://www.theguardian.com/media/commentisfree/2019/mar/24/christchurch-attacks-the-medias-rush-to-be-first-causes-its-own-kind-of-harm>.
- [8] J. Coaston, "The new zealand shooter's manifesto shows how white nationalist rhetoric spreads," 2019, <https://www.vox.com/identities/2019/3/15/18267163/new-zealand-shooting-christchurch-white-nationalism-racism-language>.
- [9] C. Timberg, D. Harwell, E. Dvoskin, and T. Romm, "How social media's business model helped the New Zealand massacre go viral," *Washington Post*, vol. 19, 2019.
- [10] B. L. Smith, "Propoganda encyclopedia," 1999, <https://www.britannica.com/topic/propaganda>.
- [11] B. Johnson, "Shared themes, tactics in white supremacist and islamist extremist propoganda – homeland security today," 2020, <https://www.hstoday.us/subject-matter-areas/counterterrorism/shared-themes-recruitment-tactics-in-white-supremacist-and-islamist-extremist-propaganda/>.
- [12] P. Wignell, K. Chai, S. Tan, K. O'Halloran, and R. Lange, "Natural language understanding and multimodal discourse analysis for interpreting extremist communications and the Re-use of these materials online," *Terrorism and Political Violence*, vol. 33, no. 1, pp. 71–95, 2018.
- [13] T. Welch, "Theology, heroism, justice, and fear: an analysis of ISIS propaganda magazines Dabiq and Rumiyah," *Dynamics of Asymmetric Conflict*, vol. 11, no. 3, pp. 186–198, 2018.
- [14] C. McCauley and S. Moskalenko, "Mechanisms of political radicalization: pathways toward terrorism," *Terrorism and Political Violence*, vol. 20, no. 3, pp. 415–433, 2008.
- [15] US Dept of Justice, "Awareness brief: online radicalization to violent extremism," 2018, <https://www.theiacp.org/sites/default/files/2018-07/RadicalizationtoViolentExtremismAwarenessBrief.pdf>.
- [16] A. T. Chatfield, C. G. Reddick, and U. Brajawidagda, "Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks," in *Proceedings of the 16th Annual International Conference on Digital Government Research*, pp. 239–249, Boston, MA, USA, June 2015.
- [17] M. S. Kimmel, "Globalization and its Mal(e)Contents," *International Sociology*, vol. 18, no. 3, pp. 603–620, 2003.
- [18] R. Torok, "ISIS and the institution of online terrorist recruitment," 2015, <https://www.mei.edu/publications/isis-and-institution-online-terrorist-recruitment>.
- [19] M. Gaikwad, S. Ahirrao, S. Phansalkar, and K. Kotecha, "Online extremism detection: a systematic literature review with emphasis on datasets, classification techniques, validation methods, and tools," *IEEE Access*, vol. 9, pp. 48364–48404, 2021.
- [20] O. De Gibert, N. Perez, A. García-Pablos, and M. Cuadros, "Hate speech dataset from a white supremacy forum," 2018, <https://aclanthology.org/W18-5102/>.
- [21] B. Kennedy, "The gab hate corpus: a collection of 27k posts annotated for hate speech," 2020, <https://psyarxiv.com/hqjxn/>.
- [22] Kaggle, "Fifth tribe, 'how ISIS uses twitter'," 2015, <https://www.kaggle.com/fifthtribe/how-isis-uses-twitter>.
- [23] J. M. Berger, "Nazis vs. ISIS on twitter: A comparative study of white nationalist and ISIS online social media networks," 2016, <https://extremism.gwu.edu/sites/g/files/zaxdzs2191/f/downloads/Nazis%20v.%20ISIS.pdf>.
- [24] J. M. Berger, "The dangerous spread of extremist manifestos," 2019, <https://www.theatlantic.com/ideas/archive/2019/02/christopher-hasson-was-inspired-breivik-manifesto/583567/>.
- [25] J. M. Berger and J. Morgan, "The ISIS twitter census," 2015, https://www.brookings.edu/wp-content/uploads/2016/06/isis_twitter_census_berger_morgan.pdf.
- [26] J. M. Berger, "The alt-right twitter census: defining and describing the audience for alt-right content on twitter," 2018, <https://www.voxpol.eu/new-research-report-the-alt-right-twitter-census-by-j-m-berger/>.
- [27] United Nations, "United nations strategy and plan of action on hate speech," 2020, <https://www.un.org/en/genocideprevention/hate-speech-strategy.shtml>.
- [28] P. R. Neumann, "The trouble with radicalization," *International Affairs*, vol. 89, no. 4, pp. 873–893, 2013.
- [29] H. Speech, "Facebook," 2020, https://www.facebook.com/communitystandards/recentupdates/hate_speech/.
- [30] W. M. Curtis, "Hate speech," 2015, <https://www.britannica.com/topic/hate-speech>.
- [31] S. Wibisono, W. R. Louis, and J. Jetten, "A multidimensional analysis of religious extremism," *Frontiers in Psychology*, vol. 10, p. 2560, 2019.
- [32] C. Winter, P. Neumann, A. Meleagrou-Hitchens, M. Ranstorp, L. Vidino, and J. Fürst, "Online extremism: research trends in internet activism, radicalization, and counter-strategies," *International Journal of Computer Vision*, vol. 14, no. 2, pp. 1–20, 2020.
- [33] Eu, "Study on right-wing violent extremism and hate speech online," 2020, <https://ec.europa.eu/newsroom/just/items/671669>.
- [34] M. Ray, "Paris attacks of 2015: additional information | britannica," 2015, <https://www.britannica.com/event/Paris-attacks-of-2015/additional-info#history>.
- [35] M. Fernandez, M. Asif, and H. Alani, "Understanding the roots of radicalisation on twitter," in *Proceedings of the 10th ACM Conference on Web Science*, pp. 1–10, Boston, MA, USA, May 2018.
- [36] O. Araque and C. A. Iglesias, "An approach for radicalization detection based on emotion signals and semantic similarity," *IEEE Access*, vol. 8, pp. 17877–17891, 2020.
- [37] Kaggle, "ActiveGalaxy, 'ISIS related dataset'," 2016, <https://www.kaggle.com/activegalaxy/isis-related-tweets>.
- [38] Kaggle, "FifthTribe, 'ISIS religious text'," 2017, <https://www.kaggle.com/fifthtribe/isis-religious-texts>.
- [39] M. Rowe and H. Saif, "Mining pro-ISIS radicalisation signals from social media users," in *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*, Cologne, Germany, May 2016.
- [40] D. O'Callaghan, N. Prucha, D. Greene, M. Conway, J. Carthy, and P. Cunningham, "Online social media in the Syria conflict: encompassing the extremes and the in-betweens," in *Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, pp. 409–416, Beijing, China, August 2014.
- [41] L. Kaati, E. Omer, N. Prucha, and A. Shrestha, "Detecting multipliers of jihadism on twitter," in *Proceedings of the 2015 IEEE International Conference On Data Mining Workshop (ICDMW)*, pp. 954–960, Atlantic City, NJ, USA, November 2015.
- [42] M. Ashcroft, A. Fisher, L. Kaati, E. Omer, and N. Prucha, "Detecting jihadist messages on twitter," in *Proceedings of the 2015 European Intelligence And Security Informatics Conference*, pp. 161–164, Washington, DC, USA, September 2015.

- [43] M. C. Benigni, K. Joseph, and K. M. Carley, "Online extremism and the communities that sustain it: detecting the ISIS supporting community on Twitter," *PLoS One*, vol. 12, no. 12, pp. e0181405–e0181423, 2017.
- [44] M. F. Abrar, M. S. Arefin, and M. S. Hossain, "A framework for analyzing real-time tweets to detect terrorist activities," in *Proceedings of the 2nd International Conference on Electrical, Computer and Communication Engineering, ECCE 2019*, pp. 1–6, March 2019, Haldia, India.
- [45] S. Ahmad, M. Z. Asghar, F. M. Alotaibi, and I. Awan, "Detection and classification of social media-based extremist affiliations using sentiment analysis techniques," *Human-centric Computing and Information Sciences*, vol. 9, no. 1, 2019.
- [46] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid, and J. Shah, "Sentiment analysis of extremism in social media from textual information," *Telematics and Informatics*, vol. 48, Article ID 101345, 2020.
- [47] I. Gialampoukidis, G. Kalpakis, T. Tsirikika, S. Papadopoulos, S. Vrochidis, and I. Kompatsiaris, "Detection of Terrorism-Related Twitter Communities Using Centrality Scores," in *Proceedings of the 2nd International Workshop On Multimedia Forensics And Security, co-located with ICMR 2017*, pp. 21–25, Bucharest, Romania, June 2017.
- [48] S. Jaki and T. De Smedt, "Right-wing German hate speech on twitter: analysis and automatic detection," 2019.
- [49] T. De Smedt, "Multilingual cross-domain perspectives on online hate speech," *CLiPS Technical Report Series*, vol. 8, 2018.
- [50] M. Heidarysafa, K. Kowsari, T. Odukoya, P. Potter, L. E. Barnes, and D. E. Brown, "Women in ISIS propaganda: a natural language processing analysis of topics and emotions in a comparison with a mainstream religious group," *Advances in Intelligent Systems and Computing*, vol. 3, pp. 610–624, 2020.
- [51] O. Araque and C. A. Iglesias, "An ensemble method for radicalization and hate speech detection online empowered by sentic computing," *Cognitive Computation*, vol. 14, no. 1, pp. 48–61, 2021.
- [52] V. Basile, "SemEval-2019 task 5: multilingual detection of hate speech against immigrants and women in twitter," in *Proceedings of the 13th International Workshop on Semantic Evaluation, 2019*, pp. 54–63, Minneapolis, MN, USA, June 2019.
- [53] T. Davidson, D. Warmesley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," in *Proceedings of the 11th International Conference on Web and Social Media, ICWSM 2017*, pp. 512–515, Montreal, Canada, May 2017.
- [54] S. Mussiraliyeva, B. Omarov, P. Yoo, and M. Bolatbek, "Applying machine learning techniques for religious extremism detection on online user contents," *Computers, Materials & Continua*, vol. 70, no. 1, pp. 915–934, 2022.
- [55] Vkontakte, "Vkontakte social network," 2014, https://vk.com/topic-78863260_30603285.
- [56] Twitter, "Updating our rules against hateful conduct," 2020, https://blog.twitter.com/en_us/topics/company/2019/hatefulconductupdate.html.
- [57] S. Agarwal and A. Sureka, *Topic-Specific Youtube Crawling To Detect Online Radicalization*, Springer, Berlin, Germany, 2015.
- [58] H. Saif, T. Dickinson, L. Kastler, M. Fernandez, and H. Alani, "A semantic graph-based approach for radicalisation detection on social media," in *Proceedings of the European Semantic Web Conference, 2017*, pp. 571–587, Portorož, Slovenia, June 2017.
- [59] M. Petrovskiy and M. Chikunov, "Online extremism discovering through social network structure analysis," in *Proceedings of the 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT)*, pp. 243–249, New York, NY, USA, March 2019.
- [60] M. Moussaoui, M. Zaghdoud, and J. Akaichi, "A possibilistic framework for the detection of terrorism-related Twitter communities in social media," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 13, Article ID 5077, 2019.
- [61] W. Sharif, "An empirical approach for extreme behavior identification through tweets using machine learning," *Applied Sciences*, vol. 9, no. 18, Article ID 5723, 2019.
- [62] S. Mussiraliyeva, M. Bolatbek, B. Omarov, and K. Bagitova, "Detection of extremist ideation on social media using machine learning techniques," in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer, Cham, Champa, 2020.
- [63] Z. Ul Rehman, S. Abbas, M. Adnan Khan et al., "Understanding the language of ISIS: an empirical approach to detect radical content on twitter using machine learning," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 1075–1090, 2021.
- [64] S. Agarwal and A. Sureka, "Using KNN and SVM based one-class classifier for detecting online radicalization on twitter," in *Proceedings of the International Conference on Distributed Computing and Internet Technology*, pp. 431–442, New York, NY, USA, June 2015.
- [65] H. S. Alatawi, A. M. Alhothali, and K. M. Moria, "Detecting white supremacist hate speech using domain specific word embedding with deep learning and bert," 2020, <http://arxiv.org/abs/2010.00357>.
- [66] M. Nouh, R. C. Jason Nurse, and M. Goldsmith, "Understanding the radical mind: identifying signals to detect extremist content on Twitter," in *Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics, ISI 2019*, pp. 98–103, Shenzhen, China, July 2019.
- [67] T. De Smedt, G. De Pauw, and P. Van Ostaeyen, "Automatic detection of online jihadist hate speech," 2018, <https://arxiv.org/abs/1803.04596>.
- [68] E. Ferrara, W.-Q. Wang, O. Varol, A. Flammini, and A. Galstyan, "Predicting online extremism, content adopters, and interaction reciprocity," 2016, <https://arxiv.org/abs/1605.00659>.
- [69] U. Kursuncu, "Modeling islamist extremist communications on social media using contextual dimensions," in *Proceedings of the ACM Hum Comput Interact*, vol. 3, pp 1–22, CSCW, New York, NY, USA, November 2019.
- [70] D. Xie, J. Xu, and T.-C. Lu, "Automated classification of extremist Twitter accounts using content-based and network-based features," in *Proceedings of the 2016 IEEE International Conference on Big Data (Big Data)*, pp. 2545–2549, Washington, DC, USA, December 2016.
- [71] I. V. Mashechkin, M. I. Petrovskiy, D. V. Tsarev, and M. N. Chikunov, "Machine learning methods for detecting and monitoring extremist information on the internet," *Programming and Computer Software*, vol. 45, no. 3, pp. 99–115, 2019.
- [72] A. Kaur, J. K. Saini, and D. Bansal, "Detecting radical text over online media using deep learning," 2019, <https://arxiv.org/abs/1907.12368>.

- [73] S. O. Mussiraliyeva, B. Bolatbek, M. Ospanov, R. Baispay, G. Medetbek, and Z. Yeltay, "Applying deep learning for extremism detection," *Communications in Computer and Information Science*, vol. 21, pp. 597–605, 2021.
- [74] P. Simi, K. Blee, M. DeMichele, and S. Windisch, "Addicted to hate: identity residual among former white supremacists," *American Sociological Review*, vol. 82, no. 6, pp. 1167–1187, 2017.
- [75] J. M. Berger, K. Aryaeinejad, and S. Looney, "There and back again: how white nationalist ephemera travels between online and offline spaces," *The RUSI Journal*, vol. 165, no. 1, pp. 114–129, 2020.
- [76] C. Charles, *(Main)streaming Hate: Analyzing White Supremacist Content and Framing Devices on YouTube*, University of Central Florida, Orlando, FL, USA, 2020.
- [77] B. Ray and G. E. Marsh, "Recruitment by extremist groups on the Internet," *First Monday*, vol. 6, no. 2, 2001.
- [78] S. T. Dumais, "Latent semantic analysis," *Annual Review of Information Science & Technology*, vol. 38, no. 1, pp. 188–230, 2005.
- [79] J. C. Chappelier and E. Eckard, "PLSI: the true fisher kernel and beyond: iid processes, information matrix and model identification in PLSI," *Machine Learning and Knowledge Discovery in Databases*, vol. 5781, no. 1, pp. 195–210, 2009.
- [80] D. Angelov, "Top2vec: distributed representations of topics," 2020, <https://arxiv.org/abs/2008.09470>.
- [81] M. Grootendorst, "BERTopic: Leveraging BERT and C-TF-IDF to Create Easily Interpretable Topics," *Zenodo*, vol. 2020, Article ID 4381785, 81 pages, 2020.
- [82] S. Gupta, "Overview of text similarity metrics in python," 2018, <https://towardsdatascience.com/overview-of-text-similarity-metrics-3397c4601f50>.
- [83] D. Kochedykov, M. Apishev, L. Golitsyn, and K. Vorontsov, "Fast and modular regularized topic modelling," in *Proceedings of the 21st Conference of Open Innovations Association FRUCT*, pp. 182–193, Helsinki, Finland, November 2017.
- [84] Scikit-Learn, "GridSearchCV," 2011, https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html.
- [85] J. Han, M. Kamber, and J. Pei, "Getting to know your data," in *Data Mining*, pp. 39–82, Elsevier, Americas Brazil, 2012.
- [86] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometric Bulletin*, vol. 1, no. 6, p. 80, 1945.
- [87] Scikit-Learn, "CountVectorizer," 2011, https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html.
- [88] Scikit-Learn, "TFIDFVectorizer," 2011, https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html.
- [89] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *Springer Series in Statistics*, vol. 50, pp. 11–28, 1992.
- [90] A. Aizawa, "An information-theoretic perspective of tf-idf measures," *Information Processing & Management*, vol. 39, no. 1, pp. 45–65, 2003.
- [91] F. Melo, "Receiver operating characteristic (ROC) curve," in *Encyclopedia of Systems Biology*, pp. 1818–1823, Springer, New York, NY, USA, 2013.
- [92] F. Melo, "Area under the ROC curve," in *Encyclopedia of Systems Biology*, pp. 38–39, Springer, New York, NY, USA, 2013.
- [93] W. Li, T. Matsukawa, H. Saigo, and E. Suzuki, "Context-aware latent dirichlet allocation for topic segmentation," in *Advances in Knowledge Discovery and Data Mining*, pp. 475–486, Springer, Cham, Champa, 2020.
- [94] L. Wu, "Word Mover's Embedding: From Word2Vec to Document Embedding," 2018, <https://arxiv.org/abs/1811.01713>.
- [95] J. Pennington, R. Socher, and C. D. Manning, "GloVe: global vectors for word representation," in *Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, Stanford University, Stanford, CA, USA, 2014.
- [96] A. Joulin, E. Grave, P. Bojanowski, M. Douze, H. Jégou, and T. Mikolov, "Fasttext.zip: compressing text classification models," 2016, <https://arxiv.org/abs/1612.03651>.

Retraction

Retracted: An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding

Computational Intelligence and Neuroscience

Received 9 December 2022; Accepted 9 December 2022; Published 9 February 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.

Computational Intelligence and Neuroscience has retracted the article titled “An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Chief Editor.

The authors do not agree to the retraction.

References

- [1] A. Alsharif, K. Aggarwal, Sonia, D. Koundal, H. Alyami, and D. Ameyed, “An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8467349, 8 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Retraction

Retracted: Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty

Computational Intelligence and Neuroscience

Received 23 November 2022; Accepted 23 November 2022; Published 28 December 2022

Copyright © 2022 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.

Computational Intelligence and Neuroscience has retracted the article titled “Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Chief Editor.

References

- [1] X. Zhang, Y. Shen, S. Yang, B. Wang, and J. Chen, “Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1054299, 6 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Retraction

Retracted: Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] T. Saba, A. R. Khan, I. Abunadi, S. A. Bahaj, H. Ali, and M. Alruwaythi, "Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8622022, 9 pages, 2022.

Research Article

Arabic Speech Analysis for Classification and Prediction of Mental Illness due to Depression Using Deep Learning

Tanzila Saba,¹ Amjad Rehman Khan,¹ Ibrahim Abunadi ,¹ Saeed Ali Bahaj,² Haider Ali ,³ and Maryam Alruwaythi¹

¹Artificial Intelligence and Data Analytics Lab CCIS, Prince Sultan University, Riyadh, Saudi Arabia

²MIS Department College of Business Administration, Prince Sattam Bin Abdulaziz University, Alkharj 11942, Saudi Arabia

³Department of Statistics, University of Gujrat, Gujrat, Pakistan

Correspondence should be addressed to Haider Ali; 18101713-006@uog.edu.pk

Received 3 February 2022; Revised 30 March 2022; Accepted 18 April 2022; Published 27 May 2022

Academic Editor: Syed Ahmad Chan Bukhari

Copyright © 2022 Tanzila Saba 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.

Depression is a global prevalent ailment for possible mental illness or mental disorder globally. Recognizing depressed early signs is critical for evaluating and preventing mental illness. With the progress of machine learning, it is possible to make intelligent systems capable of detecting depressive symptoms using speech analysis. This study presents a hybrid model to identify and predict mental illness from Arabic speech analysis due to depression. The proposed hybrid model comprises convolutional neural network (CNN) and a support vector machine (SVM) to identify and predict mental disorders. Experiments are performed on the Arabic speech benchmark data set of 200 speeches. A total of 70% of data were reserved for training, while 30% of data were to test the proposed model. The hybrid model (CNN + SVM) attained a 90.0% and 91.60% accuracy rate to predict the depression from Arabic speech analysis for training and testing stages. To authenticate the results of a proposed hybrid model, recurrent neural network (RNN) and CNN are also applied to the same data set individually, and the results are compared with each other. The RNN achieved an 80.70% and 81.60% accuracy rate to predict depression while speaking in the training and testing stages. The CNN predicted the depression in the training and testing stages with 88.50% and 86.60% accuracy rates. Based on the analysis, the proposed hybrid model secured better prediction results than individual RNN and CNN models on the same data set. Furthermore, the suggested model had a lower FPR, FNR, and higher accuracy, AUC, sensitivity, and specificity rate than individual RNN, CNN model performance in predicting depression. Finally, the achieved findings will be helpful to classify depression while speaking Arabic/speech and will be beneficial for physicians, psychiatrists, and psychologists in the detection of depression.

1. Introduction

Depression is known as a mental disorder or mental illness, and according to WHO, currently more than 300 million (4.4%) people are affected by depression [1], and its rate is continually increasing [2]. From 2005 to 2015, almost 18% of the occurrence of depression has increased worldwide. Depression leads to somatic problems, mental disorders, sleep disorders, and gastrointestinal problems. The self-confidence and rumination symptoms show in depression-related patients [3, 4]. It affects the functioning or performance of patients at school, family, and work. It may also severely impact people causing self-harm and sometimes

suicide. Mood disorder and mental illness in adult life are also associated with depressive disorder [5, 6]. From depression, people may also experience a bad mood, low self-esteem, loss of interest, low energy, and body pain without a clear cause [7]. Automatic speech recognition (ASR) is well known as speech recognition. It provides the facility of understanding the users' speech by converting the word speech into series using the computer [8]. A speech emotion recognition system is helpful in medical practice for detecting changes in mental state and emotions. For example, when a patient has mood swings, the system will react rapidly and examine their current psychological state [9]. As a result, the depression prediction methods might help

design better mental health care software and technologies such as intelligent robots.

1.1. Background. Depression rates are continually increasing in people where many issues occur from this mental disorder in daily life. Unfortunately, it is difficult to predict depression from people while neutral speaking. Machine learning can be considered one of the most common ways to look at data from different sources and figure out how people feel and speak under depression.

Early recognition of depressed symptoms, followed by evaluation and therapy, may greatly enhance the odds of controlling symptoms and the underlying illness and attenuate harmful consequences for health and social life. However, detecting depression disorder is difficult and time-consuming. Current methods primarily rely on clinical discussion and surveys conducted by a psychologist for mental disorder predictions. This method is largely based on one-on-one surveys and may generally identify depression as a mental disorder condition. Since machine learning models are increasingly being used to make essential predictions in critical situations daily, the demand for transparency from all the people in the AI industry grows in these situations. Many research projects attempt to develop an automated depression detection system [10]. The GMMs (Gaussian mixtures models) [11, 12], HMMs (hidden Markov models) [13], and SVM (support vector machine) [14, 15] were used to recognize the depressed emotions using the speech data.

Deep neural networks have lately made significant contributions to a wide range of disciplines of study, including pattern recognition, and proved a better option than traditional machine learning techniques such as SVM, ANN, HMM, and so on. Han et al. [16] proposed a DNN-ELM (extreme learning machine) based voice emotion classification system. Bertero and Fung [17] used the convolutional neural network (CNN), which has a lot of applications in this field to recognize voice-related emotions, and reported good results. In the subsequent research, RNN and LSTM (long short-term memory) were also enhanced, and GRU [18], QRNN [19], and other models were also proposed for speech data. Simultaneously, different work attempted to integrate the CNN and RNN into a CRNN model for speech emotion recognition [20]. The 1D-CNN architecture improves the individual systems' performance since it was recently developed to deal with text or one-dimension data such as human speech. However, ensemble CNN models exhibited better performance for emotions classification using speech analysis [7].

To help address these issues, we built an automated method for identifying depressive symptoms from Arabic speech analysis. The proposed automated mental illness identification technique, which describes users' concerns in Arabic, might significantly contribute to this research area. This study proposed a hybrid model (CNN + SVM) to classify depression from Arabic speech analysis and predict mental disorders. Additionally, results are compared with RNN and 1-D CNN for the same problem on the same data set.

1.2. Main Contributions. This research has the following main contributions:

- (i) The first time, CNN + SVM-based hybrid model is proposed for Arabic speech analysis to predict mental illness due to depression and attained approximately 92% accuracy
- (ii) A large Arabic speech benchmark data set is employed for experiments
- (iii) Experts from both the medical and psychology fields are consulted to derive possible symptoms of depression for best features identification
- (iv) RNN and CNN are individually applied to the same data set for analysis and comparisons of the results of the proposed hybrid model
- (v) Using our model researcher will detect depression while speaking the Arabic language with an approximately 92% accuracy rate

Furthermore, this research is divided into four main sections. Section 2 presents the proposed methodology. Section 3 details experimental results with analysis. Section 4 compares the results of the proposed hybrid model with individual RNN and CNN on the same benchmark data set. Finally, Section 5 summarizes the research.

2. Proposed Methodology

This study is designed to predict depression using recorded Arabic speech analysis or while speaking in the Arabic language with the proposed hybrid approach exhibited in Figure 1 and compare with deep learning (DL) models such as RNN and CNN.

First, we extracted the features from the speeches of both depression and nondepression groups. The Mcc, chroma_stft, chroma_cqt, tonnetz, melspectrogram, spectral_centroid, and spectral_contrast features were extracted for speeches using the Python coding.

CNN is a deep learning model used for pattern classification and is composed of an input layer, hidden layers, and output layer $F = (Y, W) = X$, where Y is the input, W is the weight vector, F is any function, and X is the output. The hidden layer contains four components: the convolution layer, pooling layer, fully connected layer, and activation function [21].

- (i) Convolution layer: a kernel is selected that goes over the input vector that produces a feature map $x_{i,j} = \sigma((W * Y)_{i,j} + b)$, where $x_{i,j}$ is the output of the convolution operator, W is the kernel with goes over, Y is the input, σ denotes the nonlinearity in the network, and b is the bias [21–23].
- (ii) Pooling layer: the dominant features are extracted by selecting a window that passes through the pooling function, average pooling, max-pooling, or stochastic pooling [24].
- (iii) Fully connected layer: the convolution and pooling outputs are included here, and the final dot product of input and weight vector is computed in this layer

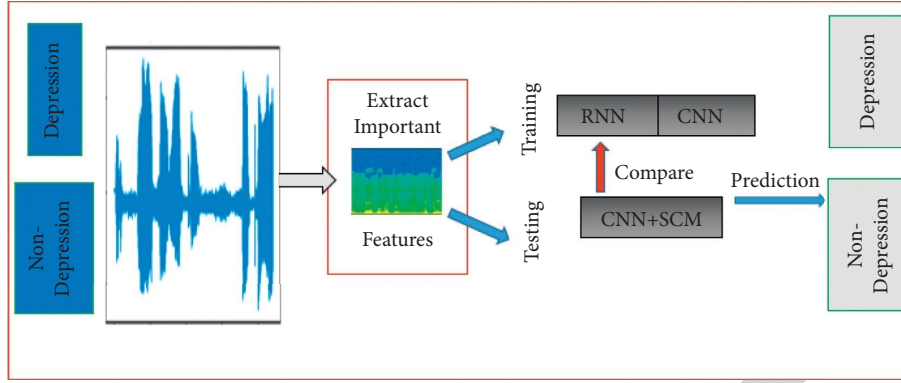


FIGURE 1: The proposed research architecture.

- (iv) Activation function: sigmoid (it takes values between $[0, 1]$) also called logistic function; in CNN, its use may cause vanishing or gradient $f(x) = (1/1 + e^{-x})$ and [14] softmax (it takes a vector argument and transforms to a vector whose elements fall in the range $[0, 1]$). When all our dependent variables are categorical, then softmax function is appropriate $f(x) = (e^{z_i} / \sum_j e^{z_j})$, and ReLU does not allow the gradient to vanish $f(x) = \max(0, x)$ for values greater than zero; it is linear [24].
- (v) Support vector machine (SVM): it is a nonparametric supervised machine learning technique employed to classify data by fitting a hyperplane to the data [25, 26]. There are different types of SVM learning mechanisms to classify the data; for this purpose, a kernel (kernel selected to make nonlinear data linearly separable) is fitted to the data; the most commonly used kernels are Gaussian $K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$ and sigmoid $K(x_i, x_j) = \tanh(\gamma(x_i, x_j) + r)$ [27]. The dense layer of the CNN model is used to make the hybrid approach for depression prediction. The architecture of the proposed model is explained in Table 1.

2.1. Recurrent Neural Network (RNN). RNN is normally used to analyze sequential data (e.g., speech, text); just like other neural networks, it contains input, hidden and, output layers [28]. The hidden layer, called the recurrent layer, keeps the same parameters in the following layers that keep on updating in its memory, $h(t) = f(Wx(t) + Uh(t-1))$, where W and U are weight matrices, $x(t)$. The input vector is $h(t-1)$, and the correlated hidden layer and f represent the nonlinear activation function [28–30]. In the hidden layer, different activation functions are used. The most commonly used are sigmoid and tanh: sigmoid function $f(x) = 1/1 + e^{-x}$ [29] and tanh function $h(t)$ with range $(-1, 1)$ [28]. In the output layer, the softmax function is used $y(t) = g(Vh(t))$, where $f(x) = e^{V_i h(t)} / \sum_j e^{V_j h(t)}$ for the final output [28, 29]. The architecture of RNN is explained in Figure 2.

The proposed hybrid approach and individual CNN and RNN are applied to diagnose depression while speaking

TABLE 1: Architecture of proposed hybrid model.

Layers	CNN + SVM		
	Results	Parameters	
Conv1D1	(Nil, 202, 32)	128	
Max pool 1	(Nil, 101, 32)	0	
Conv1D2	(Nil, 101, 64)	6,208	
Max pool 2	(Nil, 50, 64)	0	
Conv1D3	(Nil, 50, 64)	12,352	
Max pool 3	(Nil, 25, 64)	0	
Dropout	(Nil, 25, 64)	0	
Flatten	(Nil, 1,600)	0	
Dense 1	(Nil, 128)	204,928	Used for SVM
Dense 2	(Nil, 1)	129	
Total params: 223,745			
Trainable params: 223,745			

Arabic. The training-testing criteria are adopted in the analysis for 200 speeches. A total of 70% (140 speeches) of data are used as a training part, and 30% (60 speeches) of data are used as a testing part. The train data is used to train the CNN + SVM, RNN, and CNN, and test data is used to check the validity of all models and the prediction rate of the trained sample. The accuracy, area under curve (AUC), sensitivity, specificity, false-positive rate (FPR), and false-negative rate (FNR) are calculated to observe the model's performance in depth using the following equations.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

$$\text{AUC} = \int \text{TP Rate d}(\text{FP Rate}),$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{FNR} = \frac{\text{FN}}{\text{TPe} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

(1)

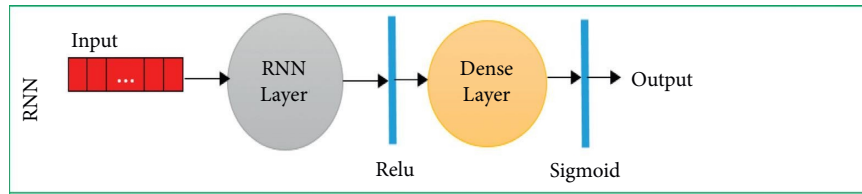


FIGURE 2: Architecture of RNN.

where FP stands for false positive, TN for true negative, TP for true positive, and FN for true negative.

The Receiver Operating Characteristic (ROC) curve is also drawn to check the model accuracy by plotting the sensitivity and specificity [31].

3. Experimental Results and Performance Analysis

Using Arabic speech analysis, the study predicts depression disorder and compares it with DL models such as RNN and CNN. Out of 100% of the data, 70% of data are used for training and 30% for testing stages.

3.1. Data Description. In this study, we used the Basic Arabic Vocal Emotions Dataset (BAVED), composed of Arabic words spelt in different levels of emotions recorded in an audio format <https://www.kaggle.com/a13x10/basic-arabic-vocal-emotions-dataset>. In experiments, we included seven words, 0 for “like,” 1 for “unlike,” 2 for “this,” 3 for “file,” 4 for “good,” 5 for “neutral” and 6 for “bad.” The seven words are further classified according to their emotional intensity: 0 denotes low emotion including tired or weary, 1 denotes neutral emotion, and 2 denotes strong emotion of happiness, joy, sadness, and anger. The categories labelled as 0 and 1 are for low and neutral emotions that represent nondepression (sadness) and negative emotions (anger).

3.2. Hybrid Model Performance. First, we applied the proposed hybrid model to the data. As a result, we attained a 90% accuracy rate to classify the depression while speaking in the training part and a 91.60% accuracy rate to predict the depression from the testing part. The graphical representation of the accuracy of the CNN + SVM model with a bar chart on train and test data is presented in Figure 3. The red color presents the accuracy of the training data and the blue color presents the accuracy of testing data.

Correctly classifying the depression speeches present in diagonal and off-diagonal values shows incorrect speech prediction. The hybrid model has accurately predicted a total of 126 (depression = 68, nondepression = 58) speeches and 14 speeches incorrectly predicted for the training data set. Similarly, the RNN model has accurately predicted 55 (depression = 31, nondepression = 24) speeches and 5 speeches not correctly predicted for the test data set. Figure 4 presents confusion matrix results of the hybrid model on train and test data.

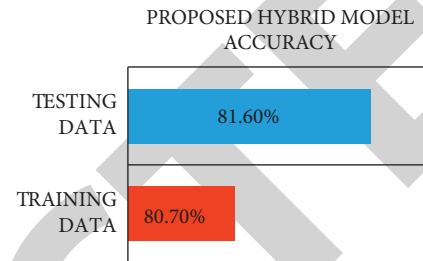


FIGURE 3: Hybrid model accuracy on training and testing data.

3.3. Individual RNN and CNN Models Performance. RNN and CNN individually applied the data where the RNN achieved an 80.70% accuracy rate to predict the depression while speaking in the training part and got an 81.60% accuracy rate for the testing part. Similarly, CNN attained an 88.5% accuracy rate to predict the depression while speaking in the training part and attained an 86.60% accuracy rate for the testing part. The accuracies attained in the training and testing stages of RNN and CNN models are exhibited in Figure 5. The red color presents the accuracy of the training data and the blue color presents the accuracy of testing data.

The training and testing loss and accuracy are measured for RNN and CNN models are plotted against the 25 epochs shown in Figure 6. The blue and red solid lines represent the accuracies of the RNN and CNN model for train and test data. The dotted blue and red solid lines present the losses of the RNN and CNN model with respect to training and testing data. It is observed that initially, network loss is higher but as epochs increase, the loss shows a decreasing trend in all models [32].

The results of RNN and CNN models with respect to the confusion matrix on train and test data are presented in Figure 7. The correctly classified depressed speeches are presented in diagonal and off-diagonal values presented as the incorrect classified prediction speech. The RNN model has accurately predicted a total of 113 (depression = 69, nondepression = 44) speeches and 27 speeches incorrectly predicted for the training data set. Likewise, the RNN model has predicted a total of 49 (depression = 31, nondepression = 18) speeches accurately and 11 speeches incorrectly predicted for the testing data set. On the other hand, the CNN model has predicted a total of 124 (depression = 66, nondepression = 58) speeches accurately and 16 speeches incorrectly on the train data set. Correspondingly, the CNN model has predicted a total of 52 (depression = 29, nondepression = 23) speeches accurately and 8 speeches incorrectly on the test data set.

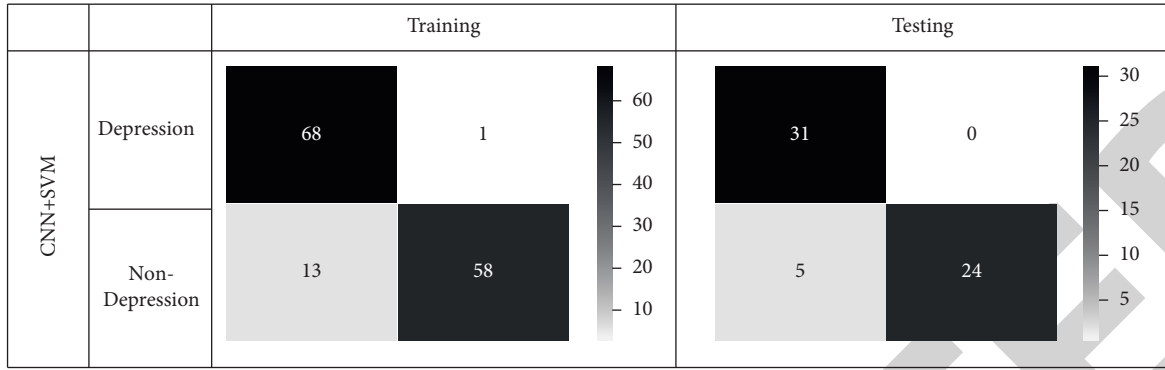


FIGURE 4: Confusion matrix results of the hybrid model on train and test data.

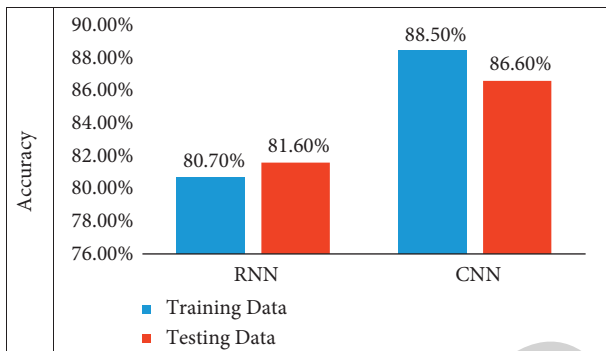


FIGURE 5: RNN and CNN accuracies comparisons for training and testing data.

4. Comparisons of Proposed Hybrid Model with RNN and CNN

4.1. Sensitivity Analysis. The assessment of the models is checked with sensitivity, specificity, FPR, and FNR for both train and test data given in Table 2. Sensitivity and specificity represent a model that correctly identifies depression and nondepression speech if it belongs to depression and nondepression speeches. The FPR and FNR are probabilities showing that a model predicts depression but it belongs to nondepression and predicts nondepression while it belongs to depression [33]. For the training data set, the RNN model achieved the 100%, 61.9%, 0.0, and 0.380 of sensitivity, specificity, FPR, and FNR, respectively. Similarly, for the testing data set, 100%, 62%, 0.0, and 0.379 of sensitivity, specificity, FPR, and FNR, respectively. The CNN model achieved the 95.6%, 81.6%, 0.043, and 0.183 of sensitivity, specificity, FPR, and FNR, respectively, for the training data set. Similarly, 93.5%, 79.3%, 0.064, and 0.206 of sensitivity, specificity, FPR, and FNR, respectively, were attained for the testing data set. The proposed hybrid model achieved the 98.5%, 81.6%, 0.014, and 0.181 of sensitivity, specificity, FPR, and FNR, respectively, for the training data set. Similarly, for testing the data set, 100%, 82.7%, 0.0, and 0.172 of sensitivity, specificity, FPR, and FNR, respectively, were attained. The performance also measured by calculating precision, recall, and F1-score. The hybrid model achieved high precision, recall, and F1-score than individually RNN and CNN. The

precision, recall, and F1-score values of the proposed hybrid model were 0.983, 0.816, and 0.892 for training data, respectively. Similarly, 1, 0.827, and 0.905 values were achieved for precision, recall, and F1-score, respectively, for testing data for the proposed hybrid model as presented in Table 3.

4.2. ROC Curve Analysis. The ROC curve is used to plot the sensitivity and specificity of training and testing data. The ROC curve values 0.70–0.80, >0.80 and >0.90 are acceptable, excellent and rarely observed [34]. The ROC with AUC of the RNN, CNN, and CNN + SVM model based on speech analysis is shown in Figure 8.

The hybrid approach provided the minimum FPR, FNR, and a higher sensitivity and specificity rate than the RNN and CNN model to predict the depression in the Arabic language.

4.3. Discussion and Comparisons. The study is designed to predict depression using speech or while speaking in the Arabic language with the proposed hybrid approach and compare it with deep learning (DL) models such as RNN and CNN. All approaches are used to diagnose depression while speaking in the Arabic language. The training-testing approach is adopted in our analysis. A total of 70% of data are used as the training part, and 30% of data are used as the testing part. The CNN + SVM is 90.0% and 91.60% that correctly predict the depression while speaking in the training and testing. Overall, the hybrid approach (CNN + SVM) provided better results than RNN and CNN in the same data set. The CNN + SVM provides better results or accuracy than the individual approach in speech data [35]. The RNN has 80.70% and 81.60% that correctly predict depression while speaking in training and testing. Comparably, the CNN has 88.50% and 86.60% that correctly predict depression while speaking in training and testing stages. While the proposed hybrid model predicted 126 speeches correctly and 14 speeches incorrectly for the training data set. Also, it has predicted 55 speeches correctly and 5 speeches not correctly for the testing data set. The RNN model mispredicted 113 speeches correctly and 27 for the training data set. Similarly, the testing data set has predicted 49 speeches correctly and 11 incorrectly. The CNN model mispredicted 124 speeches correctly and 16 for the

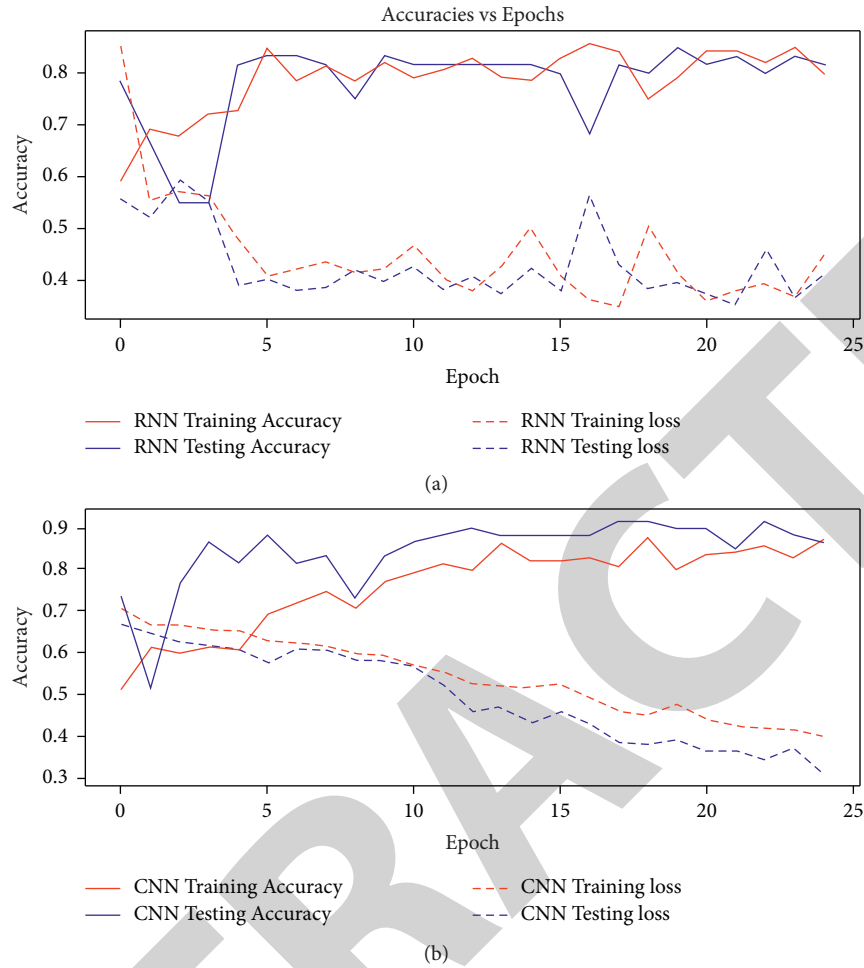


FIGURE 6: RNN and CNN accuracies vs loss against 25 epochs.

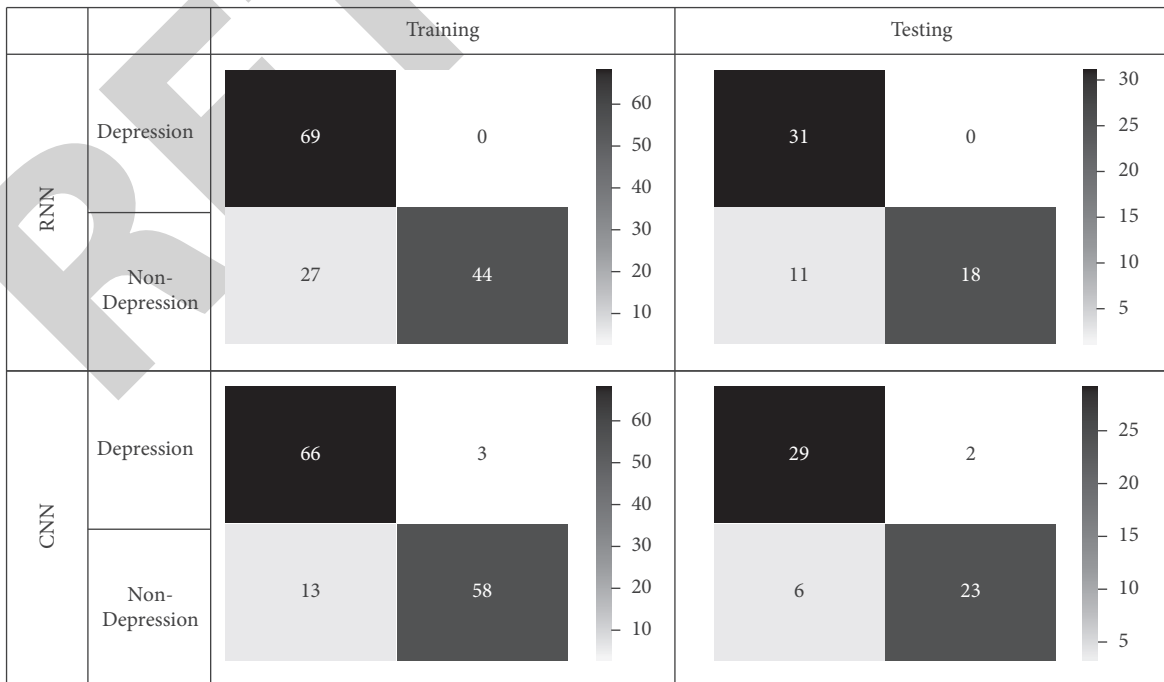


FIGURE 7: The confusion matrix results of RNN and CNN models for training and testing data.

TABLE 2: Performance comparisons of hybrid model with RNN and CNN.

		Accuracy (%)	AUC	Sensitivity (%)	Specificity (%)	FPR	FNR
Training	RNN	80.70	0.81	100	61.9	0.0	0.380
	CNN	88.50	0.89	95.6	81.6	0.043	0.183
	CNN + SVM	90	0.90	98.5	81.6	0.014	0.183
Testing	RNN	81.60	0.81	100	62	0.0	0.379
	CNN	86.60	0.86	93.5	79.3	0.064	0.206
	CNN + SVM	91.60	0.91	100	82.7	0.0	0.172

TABLE 3: Performance measured with precision, recall, and f1-score.

		Precision	Recall	F1-score
Training	RNN	1	0.619718	0.765217
	CNN	0.95082	0.816901	0.878788
	CNN + SVM	0.983051	0.816901	0.892308
Testing	RNN	1	0.62069	0.765957
	CNN	0.92	0.793103	0.851852
	CNN + SVM	1	0.827586	0.90566

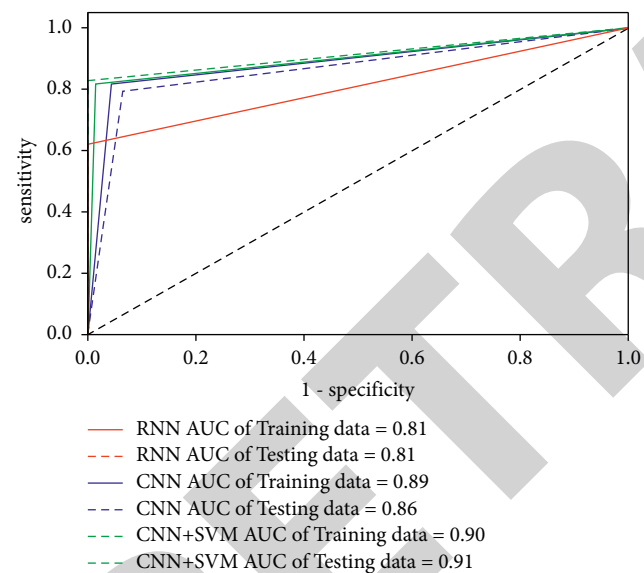


FIGURE 8: The ROC with AUC of the RNN, CNN, and hybrid model based on Arabic speech analysis.

training data set. The testing data set has predicted 52 speeches correctly and 8 incorrectly. The CNN + SVM model achieved the 98.5%, 81.6%, 0.014, and 0.181 of sensitivity, specificity, FPR, and FNR, respectively, for the training data set. Similarly, for testing the data set, it achieved the 100%, 82.7%, 0.0, and 0.172 of sensitivity, specificity, FPR, and FNR, respectively. For the training data set, the RNN model achieved the 100%, 61.9%, 0.0, and 0.380 of sensitivity, specificity, FPR, and FNR, respectively. Correspondingly, for the testing data set, it achieved the 100%, 62%, 0.0, and 0.379 of sensitivity, specificity, FPR, and FNR, respectively. The CNN model achieved the 95.6%, 81.6%, 0.043, and 0.183 of sensitivity, specificity, FPR, and FNR, respectively, for the training data set, while 93.5%, 79.3%, 0.064, and 0.206 of sensitivity, specificity, FPR, and FNR, respectively, for

testing data set. Sometimes, testing accuracy is found high than training data, but the model will consider as generalized fine. The precision, recall, and F1-score values of the proposed hybrid model were 0.983, 0.816, and 0.892 for training data, respectively. Similarly, 1, 0.827, and 0.905 values were got for precision, recall, and F1-score, respectively, for testing data for the proposed hybrid model.

The AUC value of the RNN model is found 0.81 on train and test data. Additionally, the AUC value of the CNN model is found 0.89 and 0.86 on train and test data. Comparably, the AUC value of the hybrid model is found 0.90 and 0.91 on train and test data. Based on all criteria, the hybrid model correctly identifies the depression while speaking than RNN and CNN model individually. In addition, the hybrid approach provided the minimum FPR, FNR, and higher sensitivity specificity rate than the RNN and CNN model to predict depression in the Arabic speech.

5. Conclusion

This paper has presented a hybrid model to classify depression for mental illness prediction from Arabic speech analysis. Additionally, for the same task, two deep learning models RNN and CNN are also applied individually on the same benchmark database to analyze and compare the results using standard training-testing criteria. The proposed hybrid model attained 90.0% and 91.60% correctly predicted depression while speaking on train and test data. The RNN is 80.70% and 81.60% correctly predicted depression while speaking in training and testing, respectively. The CNN has 88.50% and 86.60% that correctly predict depression while speaking in training and testing. Overall, the hybrid approach provided better results than RNN and CNN on the same benchmark database.

Moreover, the hybrid approach came out with minimum FPR and FNR. It provided a higher sensitivity and specificity rate than the RNN and CNN model to predict depression in the Arabic language. These research findings will be helpful to detect depression while speaking or in Arabic speech. Therefore, doctors, psychiatrists, or psychologists can use our approaches in healthcare applications to see depression while speaking. The doctors could also utilize the proposed approach to identify or separate the depression from neutral or normal speaking. Using our model researcher will detect depression while speaking the Arabic language with an approximately 92% accuracy rate. The proposed model could be used as a tool in the voice recognition field to detect depression while speaking the Arabic language. Depressed

persons will refer to psychiatrist for their therapies and their treatments.

Data Availability

The open-access data set employed for experiments is detailed below Basic Arabic Vocal Emotions Dataset (BAVED), composed of Arabic words spelt in different levels of emotions recorded in an audio format <https://www.kaggle.com/a13x10/basic-arabic-vocal-emotions-dataset>. The data were selected from the data source available online. However, its size was not significant enough. In the future, we will use a huge data size taken from different races (depression speeches and nondepression speeches) for the classification/identification of depression while speaking in different languages using the proposed method.

Conflicts of Interest

The authors declare that there are no conflicts of interest for this research.

Authors' Contributions

All authors contributed equally scientifically.

Acknowledgments

This research was supported by Artificial Intelligence and Data Analytics Lab (AIDA), CCIS, Prince Sultan University, Riyadh, Saudi Arabia. The authors also would like to acknowledge the support of Prince Sultan University for paying the APC of this publication.

References

- [1] O. Mohamed and S. A. Aly, "Arabic speech emotion recognition employing wav2vec2. 0 and hubert based on baved dataset," 2021, <https://arxiv.org/abs/2110.04425>.
- [2] B. Li, J. Zhu, and C. Wang, "Depression severity prediction by multi-model fusion," in *Proceedings of the HEALTHINFO 2018: The Third International Conference on Informatics and Assistive Technologies for HealthCare*, pp. 19–24, Medical Support and Wellbeing, Nice, France, 2018.
- [3] W. H. O. Depression, *Other Common Mental Disorders: Global Health Estimates*, pp. 1–24, World Health Organization, Geneva, Switzerland, 2017.
- [4] T. Saba, A. Rehman, M. N. Shahzad et al., "Machine learning for post-traumatic stress disorder identification utilizing resting-state functional magnetic resonance imaging," *Microscopy Research and Technique*, vol. 2021, no. 80, 2022.
- [5] M. N. Shahzad, H. Ali, T. Saba, A. Rehman, H. Kolivand, and S. A. Bahaj, "Identifying patients with PTSD utilizing resting-state fMRI data and neural network approach," *IEEE Access*, vol. 9, pp. 107941–107954, 2021.
- [6] P. Fusar-Poli, B. Nelson, L. Valmaggia, A. R. Yung, and P. K. McGuire, "Comorbid depressive and anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and transition to psychosis," *Schizophrenia Bulletin*, vol. 40, no. 1, pp. 120–131, 2014.
- [7] A. Vázquez-Romero and A. Gallardo-Antolín, "Automatic detection of depression in speech using ensemble convolutional neural networks," *Entropy*, vol. 22, no. 6, p. 688, 2020.
- [8] M. A. Khan, S. Kadry, Y.-D. Zhang et al., "Prediction of COVID-19 - pneumonia based on selected deep features and one class kernel extreme learning machine," *Computers & Electrical Engineering*, vol. 90, Article ID 106960, 2021.
- [9] H. Wang, Y. Liu, X. Zhen, and X. Tu, "Depression speech recognition with a three-dimensional convolutional network," *Frontiers in Human Neuroscience*, vol. 15, 2021.
- [10] A. Saidi, S. B. Othman, and S. B. Saoud, "Hybrid CNN-SVM classifier for efficient depression detection system," in *Proceedings of the 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET)*, pp. 229–234, IEEE, Manhattan, New York, December 2020.
- [11] S. Yun and C. D. Yoo, "Loss-scaled large-margin Gaussian mixture models for speech emotion classification," *IEEE Transactions on Audio Speech and Language Processing*, vol. 20, no. 2, pp. 585–598, 2011.
- [12] J. R. Williamson, T. F. Quatieri, B. S. Helfer, R. Horwitz, B. Yu, and D. D. Mehta, "Vocal biomarkers of depression based on motor incoordination," in *Proceedings of the 3rd ACM International Workshop on Audio/visual Emotion challenge*, pp. 41–48, Barcelona Spain, October 2013.
- [13] D. Le and E. M. Provost, "Emotion recognition from spontaneous speech using hidden Markov models with deep belief networks," in *Proceedings of the 2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 216–221, IEEE, Manhattan, New York, December 2013.
- [14] Y. H. Kao and L. S. Lee, "Feature analysis for emotion recognition from Mandarin speech considering the special characteristics of Chinese language," in *Proceedings of the InterSpeech*, Pittsburgh, PA, USA, September 2006.
- [15] M. Yousuf, Z. Mehmood, H. A. Habib et al., "A novel technique based on visual words fusion analysis of sparse features for effective content-based image retrieval," *Mathematical Problems in Engineering*, vol. 2018, Article ID 2134395, 13 pages, 2018.
- [16] K. Han, D. Yu, and I. Tashev, "Speech emotion recognition using deep neural network and extreme learning machine," in *Proceedings of the Interspeech 2014*, Singapore, Malaysia, September 2014.
- [17] D. Bertero and P. Fung, "A first look into a convolutional neural network for speech emotion detection," in *Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5115–5119, IEEE, New Orleans, LA, USA, 2017, March.
- [18] K. Cho, B. Van Merriënboer, C. Gulcehre et al., "Learning phrase representations using rnn encoder-decoder for statistical machine translation," 2014, <https://arxiv.org/abs/1406.1078>.
- [19] J. Bradbury, S. Merity, C. Xiong, and R. Socher, "Quasi-recurrent neural networks," 2016, <https://arxiv.org/abs/1611.01576>.
- [20] S. Basu, J. Chakraborty, and M. Aftabuddin, "Emotion recognition from speech using convolutional neural network with recurrent neural network architecture," in *Proceedings of the 2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, pp. 333–336, IEEE, Coimbatore, India, October 2017.
- [21] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network-A deep learning approach," *Procedia Computer Science*, vol. 132, pp. 679–688, 2018.

Retraction

Retracted: Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] W. Yang, X. Pan, Y. Wang, and W. Chen, "Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6861749, 7 pages, 2022.

Research Article

Comparative Meta-Analysis of the Effects of OLIF and TLIF in Lumbar Spondylolisthesis Central Nerve Injury

Wanliang Yang ¹, Xin Pan ¹, Yibo Wang,² and Wenhao Chen³

¹Qilu Hospital of Shandong University, Jinan 250000, Shandong, China

²Dingtao District People's Hospital, Jinan 250000, Shandong, China

³Qilu Hospital of Shandong University, Jinan 250000, Shandong, China

Correspondence should be addressed to Wanliang Yang; 1211410523@st.usst.edu.cn

Received 17 January 2022; Revised 14 February 2022; Accepted 18 March 2022; Published 10 May 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Wanliang Yang 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.

Objective. The main objective is to explore the efficacy of oblique anterior lumbar fusion (OLIF) and transforaminal lumbar fusion (TLIF) in the treatment of lumbar spondylolisthesis central nerve injury. **Methods.** The perioperative indexes, pain score (VAS), Oswestry dysfunction index (ODI), vertebral slip degree, slip angle, intervertebral space height, and quality of life score of the two groups were compared by meta-analysis. **Results.** According to the observation indexes, the perioperative indexes of patients in the OLIF group were better than those in the TLIF group, which showed that the effect of OLIF treatment was better than of TLIF. The pain score and ODI score of the two groups can be obtained. The one-week postoperative pain degree and ODI of patients in the OLIF group are lower than those in the TLIF group, indicating that OLIF treatment will reduce the pain of patients to a greater extent and is more conducive to the recovery of patients. There was no significant difference in vertebral slip, slip angle, and intervertebral space height between the OLIF group and TLIF group. After treatment, the quality-of-life scores of patients in the OLIF group were significantly higher than those in the TLIF group. **Conclusion.** The treatment of lumbar fusion through OLIF has irreplaceable perioperative advantages of TLIF, such as less bleeding, shorter operation time, less drainage and shorter hospital stay, less postoperative complications, less surgical wound, indirect decompression, no destruction of lumbar posterior stable structure, and maximum preservation of tissue structure. It has the advantages of reducing the intraoperative dural sac injury and nerve root traction injury and shortening the rehabilitation time of patients. It has the prospect of clinical application and can be popularized.

1. Introduction

Lumbar spondylolisthesis is divided into congenital and postnatal. Congenital spondylolisthesis is not repeated. Most acquired lumbar spondylolisthesis is caused by degenerative, chronic strain, or trauma [1]. More than 60% of the posterior lumbar spondylolisthesis is caused by degenerative factors. Therefore, the lumbar spondylolisthesis in this study specifically refers to degenerative lumbar spondylolisthesis, and the general degenerative lumbar spondylolisthesis will be accompanied by lumbar spinal stenosis, which can be recovered only through surgical treatment [2].

Lumbar interbody fusion is the main surgical method for the treatment of degenerative lumbar spondylolisthesis. The

most traditional treatment methods are posterior decompression and interbody fusion (PLIF). However, even if the treatment restores the patient's health, it will bring more serious complications to the patient. Therefore, there is the emergence of lumbar interbody fusion (TLIF) through the intervertebral foramen. It is an optimized surgical treatment based on PLIF. Jun et al. [3] showed that lumbar interbody fusion via the intervertebral foramen approach is relatively simple and safe and has the advantages of good spinal stability [3]. Oblique anterior approach lumbar interbody fusion (OLIF) was first proposed in 2012. However, it needs to be studied to determine which kind of lumbar interbody fusion is better. At present, this kind of disease mainly occurs in the middle-aged and elderly groups, which will lead to a

decline in patients' quality of life and inconvenient movement. When the effect of conservative treatment is not good, surgery needs to be taken to improve. Hongjun et al. [4] showed that in the treatment of lumbar spondylolisthesis, both OLIF and TLIF achieved good therapeutic effects, but through OLIF, the pain and dysfunction of short-term postoperative pain will be reduced to a greater extent [4].

Therefore, the purpose of this study is to analyze the therapeutic effect of OLIF and TLIF on lumbar spondylolisthesis through meta-analysis, reduce patients' pain and complications, and help patients recover faster and better.

2. Data and Methods

2.1. General Information Introduction. 120 patients with lumbar spondylolisthesis hospitalized from July 2019 to January 2020 were selected, including 34 male patients and 86 female patients, with an average age of 55.5 ± 6.1 years. There were 81 cases of L4 slippage and 39 cases of L5 slippage. The gender, age, course of the disease, and symptoms were not statistically significant ($P > 0.05$), which is comparable. Moreover, this study has submitted an application to the clinical trial ethics committee and obtained its approval. In this study, all patients were randomly divided into the TLIF group and OLIF group, with 60 cases. The TLIF group was treated with lumbar interbody fusion via foraminal approach, including 16 male patients and 44 female patients. There were 41 patients with L4 spondylolisthesis and 19 patients with L5 spondylolisthesis. Oblique anterior lumbar interbody fusion was performed in the OLIF group, including 18 male patients and 42 female patients. There were 40 patients with L4 spondylolisthesis and 20 patients with L5 spondylolisthesis. Both groups were operated by the same group of doctors (see Table 1 for detailed general information).

Case inclusion criteria were as follows: ① patients with clinical characteristics of lumbar spondylolisthesis; ② patients without absolute surgical contraindication; ③ patients whose follow-up survey at each stage has been completed; ④ patients who knew and agreed to the study and signed the informed consent form.

Case exclusion criteria were as follows: ① patients with lumbar trauma; ② patients with mental diseases; ③ patients with previous lumbar surgery; ④ patients with other surgical contraindications.

In Table 1, it can be seen that most of the hospitalized patients are over 40 years old, and the number of female patients is much more than that of male patients.

2.2. Observation Indicators and Evaluation Criteria. According to the patient's medical records, the perioperative indicators of the patient can be extracted. The pain (VAS) score is widely used in the clinic. It can also reflect the advantages and disadvantages of surgery through the evaluation of this index. Oswestry dysfunction index (ODI) is a questionnaire used to evaluate the degree of functional impairment caused by chronic low back pain. Therefore, it can be used by comparing the perioperative indexes of the OLIF group and TLIF group; pain score before and after

TABLE 1: General information of patients.

General data classification		Number of people (%)
Gender	Males	34 (28.33)
	Females	86 (71.67)
Age	<40	14 (11.67)
	40-50	41 (34.17)
	>50	65 (54.16)
Symptom	L4	81 (67.5)
	L5	39 (32.5)

operation and Oswestry dysfunction index (ODI); vertebral slip, slip angle, intervertebral space height, and quality of life score before and after operation. A visual pain simulation scale (VAS) is used to measure and evaluate the pain degree of patients before and after operation. The score range is 0-10, 0 without any pain, 10 for severe pain unbearable and seriously affecting life and sleep. ODI is composed of 10 aspects, including pain intensity, self-care, lifting, walking, sitting, standing, and interference. In terms of sleep, sexual life, social life, and tourism, there are 6 options in each aspect. The highest score of each option is 5 points, the lowest score of the first option is 0 points, and the highest score of the last option is 5 points. The scoring method is calculated according to the percentage of the total score of the number of questions answered. The higher the score, the more serious the dysfunction. The evaluation of the quality of life is based on the world general WHO quality of life-brief table (WHOQOL-BREF). The brief table includes 6 fields, with 0-20 points in each field. The higher the score, the better the quality of life.

2.3. Statistical Methods. The meta-analysis of the extracted data is carried out by using the Review Manager 52.6 software, which is combined with the SPSS 220 statistical software for analysis and processing. The measurement data are expressed in $(\bar{x} \pm s)$, the bivariate t -test is adopted, and the percentage rate of counting data is expressed in (%). When $T < 10.000$, $P < 0.05$ and the difference in statistical data is statistically significant, which can be compared by statistical methods. When $T < 10.000$, $P < 0.01$ and the statistical data are considered to be statistically significant.

3. Results

3.1. Comparison of Perioperative Indexes between the OLIF Group and TLIF Group. Perioperative indicators include surgical bleeding volume, operation time, postoperative drainage volume, and hospital stay. According to the statistical analysis of perioperative indicators of the two groups, all perioperative indicators of the OLIF group are significantly better than those of the TLIF group. The detailed data are shown in Table 2.

Table 2 shows that the perioperative index data of surgical bleeding volume, postoperative drainage volume, operation and hospitalization time show that there is statistical significance in the OLIF group and TLIF group ($P < 0.05$), and it can be seen from the data that the surgical

TABLE 2: The amount of surgical bleeding, postoperative drainage, operation, and hospital stay were compared between the two groups ($\bar{x} \pm s$).

Group	N	Surgical bleeding (ml)	Operation time (min)	Postoperative drainage (ml)	Length of stay (d)
OLIF group	60	176.4 ± 32.5	119.2 ± 8.7	162.7 ± 56.7	5.2 ± 0.6
TLIF group	60	221.8 ± 29.6	57.6 ± 8.9	248.6 ± 83.1	9.7 ± 0.5
<i>t</i> value	—	1.791	1.356	1.075	2.046
<i>p</i> value	—	<0.05	<0.05	<0.05	<0.05

bleeding volume and drainage volume of patients in the OLIF group are significantly less than those in the TLIF group, and the operation and hospitalization time are also significantly shorter than those in the TLIF group, indicating that the early curative effect of OLIF treatment is better.

Figure 1 shows that the treatment methods of the OLIF group are more conducive to the postoperative recovery of patients than those of the TLIF group.

3.2. Comparison of VAS Score and Oswestry Dysfunction Index between the OLIF Group and TLIF Group before and after Operation. Through the statistical analysis of VAS score data and ODI data of patients in the OLIF group and TLIF group before operation, after operation, and 7 d and 90 d after operation, it can be concluded that the VAS score and ODI of patients in the two groups are significantly better than those after operation, while the one-week postoperative pain course and ODI of patients in the OLIF group are lower than those in the TLIF group. It shows that OLIF treatment will reduce the pain of patients to a greater extent and is more conducive to the recovery of patients (see Table 3 for detailed data).

In Table 3, there was no significant difference in VAS score and ODI score between the two groups before operation ($P > 0.05$), but there was a significant difference in scores after 7 days of operation ($P < 0.05$), and the scores of patients in the OLIF group were significantly lower than those in the TLIF group, which showed that patients in the OLIF group recovered better after operation. However, there was no significant difference in scores after 90 days of operation.

Table 3 and Figure 2 show that the patients in the OLIF group have lower scores on postoperative pain and dysfunction than those in the TLIF group, which also shows that the treatment method of OLIF is better.

3.3. Comparison of Vertebral Slip, Slip Angle, and Intervertebral Space Height between the OLIF Group and TLIF Group before and after Operation. Through the statistical analysis of the vertebral slip degree, slip angle, and intervertebral space height of the two groups of patients with different surgical treatment before and 90 days after operation, it can be concluded that the slip degree and slip angle of patients in the OLIF group and TLIF group are significantly lower than those before operation, while the intervertebral space height is significantly higher. It was statistically significant ($P < 0.05$), but the comparison of various index data before and after operation between the two groups was not statistically significant ($P > 0.05$) (see Table 4 for details).

In Table 4, there was no significant difference between the two groups in vertebral slip, slip angle, and intervertebral

space height before and after operation, which was not statistically significant ($P > 0.05$), but there were significant differences in various indexes between the same group before and after operation, which showed that surgical treatment was positive for the course of the disease, but different surgical methods had no significant difference for the treatment effect.

In Table 4 and Figure 3, the parameters of vertebral body slip, slip angle, and intervertebral space height before and after operation in the same group were significantly compared, which showed that the surgical treatment had a significant effect on the patient's condition ($P < 0.05$ compared with that before operation).

3.4. Comparison of Quality-of-Life (WHOQOL-BREF) Scores before and after Operation between the OLIF Group and TLIF Group. Through the statistical data of WHOQOL-BREF scores of patients in the OLIF group and TLIF group before and after operation, it can be seen that the WHOQOL-BREF scores of the OLIF group and TLIF group after operation in each field are higher than those before operation, but there is no statistical significance between WHOQOL-BREF scores in each field. Compared with the two groups, the WHOQOL-BREF scores of the two groups before the operation were not statistically significant, while after the operation, the WHOQOL-BREF scores of patients in the OLIF group were significantly higher than those in the TLIF group, indicating that the OLIF operation method was more beneficial to the improvement of patients' postoperative quality of life (see Table 5 for details).

In Table 5, it can be seen that the evaluation of the quality of life in various fields after operation is significantly higher than that before operation, but there is no significant difference in the evaluation of the quality of life after operation between different fields ($P > 0.05$).

The meaning expressed in Figure 4 is the same as that in Table 5. The changes in quality-of-life evaluation in various fields before and after operation were similar, indicating that the quality of life after postevaluation operation was significantly higher than that before operation.

4. Discussion

The degenerative lumbar spondylolisthesis in this study mainly occurs in the instability of lumbar intervertebral space caused by natural degeneration and hyperplasia of the lumbar intervertebral disc, the disorder of facet joints and relaxation of surrounding ligaments, and the surface slip of upper lumbar cone and lower lumbar cone. The

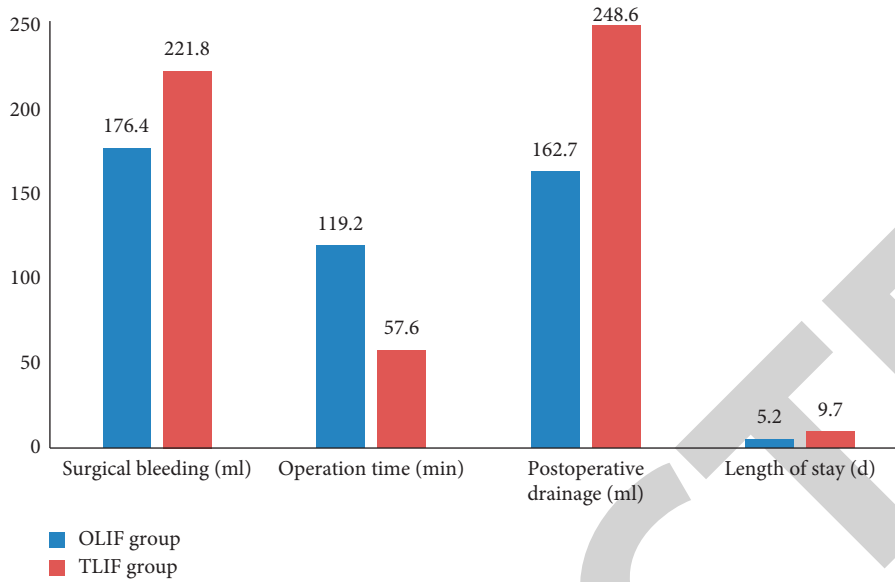


FIGURE 1: Comparison of perioperative indexes between the two groups.

TABLE 3: VAS score and ODI data analysis of patients in the OLIF group and TLIF group ($\bar{x} \pm s$).

Groups	n	VAS (branch)			ODI (%)		
		1 day before operation	7 days after operation	90 days after operation	Before operation	7 days after operation	90 days after operation
OLIF group	60	7.3 ± 1.6	3.8 ± 0.4	1.2 ± 0.2	46.3 ± 3.6	17.5 ± 2.1	13.1 ± 1.4
TLIF group	60	7.2 ± 1.8	4.7 ± 0.3	1.9 ± 0.4	44.8 ± 4.7	30.4 ± 5.2	15.6 ± 1.6
t value	—	8.624	7.301	1.254	7.235	1.062	8.621
p value	—	0.57	0.008	0.004	0.006	0.005	0.65

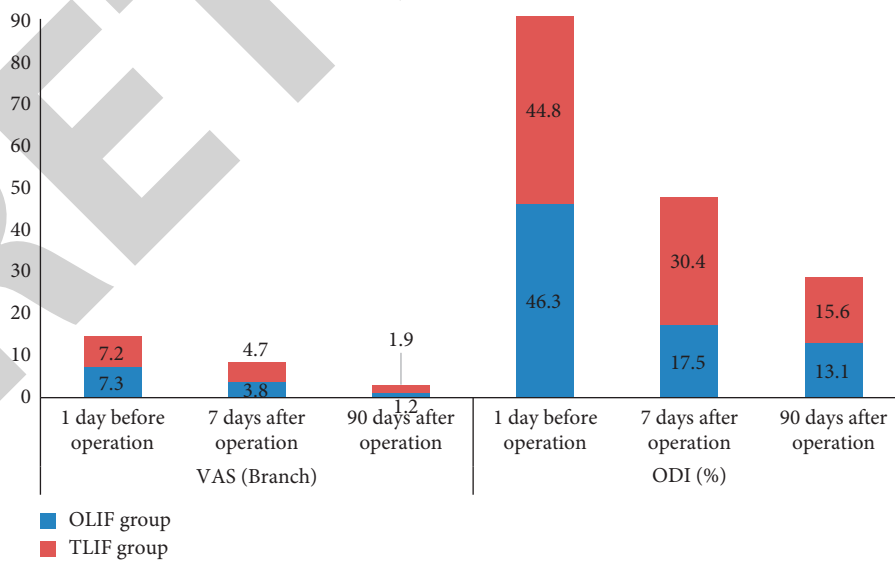


FIGURE 2: VAS score and ODI data analysis chart of patients in the OLIF group and TLIF group.

spondylolisthesis of lumbar spondylolisthesis is common in L4, L5, and sacral 1 segments, accompanied by secondary lumbar spinal stenosis. At present, the most important and

effective method for the treatment of lumbar spondylolisthesis is lumbar interbody fusion, but the method of surgical treatment cannot avoid excessive interference to the nerve

TABLE 4: Comparison of vertebral slip, slip angle, and intervertebral space height between the OLIF group and TLIF group before and after operation ($\bar{x} \pm s$).

Group	N	Slip degree (mm)		Slip angle (°)		Disc height (mm)	
		1 day before operation	90 days after operation	1 day before operation	90 days after operation	1 day before operation	90 days after operation
OLIF group	60	17.4 ± 8.4	7.9 ± 5.3	23.1 ± 7.6	10.3 ± 2.8	7.2 ± 3.4	16.1 ± 2.4
TLIF group	60	17.6 ± 9.1	8.4 ± 5.6	23.4 ± 8.1	11.8 ± 3.1	7.1 ± 3.0	15.5 ± 2.2
t value	—	8.324	7.568	8.461	7.914	8.649	6.564
p value	—	0.75	0.64	0.87	0.78	0.89	0.67

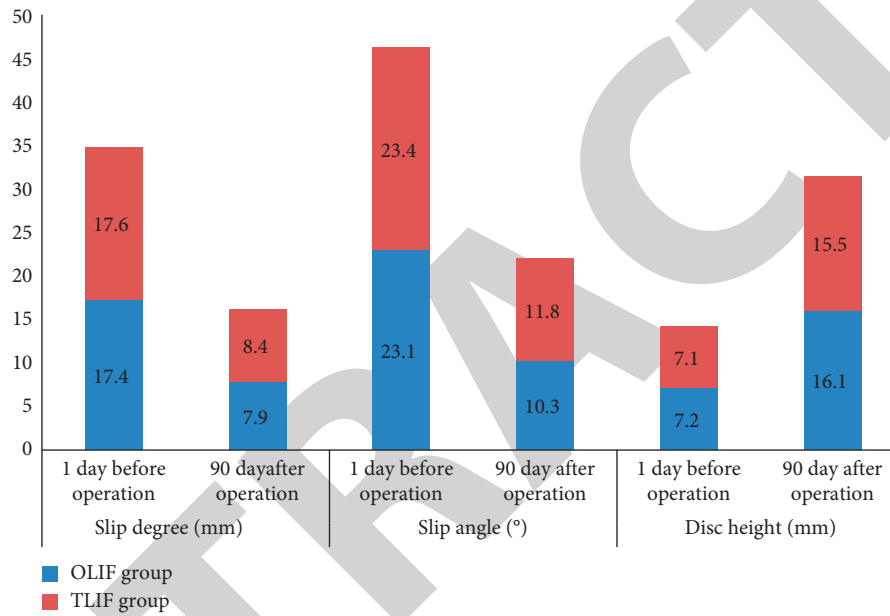


FIGURE 3: Comparison of vertebral slip, slip angle, and intervertebral space height between the two groups before and after operation.

TABLE 5: Comparison of quality-of-life (WHOQOL-BREF) scores between the two groups before and after operation ($\bar{x} \pm s$).

Groups	Physiology				Psychology			
	Before operation	After operation	t value	p value	Before operation	After operation	t value	p value
OLIF group	9.75 ± 0.8	15.74 ± 1.4	0.782	<0.05	4.863 ± 0.5	15.953 ± 0.9	0.583	<0.05
TLIF group	9.58 ± 0.9	12.03 ± 0.7	1.075	<0.05	4.906 ± 0.5	11.642 ± 0.7	0.862	<0.05
t value	11.874	8.592	—	—	11.374	8.042	—	—
p value	>0.05	>0.05	—	—	>0.05	>0.05	—	—
Group	Sociology				Environment			
	Before operation	After operation	t value	p value	Before operation	After operation	t value	p value
OLIF group	5.92 ± 0.7	16.05 ± 1.1	0.753	<0.05	4.653 ± 0.9	15.691 ± 1.0	0.795	<0.05
TLIF group	5.39 ± 0.6	12.21 ± 0.9	0.799	<0.05	4.836 ± 0.6	11.075 ± 1.1	0.804	<0.05
t value	11.742	7.989	—	—	13.659	9.004	—	—
p value	>0.05	>0.05	—	—	>0.05	>0.05	—	—
Group	Independent				Religious belief			
	Before operation	After operation	t value	p value	Before operation	After operation	t value	p value
OLIF group	7.05 ± 1.8	16.63 ± 0.4	1.491	<0.05	8.04 ± 0.8	15.97 ± 0.9	1.056	<0.05
TLIF group	7.48 ± 1.6	12.09 ± 0.6	1.832	<0.05	8.01 ± 0.7	11.58 ± 0.7	1.472	<0.05
t value	11.905	8.504	—	—	13.975	8.084	—	—
p value	>0.05	>0.05	—	—	>0.05	>0.05	—	—

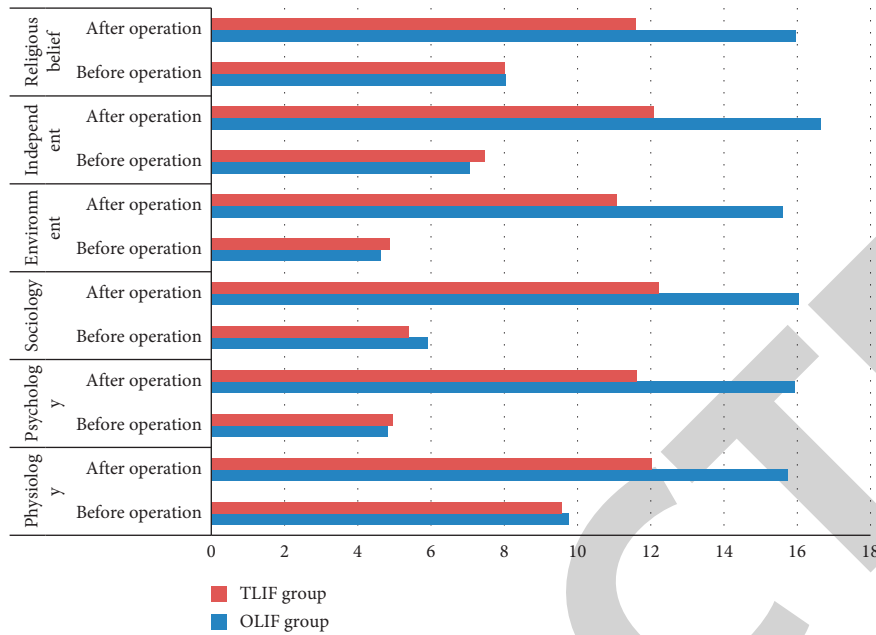


FIGURE 4: Comparison of quality of life between the two groups before and after operation.

root and damage to the structure of the posterior column of the spine, and even postoperative complications of lumbar surgery. At present, the most commonly used surgical methods are OLIF, TLIF, PLIF, and ALIF. This study mainly compares and determines the effects of OLIF and TLIF through meta-analysis.

Wu et al. [5] showed that posterior interbody fusion (PLIF) and posterolateral fusion (PLF) are internal fixation combined with bone graft for spinal fusion, which can obtain satisfactory spinal stability. However, compared with PLF, PLIF can effectively improve the degree of vertebral spondylolisthesis and does not increase the incidence of postoperative complications on the premise of ensuring the surgical effect. Therefore, it is recommended to use PLIF in the clinical treatment of lumbar spondylolisthesis [5]. Li and Wang [6] showed that although OLIF and PLIF have the same therapeutic effect in the treatment of degenerative lumbar spondylolisthesis, OLIF can more effectively shorten the operation time and reduce the amount of intraoperative bleeding than PLIF and has less soft tissue injury and less complications such as nerve injury [6]. The research of Liao et al. [7] also shows that the use of posterior interbody fusion significantly improves the clinical efficiency and bone fusion rate, reduces the incidence of complications and the probability of reoperation, and is more conducive to the treatment of lumbar spondylolisthesis. However, when selecting the surgical method, the age of the patient and the tolerance of the operation should be more considered [7]. The treatment methods used in this study are OLIF and TLIF. OLIF is a minimally invasive fusion technique to expose the lumbar spine through the space between the psoas major muscle and abdominal great vessels. Compared with TLIF, OLIF avoids the interference with the spinal canal, dural sac, or nerve root.

It can be seen from the clinical experimental data of this study that in the comparison of perioperative bleeding,

operation time, postoperative drainage, and length of stay between the two groups, it can be concluded that the perioperative indexes of patients in the OLIF group are better than those in the TLIF group, the curative effect of lumbar spondylolisthesis was studied and analyzed in the study of Xing et al. [8]. Which shows that the effect of OLIF treatment is better than the TLIF group. The pain score and ODI score of the two groups can be obtained. The one-week postoperative pain degree and ODI of patients in the OLIF group are lower than those in the TLIF group, indicating that OLIF treatment will reduce the pain of patients to a greater extent and is more conducive to the recovery of patients. There was no significant difference in vertebral slip, slip angle, and intervertebral space height between the OLIF group and the TLIF group. Wang et al. (2021) showed that the short-term treatment was compared with the two fusion [9]. Compared with the two groups, the quality-of-life scores of the two groups before treatment were not statistically significant, while after treatment, the quality-of-life scores of patients in the OLIF group were significantly higher than those in the TLIF group, which also shows that the surgical treatment of OLIF can better improve the postoperative recovery and quality of life of patients.

5. Conclusion

In this study, the patients participating in the clinical experiment were divided into the OLIF group and TLIF group by meta-analysis. Through the statistical analysis of various observation indexes of the two groups, it is concluded that the treatment method of lumbar fusion through OLIF has the advantages of perioperative surgery that cannot be replaced by TLIF, such as less surgical bleeding, shorter operation and hospital stay, and less drainage. Moreover, the treatment method of lumbar fusion of OLIF has the

Retraction

Retracted: Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Londhe, P. V. R. D. P. Rao, S. Upadhyay, and R. Jain, "Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8579640, 14 pages, 2022.

Research Article

Extracting Behavior Identification Features for Monitoring and Managing Speech-Dependent Smart Mental Illness Healthcare Systems

Alka Londhe ¹, P. V. R. D. Prasada Rao ¹, Shrikant Upadhyay ², and Rituraj Jain ³

¹Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

²Department of Electronics & Communication Engineering, Cambridge Institute of Technology, Ranchi, Jharkhand, India

³Department of Electrical and Computer Engineering, Wollega University, Nekemte, Ethiopia

Correspondence should be addressed to Rituraj Jain; jainrituraj@wollegauniversity.edu.et

Received 1 February 2022; Revised 5 March 2022; Accepted 14 March 2022; Published 6 May 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Alka Londhe 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.

Speech is one of the major communication tools to share information among people. This exchange method has a complicated construction consisting of not the best imparting of voice but additionally consisting of the transmission of many-speaker unique information. The most important aim of this research is to extract individual features through the speech-dependent health monitoring and management system; through this system, the speech data can be collected from a remote location and can be accessed. The experimental analysis shows that the proposed model has a good efficiency. Consequently, in the last 5 years, many researchers from this domain come in front to explore various aspects of speech which includes speech analysis using mechanical signs, human system interaction, speaker, and speech identification. Speech is a biometric that combines physiological and behavioural characteristics. Especially beneficial for remote attack transactions over telecommunication networks, the medical information of each person is quite a challenge, e.g., like COVID-19 where the medical team has to identify each person in a particular region that how many people got affected by some disease and took a quick measure to get protected from such diseases and what are the safety measure required. Presently, this task is the most challenging one for researchers. Therefore, speech-based mechanisms might be useful for tracking his/her voice quality or throat getting affected. By collecting the database of people matched and comparing with his/her original database, it can be identified in such scenarios. This provides the better management system without touching and maintains a safe distance data that can be gathered and processed for further medical treatment. Many research studies have been done but speech-dependent approach is quite less and it requires more work to provide such a smart system in society, and it may be possible to reduce the chances to come into contact with viral effected people in the future and protect society for the same.

1. Introduction

1.1. Background and Principle of Speaker Recognition. In this respect, speech supplies naturally and handily the shape of an entry that conveys a remarkable amount of speaker supporting information, and it is cheap to analyze and collect.

Voice is an occurrence that extremely relies on the speaker who generates it. Various physical outlooks of speech such as tone, timber, or intensity vary plenty from a speaker to another. The similar happens with supplementary

linguistic aspects such as the individual intonation and expressions a speaker normally uses or a range of vocabulary [1, 2]. All these belongings make voice a very strong biometric essential to be implemented in security structure since the physical attribute of speech is easy to calculate and compare in comparison with other biometric essentials and other medical issue analysis. The speed-dependent technology leads to better control without touching and maintaining a safe distance for information gathering for subsequent medical treatment in the healthcare system. It also protects society by decreasing the possibility of coming

into touch with a virally affected individual in the future. In inclusion, the speech wave is quite well known and has been greatly studied for many years, so many powerful algorithms can be found to deal with this kind of signal [3].

Textual content-dependent project entails a form of predicated or triggered passwords as a way to achieve the required textual content. It could be used for packages consisting of voice method signature or password confirmation. In aforesaid programs, there is a need to alternate the key or password; often it is smart to make completed without difficulty through converting the predicated text. In text-based speaker verification, at some point of the enrolment segment, a limited number of utterances of constant textual content are amassed [4]. The proposed model is more secure than the previously implemented models. This management approach, which uses speech samples, reduces the number of persons in contact, breaks the mass into pieces, and provides a better option for dealing with such a situation. Consequently, approaches primarily based on template matching are used for sample contrast rather than procedures primarily established on facts having feigned neural grids, which requires a massive quantity of learned records. Studies in the processing of speech and verbal exchange, as the maximum element, influenced by means of humans, wish to construct mechanicals prototypes to imitate human lexical conveying [5–7]. Speech processing research interest in these days is performed properly beyond the conception of copying human spoken equipment. Biomedical gains attention in different applications such as real voice analysis, operation work functioning, and viral infected throat identification [8, 9].

Addresser popularity is a system of unconsciously spotting who is expressing through the use of speaker unique statistics blanketed in speech gestures to confirm the recognition being stated via human being retrieving systems; that is, it permits to get admission to manage of diverse offerings by way of voice. Speaker identification is matched up to the anatomical and behavioural properties of the speaking manufacturing machine of a separate speaker. These properties deduce from each insubstantial pouch (vocal span properties) and prosodic capabilities (input supply traits) of articulation. The bulk usually worn quick-time period ghastrly estimations cestrum measures and their throwback measures. Therefore, for the throwback measures, usually the primary and 2nd-order measures, this is subordinate to the pulse functions of cestrum measures, which are retrieved at each body cycle, to symbolize the ghastrly energetic. Those throwback measures are also known as delta-cestrum and delta2cepstralmeasures [10].

The present article has been planned into various sections. Section 1 deals with introducing the concept and principle of speaker recognition. Section 1.1 puts light on the discussion of related research work. Section 2 illustrates the automatic addresser identification. Section 3 describes the recognition, spotting, and validation. Section 4 describes the preparation of database. The proposed model and results are described in Section 5, and finally Section 6 portrays the conclusion and possible future works based on the proposed framework.

1.2. Literature Review. Speaker verification stands on the belief that there is always some quality in every person's speech and that it can be used to verify his/her identity. The quality features in speakers' voices are used both to train a user model and to make up a reference identity for the claimed cause to be used for verification against the user's model. Feature extraction by Schmid and Gisht, 1994, suggested that the first stage of taking out information from a speech signal is known as feature extraction [10]. Both have also discussed that speech information is essential for different speech processing tasks such as speaker verification is taken as a short spectrum; as per the same author, spectral information is captured during a period of about 20 ms [11].

On the other side, Furui, 1997, discussed that an effective method to differentiate between users in a system is fundamental frequencies and different measurements of signals similar to long-term spectrum and short-term spectrum and in all powers. Gold and Morgan, 2000, discussed that text-independent and text-dependent focus on a client's input (voice) is naturally relied on process for verification [12]. For speech recognition, the correct action of speaking in terms of features is contingent on the classification technique being used. Anyhow, in some system types, Morgan and Gold, 2000, whether the verification is text-dependent, text-independent of some different group, the estimated claimant of a user model always needs the score calculation that informs the system about the distinguishable/indistinguishable vocal sounds, or statements of the claimant are matched to the model [13]. Clearly, a comparison is performed between the feature extracted and the features that the speaker model for the claimant is standing on the comparison of their spectra or both can be evaluated using statistical measures, on the basis of certainty/uncertainty between the model and the claimant's voice [14]. Morgan and Gold, 2000, also investigated that acoustic aspect degree is not frequently used as the statistical degree. Doddington, 1998, Furui, 1997, and Schmidt and Gish, 1994, discussed the environmental effect on the speech signal coming from different sources, for example, acoustic noise Doddington, 1998, and different transmission and recording environments for speech Doddington 1998, Furui, 1999 and environmental sound or other interfering voices Schmidt and Gish, 1994 [15].

A text-dependent system has no advance knowledge; whatever it may be, the claimed will throw during the session (Den Os and Boves & Doddington, 1998); therefore, it will accept or reject his/her request on the characteristics of his/her voice. This is because of reality that the system does not use vocabulary understanding, and achievements become poor compared with the text-dependent system [16].

Blombug in 2002 has underlined the major structure classes from the number of text-reliant to the number of text-self-reliant.

- (1) Text-reliant has a prearranged (fixed) key
- (2) Text-reliant defined key for every user
- (3) One's voice dependent
- (4) Action dependent, i.e., certain phonemes
- (5) System chooses the text (text-independent)

Fletcher, 1938, located the idea of a critical band that the edge of listening to change into detected for a sinusoid wave as a purpose of transmission capability of a band-allowed noise and counselled that the human auditory machine behave. Although it consisted of a group of round-skip filters having overlie skip band which introduces the period essential bandwidth [17].

1.3. Humans Associated with Speaker Recognition. Human beings can reliably discover acquainted voices and approximately two to three seconds are sufficient to perceive a voice despite the fact that altogether showing decrease in unusual inputs (voice). Surprisingly, if length of the statement becomes expanded, however, performed in reverse (which disturbs the schedule and utterance prompts), the precision reduces extensively. Extensively various overall performances in the historical past challenge recommended, that prompts to input voice popularity range from voice to voice, and that voice styles might also include a fixed set of auditory prompts from where audiences pick out a subgroup to apply in figuring out separate inputs (voices) [18]. Reputation frequently drops drastically while addressers strive to discriminate their inputs (voices). This pondered in appliances, in which precision is reduced while imitating behaves as hoaxers. Speaker reputation is a single location for synthetic observation where gadget overall execution can excel human overall execution, with the usage of quick take a look at observations and N-quantity of audio system, the system precision regularly excels that of humans. This is in particular for a strange audio system in which the education time for human beings to adopt or examine an updated input (voice) is generally too lengthy as contrast to that for systems (machines). Persons overall production in unfavourable situations changed into additionally judged wherein it changed into stating definitely that human listeners adopt the use of numerous cues to affirm the audio system in the presence of acoustic mismatch [19, 20].

1.4. Recognition Process. Speaker reputation of the system (machine) requires 3 ranges; they are (1) removal of functions to denote the addresser details near the articulation signal; (2) designing of addresser qualities; and (3) choice common sense put into effect for identity or validation challenge. The issues related to each of these stages are discussed as follows.

The primary project in an addresser popularity machine or model is to remove attributes capable of indicating the addresser facts existing in the articulation sign. It is regarded that human beings use high-stage capabilities along with the fashion of speech, speech, dialect, and verbal mannerisms (as an example, a particular kind of amusing or use of unique idioms and words) to understand speakers. Intuitively, it is far clear that those functions constitute crucial speaker statistics. Issue arises because of obstacles of the present characteristic extraction strategies. Current speaker recognition systems comply with segmental features inclusive of a pattern for the vocal tracts to symbolize speaker precise information[21]. These capabilities show sizeable versions

across speakers; however, they also display sizable versions now and then for an unmarried speaker. In addition to this, the traits of the recording device and transmission channel are also meditated in these capabilities. As soon as a right set of function vectors is acquired, then within the next section the assignment of addresser identification is to generate a version (archetype) for each addresser [22]. The growth of addresser modelling is referred to as the training section. The schematic diagram of the learning phase is shown in Figure 1.

Characteristic polar coordinates showing the voice quality of the addresser are taken out and worn for constructing the testimonial fashions. Performance of a speaker popularity machine depends in general on the effectiveness of the model in taking pictures and speaker-specific records, and therefore this phase performs a chief function in figuring out the overall performance of a speaker reputation device [23]. The final level inside the improvement of a speaker reputation machine is the decision common sense degree, where a choice to either take delivery of or to refuse the affirm of an addresser is received based totally on the end answer of the complement strategies used. The illustration of choice, good judgment, and testing segment method is shown in Figure 2.

2. Automatic Addresser Identification

Addresser reputation generally consists of various extraordinary approaches of discerning humans depending totally on their voices. The main classes are addresser recognition and addresser verification [24].

In addresser recognition, a speech observation from the hidden addresser is judged and compared with the prototypes of all acknowledged speakers. The hidden addresser is recognized as the speaker whose prototype best suits the input observation. Figure 3 suggests the plain structure of the addresser identity machine [25].

Addresser identity can be a confined set identity or unlock set identity. In a lock set spotting, it is far considered that the check utterance belongs to one in all N enrolled speakers (N decisions). In the case of open set identity, there may be an extra decision to be made to decide whether to take a look at the utterance changed into uttered by way of one of the N enrolled speakers or now not; this is, there are $N+1$ selection degrees [24, 25]. The addresser's speech character is represented by characteristic polar coordinates, which are removed and used while generating the testimony. The success of the algorithms in collecting the speaker-specific defines the overall efficiency of a speaker popularity machine; hence, this stage is important in determining the actual quality of a speaker popularity machine.

Addresser verification objective is to just allow or disallow the affirm of the addresser based totally on the samples of his speech. If the healthy check and reference are above a positive entrance, declare is regular. A high entrance makes it hard for pretenders to be normal by using the machine but with the danger of disallowing the real individual. Conversely, a low entrance ensures that the real person is widely wide-spread continuously, however, with the chance of



FIGURE 1: Training in a speaker recognition system.

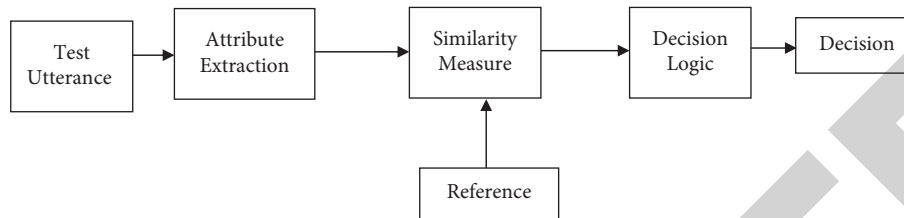


FIGURE 2: Speaker recognition testing system.

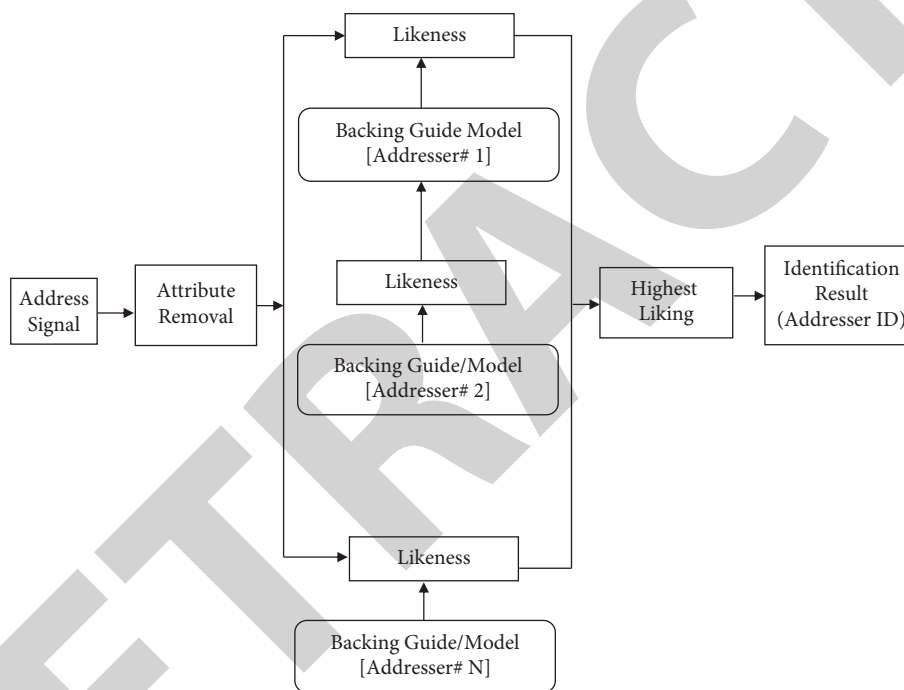


FIGURE 3: Principle shape of a closed set recognition system.

accepting impostors [26]. Figure 4 suggests the simple structure of a speaker verification machine.

2.1. Approaches Adopted for Speaker Recognition. Early research on text-independent speaker reputation is used averaging of function vectors to create a reference template. In the correlation matrices derived from the spectra of an incredibly long length of speech, alerts are used to specify speaker differences. In this research, speech is very useful for remote attack transactions across telephone networks, medical data, and healthcare systems. As a result, a speech-based technique might be effective in detecting changes in voice quality or throat irritation. Persons can be identified in such a setting by gathering a database of people who matches the original database. This allows for a better management system in healthcare. Such techniques will not correctly represent the distribution of characteristic vectors which are

modelled by means of parametric or nonparametric methods. Fashions which expect a chance density feature are termed parametric. In nonparametric modelling, minimal or no assumptions are made regarding the chance distribution of function vectors. In this segment, we in brief evaluate the exclusive version like Gaussian combination version (GMM), Hidden Markov model (HMM), Vector Quantization (VQ), and neuronal community primarily based procedures for speaker reputation. GMM and HMM are parametric fashions [27–29]. VQ and neural network models are dealt as nonparametric models. The Hidden Markov model (HMM) has been the foundation of a number of effective acoustic modelling approaches in speech recognition systems. The model's analysis skill in phenomena, as well as its effectiveness in real voice recognition systems, is the prime cause for its achievement.

Addresser authentication is an unlatched problem. Addresser authentication structures can be additions

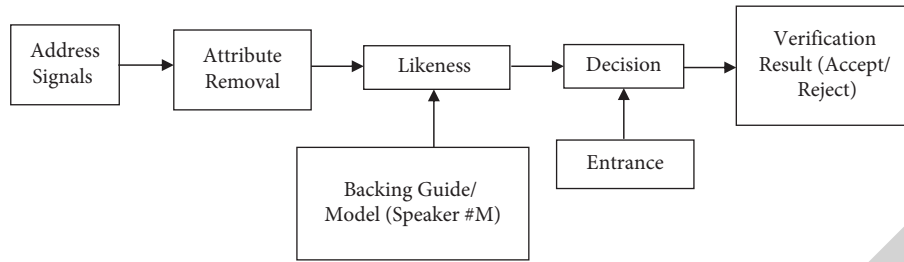


FIGURE 4: Principle structure of an addresser recognition system.

categorized as textual content-based and textual content-impartial process. Inside the text-attached addresser recognition device, the word to be spoken is fixed. In a self-reliant text addresser recognition device, there may be no restrictions on the text to be spoken. Usually, a textual content-based speaker verification machine performs better than a textual content-impartial speaker verification system due to the high degree of control exercised over the speech sign conditions [30]. The text-dependent and text-independent can be described with the help of an example to make it clearer as shown in Table 1. In the initial stage, the agents are verified by some mean of password, and in the second step, the agent will be verified by using the voice giving some numeric value to the system. The most important comparison among HMM, VQ, and neuronal communities is the Hidden Markov model (HMM) dealing with limited factions, while vector quantization and neuronal community are nonparametric models.

2.2. Importance of Feature Extraction. Speech signal carries the following information:

- (i) The intended message
- (ii) Language spoken
- (iii) The speaker's Identity
- (iv) The emotional and physical state of the speaker

Feature removal is the approximation of variables called feature vectors, which loyally describe the pattern (speech pattern) and the problem under deliberation (speaker recognition). The main purpose of the feature removal phase in speaker identity is to withdraw the speaker precise satisfactory [31].

Voices of any two persons vary due to the dissimilarity in the diameter of the spoken space, dissimilarity in the size of the spoken strings, and the fashion in which they are used to generate speech. The short-duration spectrum is chiefly identified by the vocal stretch. Parameters of the characteristics interconnected with the spectrum are acquired from a tiny segment (typically 10–30 msec) of the speech signal, and that is why they are called segmental characteristics. The dissimilarity in the speaking way is due to the way in which the speaker has learned to apply his/her speech production mechanism [32–34]. The temporal dissimilarities of speech traits of various individuals are indicated by these differences. These characteristics are generally eliminated from a comparatively long segment (usually 100–300 msec) of the

speech signal, and thus these characteristics are called suprasegmental characteristics. It is visible that all characteristics taken out should be chosen for speaker acceptance. Therefore, as the feature vector size increases, the computation and storage necessity also increase. For that reason, there is a demand for quality elections. The intention of quality choice is to take a look at the changes in a comparatively low dimensionally quality space that holds the data suitable for the software. At the similar duration, it should also be feasible to accomplish meaningful differentiation using simple actions.

3. Recognition, Spotting, and Validation

Addresser reputation may be divided into subtasks, i.e., addresser spotting and validation. Addresser identity indicates identity for the addressers, wherein addresser validation is a system of accepting or reusing as the recognizer of declared users. Morgan in 2000, Gold, 1997, and Doddington, 1998, used terms when talking about speaker perception and speaker spotting.





Speaker verification is basically a category with two sections; the training session: it is when a model of the user's voice is built up and the real verification is done. The system is thus trained first for a new user's voice that can be performed in many sections, which mean that a spectral analysis is done from which features are extracted to generate a speaker model [35].

Secondly, the user's voice can be verified by comparing the claim's voice with the trained database of user models. On a comparison basis, the system will decide whether the claim's identity is the one modelled by the training material or not. Speaker verification stands on the belief that there is always some quality in every person's speech and that it can be used to verify his/her identity. The quality features in speakers' voices are used both to train a user model and to make up a reference identity for the claimed cause to be used for verification against the user's model [36].

Murray et al., suggested that the first stage of taking out information from speech signal is known as feature extraction. Both have also discussed that speech information is essential for different speech processing tasks [37] like speaker verification is taken as a short spectrum; as per the same author, spectral information is captured during a period of about 20 ms.

On the other side, Furui, 1997, discussed that an effective method to differentiate between users in a system is

TABLE 1: Text-attached and self-reliant text process.

Verification of text-dependent	Verification using speech recognition	Verification with text-prompted	Verification with text-independent
 <p>Phase 1: S1: Using touch-tone keypad kindly enter your account number Caller1: 1234 * # Step2: S1: Please say your password Caller1: INDIAN S1: Thanks!</p>	 <p>S1: Please say your account number. Caller1: 1234 * # S1: Thank you!</p>	 <p>Phase 1: S1: Using touch-tone keypad kindly enter your account number Caller1:1234 * # Step 2: Computer please repeat 43-69 Caller: 43-69 S1: Kindly repeat 82-24 Caller1: 82-24 S1: Thanks!</p>	 <p>Call centre ID: How can I help you Mr. Shrikant? Caller1: I want to transfer the amount 10,123 from my savings account with you to my foreign account in Nepal that I have opened. Agent: Please hold while I am fetching the information and have a look at that. Wait for a while, the computer is matching the identity of calling agent. Agent: Thanks! Process request</p>

fundamental frequencies and different measurements of signals similar to long-term spectrum and short-term spectra and in all powers.

Gold and Morgan, 2000, discussed that text-independent and text-dependent focus on a person's voice is naturally relying on the process of verification. For speech recognition, the correct action of speaking in terms of features is contingent on the classification techniques being used [38].

Doddington, 1998, Furui, 1997, and Schmidt and Gish, 1994, discussed the environmental effect on the speech signal coming from different sources, for example, acoustic noise by Doddington, 1998, and different transmission and recording environments for speech by Doddington 1998, Furui, 1999, and environmental sound or other interfering voices by Schmidt and Gish, 1994.

Doddington, 1998, also stated that differences in the speaker's voice or between different speakers affect the quality of the speaker verification system and suggested that some of the factors that influence and highlight that one justification for speaker voice changes is that the use of speech is a consequence of a person [39]:

- (1) Psychical health and physical condition
- (2) Age factor (human's voice changing as he/she is getting older)
- (3) Speaking rate and level of speech effort
- (4) Intelligence and educational level
- (5) Verification system is experienced

3.1. Model for Speech Recognition. The GMM became first added through Rose and Reynold, 1995, and was an especially used speaker model as it has the potential to model random desire, fashioning possibility density capabilities (pdfs), with the usage of superposition of multivibrate Gaussian. For the diagonal covariance matrix, this is even

genuine when the loss in expressible induced by using the Gaussians being restrained to a round area which may be struggling with the usage of more Gaussians. Using diagonal covariance will help to enhance reputation performance, and much less parameters of the model may be predicted greater comfortably from the limited education facts. The main purpose for deciding on such a model system is that everyone aggregate fashions a hid massive speech sound class [16]. GMM consists of a combination of M Gaussians, in which M absolutely relies nonlinearly on the context and length of the training records supplied through the user [40]. The proposed method is based on speech-dependent health monitoring and management system. In this system, the speech signal may be used to identify a variety of symptoms, such as throat infection, speech pattern alignment, or any voice-related issues.

The classical value of "M" is 32 for characteristics attribute dimensions in the span of 12 to 26. D-dimension is employed for each mixture with mean vector $\square(\rightarrow \tau\mu)$ and diagonal covariance vector $\square(\rightarrow \tau\sigma^2)$, measured by the agent "w" so that the total mass is 1 and modelling shape is scattering. The log-likelihood \ln_{GMM} for D-dimensional feature vector $X = \{\square(\rightarrow \tau y_{-t}) | 1 \leq t \leq T \wedge \square(\rightarrow \tau y_{-t}) \in S D\}$ given by the probability $P(X|\lambda_{GMM})$ is represented by its parameter. Gaussian chance distribution and capabilities belonging to the clusters may be satisfactory and acceptable through their opportunity values. The best problem includes green category of feature vectors [41, 42]. Figure 5 shows the Gaussian Mixture Model with its feature space and corresponding 2-dimension.

GMM used for speaker identification is influenced by two facts as follows Gaussian individual classes are explained to represent the acoustic class, and its set vocal tract information represents the acoustic classes.

GMM density gives a uniform approximation to the distribution of feature vectors in a multidimensional region.

3.2. Functioning of Health Unit. Health centre generates an authenticated report on the basis of his/her voice quality or speech analysis and generated throat quality analysis parameters using a feature extraction method which may or may not be the same as per the original speech quality as saved in database.

Monitoring will be done on the basis of speech data collected per data, and its testing will be done n number of times using training, training, and sampling techniques which may hardly take a few milliseconds. Smart management is the need of an hour where the situation changes suddenly and monitoring is required at every point of time. Pandemic situations may occur in the future, and such challenges require regular monitoring and real-time management systems to deal with such scenarios. COVID-19 [43] affected the entire nation, and India is the only country where the recovery rate is quite high compared with other developed nations in spite of low medical facilities, and everybody is not capable to have a medical test in terms of blood samples and get the result fast. However, many people face difficulties in rural areas where the situation is the worst, and the testing rate is very slow and needs more attention. In such condition, voice can be used to collect the sample of each person by taking the sample just by making a call and matching process by extracting the feature, and sampling may be useful. The only thing required is to make a database of villagers in terms of speech and store in the cloud and can be accessed anywhere and verification can be done anytime. The data collection centre can gather the information in terms of speech, and that sample will be analyzed using its original sample and trained by the system and matched with the saved database sample by extracting its original frequency spectrum variation, and it will be verified, and finally the control room of the hospital will generate a report by analyzing various parameters of voice and sent to the person whose voice really gets disturbed and might not be affected by some viral disease or voice disorder as shown in Figure 6.

4. Preparation of Database

The bottom has been organized for fifty audio systems collecting the samples (spoken) of males and females of various group institutions, with the usage of gold wave 5.5 and cool tool program using microphones. The directory collected outlying Hindi numbers along with 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. Every addressers speak 0–9 digits (in Hindi), each number for 10 instances, then changed into a note in a noise-unfastened condition, and kept as a wave sound file (.wav) recording the desired folder. Table 2 shows the full database of 50 different speaker users including both males and females of different groups. The parameters of data/sample collection are shown in Table 3.

4.1. Procedure for Data Collection. The spoken of 50 distinctive audio systems has been put down the usage of taxon speakers. Every spectrum of Hindi text from 0 (SHUNYA) to

9 (NAU) was spoken 10 times, and every text has been put down inside the shape of the scale by reducing each scale (frame), and via choosing “store alternative” from the record menu, each scale is stored using .wav file in a suitable folder. The spectrum of Hindi digits is recorded using cooling software which reflects different energies.

Steps involved are as follows:

- (1) First, choose the application type; it will give the blank space in Graphical User Interface type (GUI).
- (2) Go to the file menu and select new.
- (3) This will give the option to choose sampling rate, channels, and resolution.
- (4) Then, press OK by selecting the proper sampling rate, channel, and resolution; it will give a blank space.
- (5) Now, go the record option indicated in the left bottom corner and press the record button to record the Hindi spoken digits.
- (6) Press the stop button to examine the frame collected.

Similarly, the entire frame from “SHUNYA to NAU” has been collected by repeating the abovementioned steps. Figures 8–10 show the spectrum of the spoken digits for different variations.

5. Proposed Model and Results

Suppose “S” is the target person’s speech represented by $S = \{s_1, s_2, \dots, s_T\}$ and “V” is the unobservable part of the speech dependent represented by $V = \{v_1, v_2, \dots, v_T\}$, then for statistically independent S and V, probability $P(x_t, j, k)$ can be calculated as follows:

$$P(x_t, j, k) \int_{-\infty}^{\infty} M\left(S; \sum_{s,j} \mu_{s,j}\right) M\left(x_t - S; \sum_{v,k} \mu_{v,k}\right), \quad (1)$$

where $\sum_{s,j}$ and $\sum_{v,k}$ are weight covariance matrix and $\mu_{s,j}$ and $\mu_{v,k}$ are mean vectors, and M denotes the Gaussian density matrix.

Speaker-specific model λ_s in the maximum likelihood is given by

$$\begin{aligned} \lambda_s^* &= \max_{\arg} P(XI\lambda_x, \lambda_v), \\ &= \max_{\arg} \prod_{j=1}^J \prod_{k=1}^K \prod_{t=1}^T w_{z,j} w_{z,k} P(x_t I j, k). \end{aligned} \quad (2)$$

So, given λ_v in (2) can be solved using the expectation maximization algorithm, starts with the initial model λ_s and estimate new one λ_s^1 . The proposed model for verification is shown in Figure 7.

Here, $P(XI\lambda_v, \lambda_s^1) \geq P(XI\lambda_v, \lambda_s)$, and when testing is found to be unknown speech, then the system determines whether or not it is produced by the claimed speakers using the following equation:

$$\log P(YI\lambda_v, \lambda_s) - \log(YI\lambda_v). \quad (3)$$

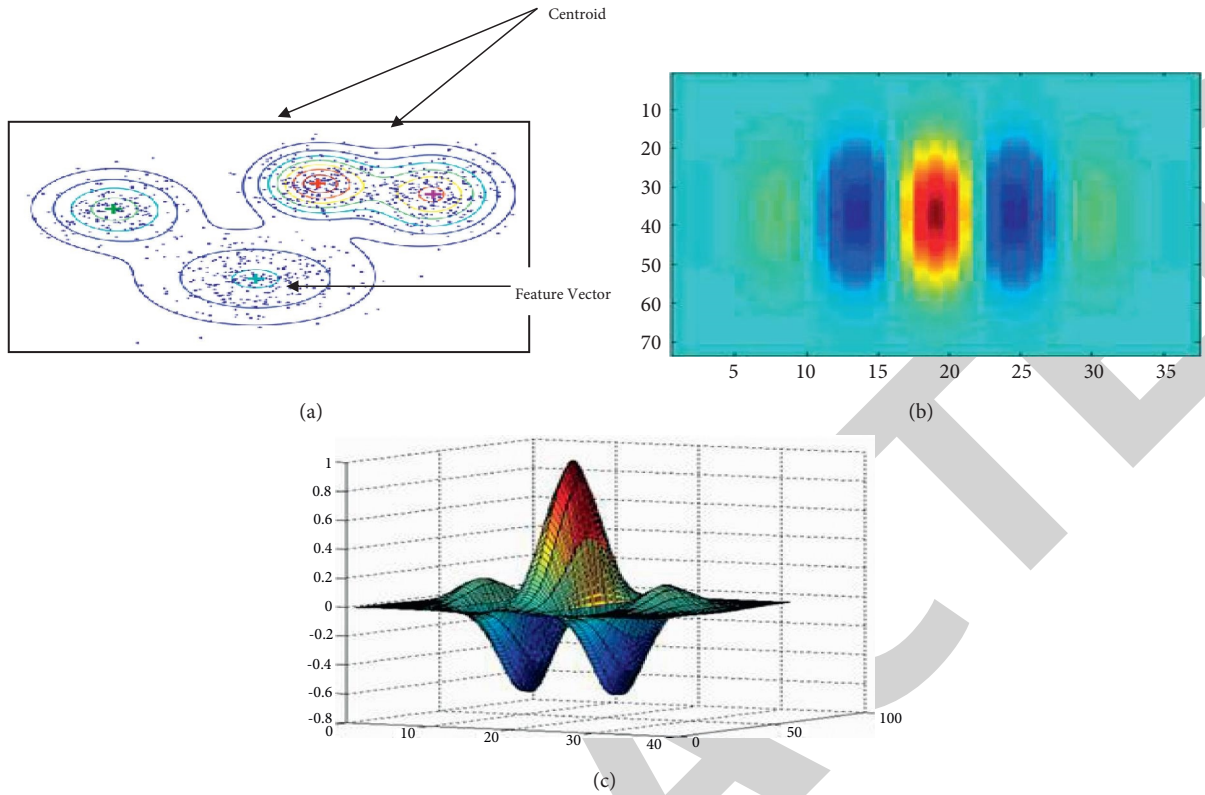


FIGURE 5: Gaussian mixture model with its feature space and corresponding 2-dimension.

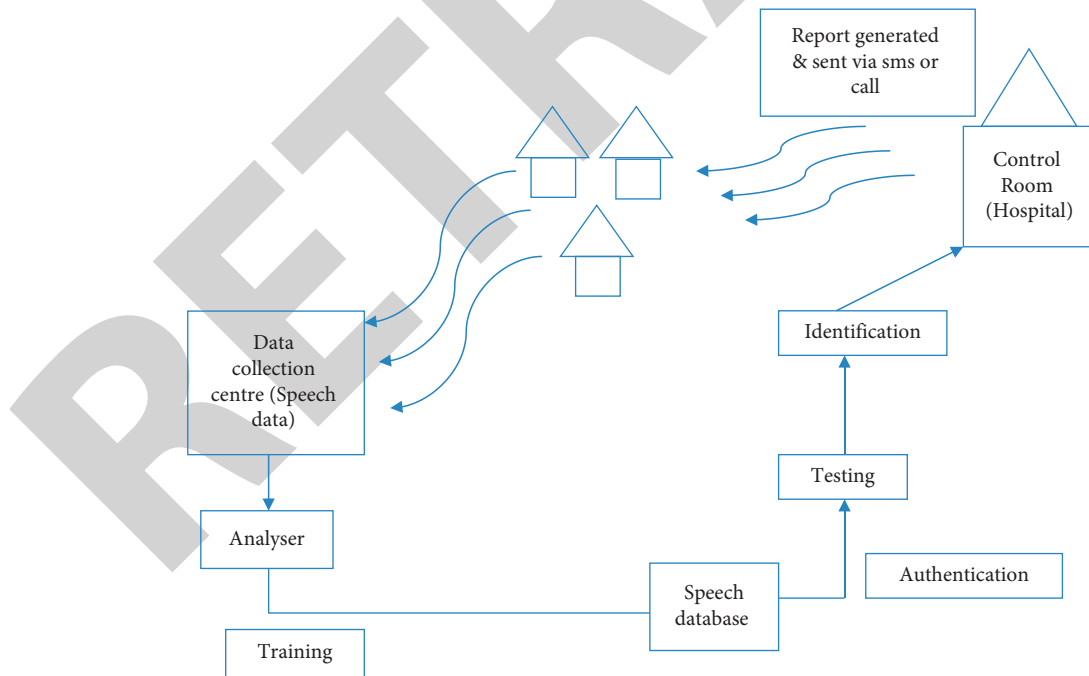


FIGURE 6: Model for monitoring and management (assembled with the proposed model, Figure 7).

It is satisfied when it is less than θ , and unsatisfied when greater than or equal to θ . Here, Y is the utterance cestrum feature and θ is the threshold level.

The efficiency of different coding techniques is shown in Table 4 represented graphically in Figure 11, where LPC means linear predictive coding, MFCC means Mel-

TABLE 2: Database of 50 speakers.

SL. No.	Name
0	Ambika_25F
1	Aditi_28F
2	Anil_24M
3	Anju_23F
4	Ankitraj_24M
5	Chandu_26M
6	Choti_18F
7	Deepak_26M
8	Deepmala_23F
9	Devmal_24M
10	Gauri_22F
11	Gopal_24M
12	Goutmi_23F
13	Hitan_24M
14	Isha_18F
15	Indira_24F
16	Kamal_25M
17	Kaushal_25M
18	Kiraan_18F
19	Lovley_26F
20	Mantu_24M
21	Navin_26M
22	Neetu_25F
23	Nitika_24F
24	Pankaj_27M
25	Parvathi_22F
26	Puja_24F
27	Prarthana_19F
28	Ramol_24F
29	Ravikant_24M
30	Ribha_21F
31	Sakshi_24F
32	Sanhita_23F
33	Sanjay_24M
34	Satyas_25M
35	Shaleen_23F
36	Shallu_23F
37	Shantanu_23M
38	Shipra_26F
39	Shital_24M
40	Sourav_25M
41	Sujjal_22M
42	Swati_23F
43	Sweta_22F
44	Toral_25F
45	Usharani_26F
46	Uttkal_25M
47	Vandana_19F
48	Vibha_19F
49	Vikash_26M

TABLE 3: Parameters used for collecting data.

Parameters	Values
Total number of speakers	50
Single channel	16 KHz
Sampling rate	16 bit (for less quantization error)
Language	Hindi (phonetics is rich)
Repetition rate of each speaker	10 digits/minute

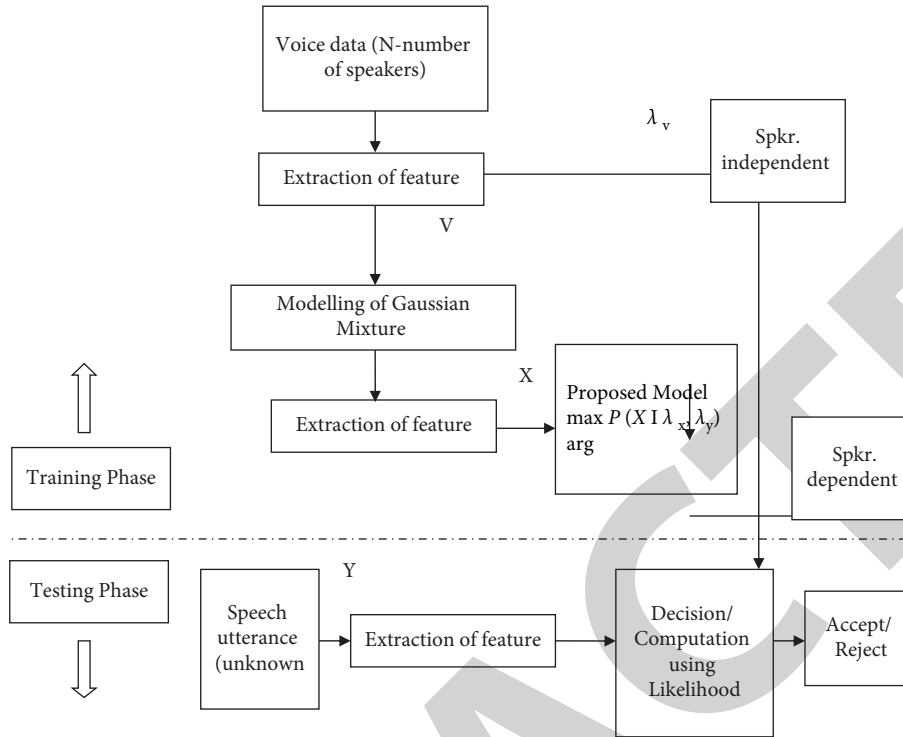


FIGURE 7: Proposed model for verification and authentication.

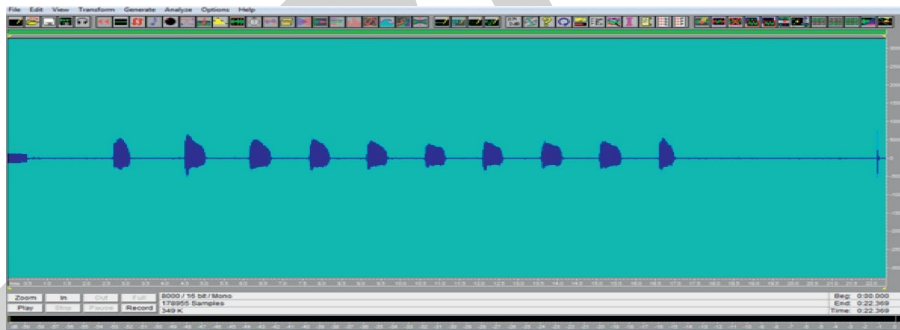


FIGURE 8: Spectrum of spoken digits.

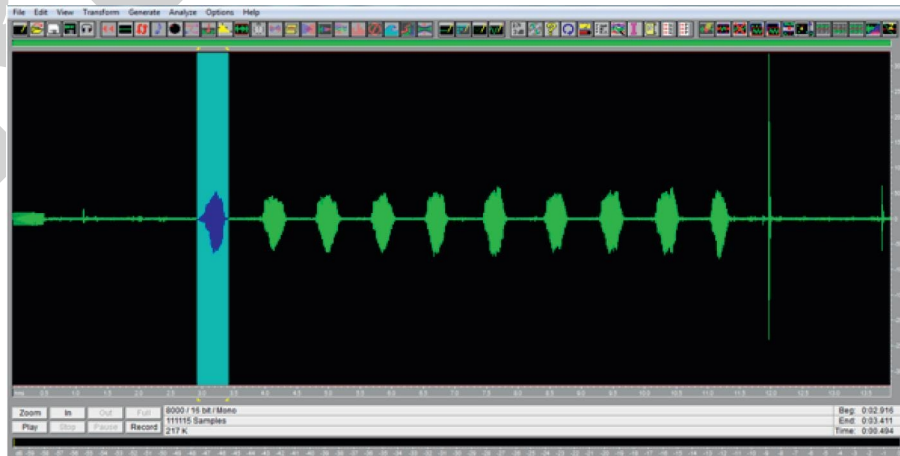


FIGURE 9: Sampling of spoken digits.

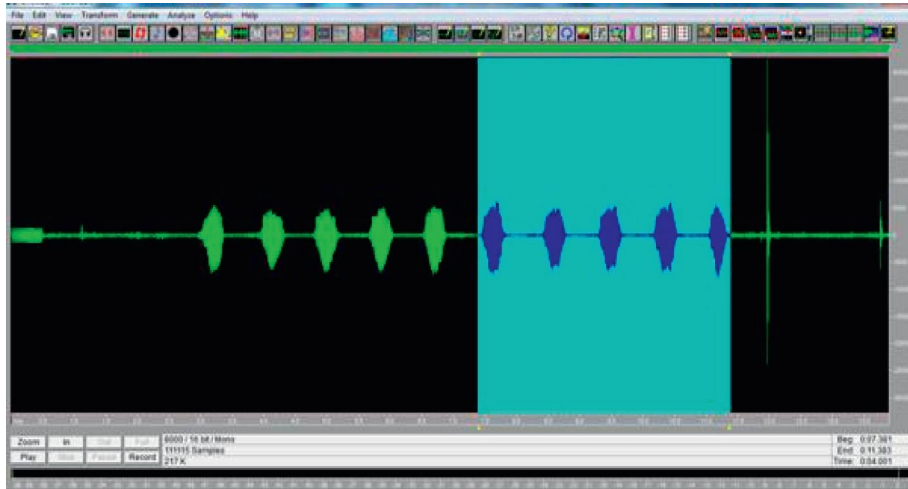


FIGURE 10: Spectrum selection for complete spoken digits.

TABLE 4: Efficiency rate in % of individual feature extraction.

Hindi dialects	LPC (%)	MFCC (%)	LPCC (%)
0 (SHUNYA)	89.99	91.99	90.89
1_EK	88.12	92.23	89.98
2_DO	90.10	89.97	88.27
3_TEEN	87.26	89.25	88.12
4_CHAR	90.12	91.25	90.22
5_PANCH	92.29	90.23	91.67
6_CHEH	90.54	93.44	92.81
7_SAAAT	89.11	90.23	92.64
8_AATH	90.78	94.10	90.21
9_NAU	91.34	93.43	87.66

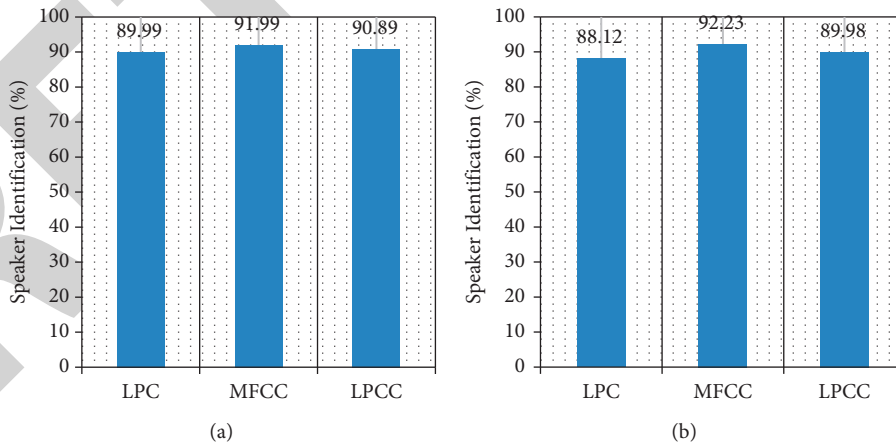


FIGURE 11: Efficiency of “SHUNYA” and “EK.”

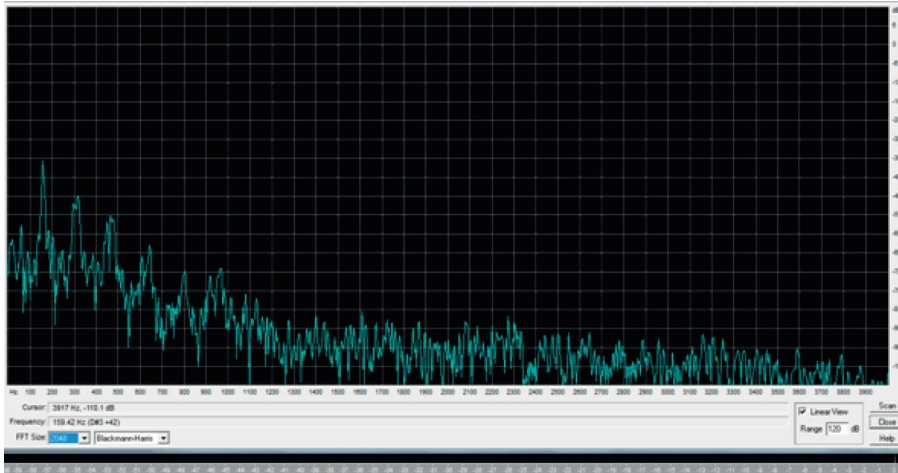


FIGURE 12: Variation in the frequency spectrum of spoken digit “SHUNYA.”



FIGURE 13: Variation in frequency spectrum of spoken digit “EK.”

frequency cepstral coefficient, and LPCC means linear prediction cepstral coefficient.

The linear predictive coding (LPC) is a technique for compressing the frequency spectrum of a digitized voice signal utilizing data from a linearly statistical model. It is broadly applied in auditory signal processing and speech analysis. The LPCC is based on linear prediction that provides the people articulatory system.

Moreover, the variation in the frequency spectrum is shown in Figures 12 and 13.

6. Conclusion

Using this speech-dependent health monitoring and management system, speech data can be collected from remote locations and can be accessed. As per the study, it is found that COVID-19 virus first affects our throat, and then it enters into the pumping area (heart). Therefore, in such a condition, a sample in the form of voice is collected from particular areas where COVID-19 patient is large in number and can be analyzed using such a model which helps to first

detect his voice quality and matched with the original voice as saved in database. The proposed model provides a better management system without touching and maintains appropriate distance information that can be collected and processed for medical treatment. The future scope of this research is that it will reduce the chances to come into contact with a viral effected people in future. Any feature extraction approach can be imposed on that proposed model whose efficiency is found to be better to sample and analyze the speech signal to achieve good authentication. As per our analysis, MFCC found a good efficiency rate, i.e., 93.43% for the spoken digit, which may suit for further analysis. Blood sampling can be avoided using this method, and of course social distancing can also be maintained by remotely accessing the voice sample and proven to be a much better option for the society. Smart solution for the upcoming generation is required in terms of better facility, less pain, economical, and fast results. Therefore, such health care and management system is proved beneficial, and its implementation changes the face of health units in rural as well as urban areas.

Data Availability

The data are used to support this study are available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work is self-funded.

References

- [1] D. G. Childers, R. V. Cox, R. Demori et al., "The past, present & future of speech processing," *IEEE Processing Magazine*, vol. 15, pp. 24–48, 1998.
- [2] D. Lancker, J. Kreiman, and K. Emmorey, "Familiar voice recognition: pattern & parameters- recognition of backward voices," *Phonetics*, vol. 13, no. 1, pp. 19–38, 1985.
- [3] A. R. Reich and J. E. Duke, "Effects of selected vocal disguises upon speaker identification by listening," *Journal of the Acoustical Society of America*, vol. 66, no. 4, pp. 1023–1028, 1979.
- [4] M. Sigmund, "Speaker recognition identifying people by their voices," Habilitation Thesis, Brno University of Technology, Institute of radio electronics, Brno, Czechia, 2000.
- [5] G. R. Doddington, "Speaker recognition- identifying people by their voices," *Proceedings of the IEEE*, vol. 73, no. 11, pp. 1651–1664, 1985.
- [6] J. D. Markel and B. Oshika, "Long-term feature averaging for speaker recognition," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 25, no. 4, pp. 330–337, 1977.
- [7] W. A. Hargreaves and J. A. Starkweather, "Recognition of speaker identity," *Language and Speech*, vol. 6, no. 2, pp. 63–67, 1963.
- [8] K. P. Li and G. W. Hughes, "Talker differences as they appear in correlation matrices of continuous speech spectra," *Journal of the Acoustical Society of America*, vol. 55, no. 4, pp. 833–837, 1974.
- [9] D. O'Shanughnessy, "Speaker recognition," *IEEE ASSP Magazine*, vol. 3, pp. 4–17, 1986.
- [10] G. Fant, A. Kruckenberg, and L. Nord, "Prosodic and segmental speaker variations," *Speech Communication*, vol. 10, no. 5-6, pp. 521–531, 1991.
- [11] N. Malayath, "Voice coversion," MS Dissertation, Indian Institute of Technology, Department of Computer Science & Engineering, Madras, India, 1995.
- [12] J. P. Campbell, "Speaker recognition: a tutorial," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1437–1462, 1997.
- [13] A. Sutherland and M. L. Jack, *Speaker Verification in Aspects of Speech Technology*, M. A. Jack and J. Laver, Eds., Edinburgh University Press, Edinburgh, Scotland, 1988.
- [14] F. Nolan, *The Phonetic Bases of Speaker Recognition*, Cambridge University Press, England, UK, 1983.
- [15] M. Pandit and J. Kittler, "Feature selection for a dtw-based speaker verification system," vol. 2, pp. 769–772, in *Proceedings of the IEEE International Conference Acoust., Speech & Signal Processing*, vol. 2, pp. 769–772, IEEE, Seattle, WA, USA, May 1998.
- [16] S. Rydin, *Text Dependent & Text Independent Speaker Verification Systems: Technology & Applications*, Center for Speech Technolgy KTH, Stockholm, Sweden, 2001.
- [17] S. Sharma, P. Rattan, A. Sharma, and M. Shabaz, "Voice activity detection using optimal window overlapping especially over health-care infrastructure," *World Journal of Engineering*, vol. 19, no. 1, pp. 118–123, 2022.
- [18] Y. Ittana, "An intelligent text speaker identification using vq-gmm model based multiple classifier system," Master Science Thesis, Department of Information Engineering, Brescia, Italy, 2010.
- [19] L. R. Rabiner and H. H. Juang, *Fundamental of Speech Recognition*, Prentice-Hall, New Jersey, NJ, USA, 1993.
- [20] B. S. Atal, "Automatic speaker recognition based on pitch contours," *Journal of the Acoustical Society of America*, vol. 52, no. 6B, pp. 1687–1697, 1972.
- [21] A. Tiwari, V. Dhiman, M. A. M. Iesa, H. Alsarhan, A. Mehbodniya, and M. Shabaz, "Patient behavioral analysis with smart healthcare and IoT," *Behavioural Neurology*, vol. 2021, Article ID 4028761, 9 pages, 2021.
- [22] S. Furui, "Cepstral analysis technique for automatic speaker verification," *IEEE Transactions on Acoustics, Speech, & Signal Processing*, vol. 29, pp. 254–272, 1984.
- [23] C. M. Bishop, *Neural Network for Pattern Recognition*, Oxford University Press, Oxford, UK, 1995.
- [24] B. Yegnanarayana, *Artificial Neural Networks*, Prentice Hall of India Pvt. Limited, Hoboken, NJ, USA, 1999.
- [25] S. Upadhyay, S. K. Sharma, and A. Upadhyay, "Performance analysis of Hindi of Hindi voice for speaker recognition and verification using different feature extraction," *International Journal of Electronics Engineering Research*, vol. 9, no. 2, pp. 241–254, 2017.
- [26] A. Upadhyay, S. Kumar Sharma, and A. Upadhyay, "Speaker Identification and verification using different Model for text dependent," *Indian Journal of Science and Technology*, vol. 10, no. 47, pp. 1–6, 2017.
- [27] E. A. Patrick, *Fundamental of Pattern Recognition*, Prentice Hall Inc, New Jersey, NJ, USA, 1972.
- [28] G. T. Toussaint, "Bibliography on estimation of misclassification," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 472–479, 1974.
- [29] G. S. Sebestyen, *Decision Making-Process in Pattern Recognition*, Macmillan company, New York, NY, USA, 1972.
- [30] J. Oglesby, "What's n a number? moving beyond the equal error rate," *Speech Communication*, vol. 17, no. 1-2, pp. 193–208, 1995.
- [31] A. Kakti, S. Kumar, N. K. John, V. V. Ratna, S. Afzal, and A. Gupta, "Impact of patients approach towards healthcare costs on their perception towards health: an empirical study," *Tobacco Regulatory Science*, vol. 7, no. 6-1, pp. 7380–7390, 2021.
- [32] S. Upadhyay, S. K. Sharma, S. Vijay, and A. Upadhyay, "Analysis of feature extraction behavior for speaker identification and verification in adverse acoustic condition," *International Journal of Control Theory and Applications*, vol. 10, no. 9, pp. 25–31, 2017.
- [33] H. Gish, "Robust discrimination in automatic speaker identification," vol. 1, pp. 289–292, in *Proceedings of the IEEE International Conference Acoust., Speech, & Signal Processing*, vol. 1, pp. 289–292, IEEE, Albuquerque, NM, USA, April 1990.
- [34] V. Raman and J. Naik, "Noise reduction for speech recognition & speaker verification in mobile telephony," in *Proceedings of the International Conference Spoken Language Processing*, pp. 1839–1842, West Bengal, India, 1994.
- [35] H. Z. Almarzouki, H. Alsulami, A. Rizwan, M. S. Basingab, H. Bukhari, and M. Shabaz, "An internet of medical things-

Retraction

Retracted: Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] J. Srikanth, A. Damodaram, Y. Teekaraman, R. Kuppusamy, and A. R. Thelkar, "Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8898100, 11 pages, 2022.

Research Article

Sentiment Analysis on COVID-19 Twitter Data Streams Using Deep Belief Neural Networks

Jatla Srikanth,¹ Avula Damodaram,² Yuvaraja Teekaraman ,³ Ramya Kuppusamy ,⁴ and Amruth Ramesh Thelkar ⁵

¹Department of Computer Science and Engineering, Aurora's Technological and Research Institute, Hyderabad 500098, TS, India

²School of Information Technology (SIT), JNTUH, Hyderabad 500085, TS, India

³Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 3JD, UK

⁴Department of Electrical and Electronics Engineering, Sri Sairam College of Engineering, Bangalore 562106, India

⁵Faculty of Electrical & Computer Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

Correspondence should be addressed to Yuvaraja Teekaraman; yuvarajastr@ieee.org and Amruth Ramesh Thelkar; amruth.rt@gmail.com

Received 23 January 2022; Accepted 16 March 2022; Published 6 May 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Jatla Srikanth 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.

Social media is Internet-based by design, allowing people to share content quickly via electronic means. People can openly express their thoughts on social media sites such as Twitter, which can then be shared with other people. During the recent COVID-19 outbreak, public opinion analytics provided useful information for determining the best public health response. At the same time, the dissemination of misinformation, aided by social media and other digital platforms, has proven to be a greater threat to global public health than the virus itself, as the COVID-19 pandemic has shown. The public's feelings on social distancing can be discovered by analysing articulated messages from Twitter. The automated method of recognizing and classifying subjective information in text data is known as sentiment analysis. In this research work, we have proposed to use a combination of preprocessing approaches such as tokenization, filtering, stemming, and building N -gram models. Deep belief neural network (DBN) with pseudo labelling is used to classify the tweets. Top layers of the base classifiers are boosted in the pseudo labelling strategy, whereas lower levels of the base classifiers share weights for feature extraction. By introducing the pseudo boost mechanism, our suggested technique preserves the same time complexity as a DBN while achieving fast convergence to optimality. The pseudo labelling improves the performance of the classification. It extracts the keywords from the tweets with high precision. The results reveal that using the DBN classifier in conjunction with the bigram in the N -gram model outperformed other models by 90.3 percent. The proposed approach can also aid medical professionals and decision-makers in determining the best course of action for each location based on their views regarding the pandemic.

1. Introduction

The new coronavirus illness (COVID-19) is an ongoing pandemic that has sparked widespread concern around the world. Spreading misleading information on social media platforms like Twitter, on the other hand, is exacerbating the disease's concern. People can stay safe, informed, and connected with the help of technological improvements and social media. The same instruments, on the other hand, enable and intensify the current infodemic, which continues

to undermine the global response and risk pandemic-control measures. Despite the fact that young people are at a lower risk of serious sickness from COVID-19, they are an important group in the context of the pandemic and share in the communal responsibility of helping us stop transmission. They are also the most active online, using different digital channels per day (such as Twitter, Facebook, TikTok, WeChat, and Instagram).

Social media has swiftly grown into a crucial communication tool for the production, transmission, and

consumption of information since the start of the COVID-19 epidemic. Several studies have already used social media data to aid in the detection and identification of infectious disease epidemics, as well as the interpretation of public attitudes, behaviours, and perceptions. As social media content is provided by users, it can be subjective or erroneous, and it commonly contains misinformation and conspiracy theories [1]. Sentiment analysis, also known as opinion mining, is the method of classifying emotions in subjective data using machine learning (ML) and natural language processing (NLP).

By analysing and sharing information from peer-reviewed, published research, policymakers and public health organizations may design efforts for accurate and timely knowledge translation to the general population. Individual profiles have evolved, reinforced by social networks, which have had a similar influence on the more specialized communication medium, in addition to traditional media as the main drivers of social communication in crisis situations. The findings suggest to set a new communications paradigm that creates a new area for agents whose material has a level of engagement equivalent to, if not greater than, that of digital health media [2].

The current COVID-19 problem is resulting in a socially advanced condition that is unprecedented for health communities. We live in a highly globalized society, where mobility, unrestricted travel between countries, and the evolution and application of Information and Communication Technologies (ICTs) are all highly developed. Similarly, there is a continual stream of bogus news and criticism of movements on social media sites as a result of the viral pandemic. It stifles public health professionals' communication and incites broad public indignation [3].

A convolutional neural network (CNN or ConvNet) is a type of deep neural network used to interpret visual imagery in deep learning. Artificial neural networks have ushered in new ideas and procedures in machine learning in a variety of fields during the last two decades. Many existing processes have been replaced by them. ConvNets are made up of filter layers (known as convolutional layers) and aggregation layers (known as pooling layers), which are alternately repeated. One or more fully connected layers follow them.

Text classification aims to categorize text documents into one or more predetermined categories automatically. Understanding audience sentiment (happy, sad, and furious) from social media, detecting spam and nonspam emails, auto tagging of customer enquiries, categorization of news items into predetermined subjects, and so on are some instances of text classification [4].

Text encoders are generally pretrained to handle text as a sequence of tokens corresponding to small text units, such as English word parts. In many NLP jobs, where a powerful encoder is required to model more contextual information, a high quality of text representation plays a key part in achieving good performance. Text mining and information retrieval need the selection of text feature items, which is a fundamental and significant task. Deep learning differs from traditional methods in that it automatically learns features from huge data rather than adopting handmade

characteristics, which rely heavily on past knowledge of designers and make it nearly impossible to take advantage of massive data [5]. Text classification framework using a supervised learning model is shown in Figure 1.

The deep counterpart of an auto encoder is a stacked auto encoder, which may be produced simply by stacking layers. The learnt representation of the previous layer is the input for each layer, and it learns a more compact representation of the existing learned representation. We propose to use feature extraction from the twitter data streams using n -gram stack encoder followed by setting the rules and training the neural network for classifying streaming data. We will concentrate on the data stream classification task, where the parameters of a classification model may change over time, requiring the model to adapt.

2. Related Work

The COVID-19 pandemic has resulted in a massive loss of human life around the world and poses a significant threat to global health, food chain, and the workplace. The COVID-19 epidemic has claimed many lives and poses a serious danger to world health, the food supply, and the workplace. The response differences between social media and financial markets as a result of the severe viral spread's after-effects revealed the dynamics of COVID-19, including mortality, contagion variables, time of country virus, and early fatalities. During the lockdown, social media channels were critical in spreading information about the pandemic around the world, as people used the platforms to express their emotions. In light of this severe situation, it is necessary to examine people's reactions on Twitter, taking into consideration popular words that are directly or indirectly related to the epidemic [6].

It would be practically impossible for a human to read and comprehend everything that has been said about COVID-19 vaccinations on Twitter. Fortunately, using textual feature extraction, sentiment analysis, and word cloud visualizations, we can look into an incredibly complicated and wide-ranging dialogue using natural language processing (NLP) approaches. Bhatia et al. made a study on sentiment analysis of COVID-19 tweets with the help of deep learning classifiers [7]. People were not guided by tweets about the COVID-19 epidemic, according to this research. The findings show that neither WordCloud nor the frequency of terms in tweets contains any useful words. A proposed deep learning classifier model with an accuracy of up to 81 percent validates the claims. The authors claim that a fuzzy rule based on Gaussian membership correctly recognizes sentiments from tweets.

Amit et al. have proposed a sentiment analysis on the impact of coronavirus in social life using the BERT model. They claim that, because Twitter has become one of the most prominent social media sites, the authors conducted sentiment analysis on tweets using the BERT model, based on people's excitement and opinions to better understand their mental state [8]. The authors conducted a sentiment analysis on two data sets in this paper: one data set contains tweets from people all over the world, while the other data set

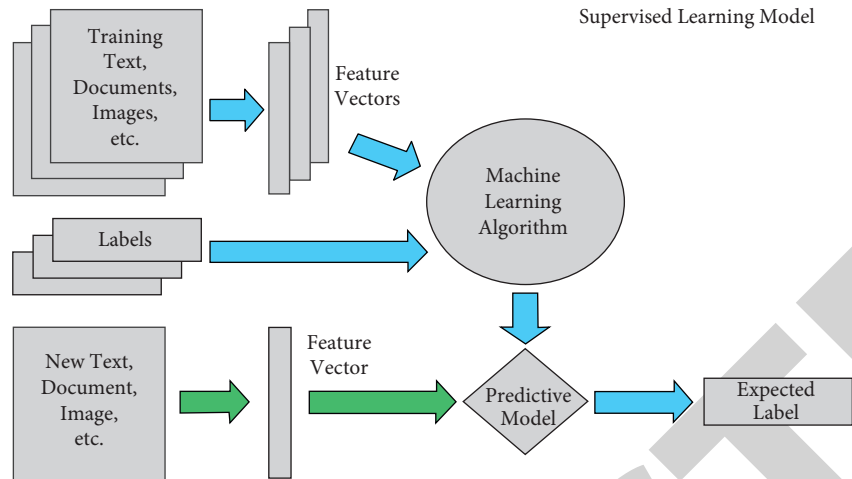


FIGURE 1: A typical framework with supervised learning model for text classification.

contains tweets from people in India. They used the GitHub repository to validate the emotion classification's accuracy. The validation accuracy is 94 percent, according to the experimental data.

With the growth in COVID-19 cases, a strange position of pressure was placed on each country to manage the problem and make the best use of available resources. People experienced panic, anxiety, and depression as the number of positive cases increased rapidly around the world. The impact of this lethal disease was discovered to be directly proportionate to the population's physical and mental health. Social media has been the most prominent tool that has disrupted human life during this period. The tweets about COVID-19, whether they were about a large number of positive cases or deaths, caused a surge of panic and concern among people all across the world. Researchers and data scientists will be able to access the data for academic and research purposes. Many data points related to real-life occurrences, such as COVID-19, can be found in social media data. Harleen Kaur and others have used the R programming language to analyse the Twitter data. The authors gathered data from Twitter using hashtag terms such as COVID-19, coronavirus, deaths, new case, and recovered cases. They used an algorithm called hybrid heterogeneous support vector machine (H-SVM) for sentiment classification and categorised the scores as positive, negative, or neutral in this study [9]. They also compared the suggested algorithm's performance to that of the recurrent neural network (RNN) and support vector machine on specific criteria such as precision, recall, *F1* score, and accuracy (SVM).

In a very short period of time, the social networking site Twitter saw an extraordinary rise in tweets pertaining to the novel Coronavirus. Gurumurthy et al. conducted a global sentiment analysis of coronavirus tweets, demonstrating how public opinion in various nations has changed over time. Furthermore, tweets relating to work from home (WFH) and online learning were scraped and the change in sentiment over time was studied to identify the impact of coronavirus on daily elements of life. The authors also tested the accuracy of several machine learning models for

sentiment categorization, such as long short term memory and artificial neural networks [10].

Kuhn et al. have discussed cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic in their work. The difficulty in analysing the data source stems from the fact that social media communications are exceedingly noisy and idiosyncratic, and the volume of incoming data is far too enormous to manually evaluate. As a result, the authors found that automated methods for extracting useful information are required. The authors examine the emotion of Twitter messages collected during the first months of the COVID-19 outbreak in Europe [11]. This is done utilizing multilingual sentence embedding and a neural network for sentiment analysis. The researchers categorize the findings by place of origin and compare their chronological progression to events in those countries. This allows them to investigate how the situation affects people's moods. They discovered, for example, that lockdown announcements are associated with a drop in mood in almost all of the countries studied, which recovers quickly.

Anupam Mondal and others have presented a new approach to classify COVID-19 tweets using machine learning algorithms. The documentation goes over the methods for preprocessing tweets, feature extraction, and developing machine learning models in detail. When the organizers evaluated both of the created learning models, they received *F1* ratings of 0.93 and 0.92, respectively [12]. They passed the words, POS tags, and TF-IDF values through their default TensorFlow embedding layer as separate inputs, then concatenated the outputs to train the DL model. The output of the concatenation layer was then sent to a dense layer that mapped the tensors to their labels via two layers of bidirectional long-short term memory.

Ching have proposed a *n*-Gram statistics for natural language understanding and text processing [13]. For applications in natural language understanding and text processing, *n*-gram statistics and other features of the English language were determined. They were calculated using 1 million word samples from a well-known corpus.

Similar features were discovered in three more corpuses' most frequently used 1000 words. The n -gram positional distributions produced in this investigation are discussed. There are statistical studies on word length and n -gram frequency trends versus vocabulary. A collection of n -gram statistics obtained by different researchers is evaluated and compared in addition to a study of n -gram statistics found in the literature.

With the recent growth of data in motion, there is a rising academic interest in analysing streaming data, which has resulted in the publication of numerous new papers on data stream analytics. But, investigating the capabilities of traditional recurrent neural networks in the context of streaming data classification is still a relatively unexplored area. Monidipa Das et al. have presented a model called FERNN, a unique RNN variant with single-pass learning capability and self-evolution attribute [14]. FERNN is well suited to working with streaming data due to its online learning abilities, while its self-organizing trait makes it adaptable to a fast changing environment. In the hidden layer, FERNN employs hyperplane activation, which not only reduces network parameters but also causes the model to act as a teacher forcing mechanism by default, automatically addressing the vanishing/exploding gradient issues that can arise with traditional RNN learning based on back-propagation-through-time policy. As it is not limited by the normal distribution assumption for streaming data, FERNN is more flexible than the majority of the existing autonomous learning models. Under the test-then-train approach, the efficacy of FERNN is assessed in terms of classifying six publicly available data streams. The experimental findings suggest that FERNN can achieve state-of-the-art classification accuracy with a relatively low computation cost.

A methodology is proposed to analyse COVID-19 from Arabic twitter tweets. Naïve Bayes classification is used for basic classification. Later ensemble methods are used for analysis. Ensemble methods with SMOTE is performed better than basic classifiers [15]. The polarization is performed using machine learning algorithms, deep learning, and TextBlob. The deep learning approach such as Bi-LSTM performed when compared to other methods [16]. People emotions are analysed and classified as positive, negative, and neutral. The machine learning algorithms, deep learning algorithms, and ensemble methods are used [17].

In this work, we propose to use a feature extraction using n -gram stack encoder method first. Data collection and preprocessing are the first steps in our research work. The preprocessing involves different phases including data cleaning, finding polarity, finding sentiments, and combining the data sets. To get the sentiments, we then use different methods including parts of speech tagging, lemmatization, stemming, stop words removal, and finding the keywords. Once the features are extracted, we then use the convolutional neural network for classifying the streaming data. In machine learning, we utilize the deep belief neural network, which is a generative graphical model, or a form of deep neural network, made up of several layers of latent variables with connections between levels but not between units within each layer. Our results are found to be

promising and better as compared to the existing methods in the literature that we have discussed in this section.

3. Data Collection and Preprocessing

Businesses may interact with customers on Twitter in a more personal way. However, because Twitter has so much data, it can be difficult for marketers to pick which tweets to respond to first. As a result, sentiment analysis has become an important tool in social media marketing campaigns. Sentiment analysis is a technique that automatically tracks sentiments in social media interactions. Data collection is the first step in this research problem of sentimental analysis of the COVID-19 twitter data [18]. To follow certain terms and accounts that were trending at the time of data collection, we used Twitter's streaming application programming interface (API) and Tweepy. There are several steps involved in the process of using Twitter's API to build the data set. This includes creating an account, installation of Tweepy, a quick test run, inspecting a tweets JSON, parsing out the data, identifying and collecting the data. Once the data are collected, it should be preprocessed before further steps.

Both data preprocessing and exploratory data analysis are included in the second phase. The raw tweets were incapable of creating unbiased results in sentiment prediction during the process of cleaning the data to uncover useable characteristics. #tags, @mentions, URLs, and stop words in tweets were the biggest roadblocks. Regular expression-based substitution were used to remove the #tags, @mentions, and URLs from the tweets. Stop words were handled by NLTK library in python.

The polarity scores of each tweet in the dataset can be computed using Python's TextBlob package [19]. The cleaned tweets from the previous phase were subjected to multiple assessment models using TextBlob, and a generic polarity score for each tweet was calculated. This score was linked to the type of words in the text, such as Unigrams, Bigrams, and Trigrams.

By classifying the polarity ratings of tweets into three groups, the sentiments phase expanded on the preceding phase. Positive emotions are ones that have a wide range (0, 1). Negative sentiments have a polarity of (-1, 0), while neutral sentiments have a polarity of 0.0. These three classes were saved in the dataset as a separate feature called "Sentiments." Parts of speech tagging (POS), lemmatization, stemming, and stop words removal were the different techniques used in the framework process as shown in Figure 2.

A part-of-speech tag (or POS tag) is a particular label assigned to each token (word) in a text corpus to denote the part of speech as well as other grammatical categories such as tense, number (plural/singular), and case [20]. POS tags are employed in text analysis tools and algorithms, as well as in corpus searches.

In linguistics, lemmatization is the act of bringing together the inflected forms of a word so that they may be examined as a single item, identified by the word's lemma, or dictionary form [21]. The process of reducing inflected

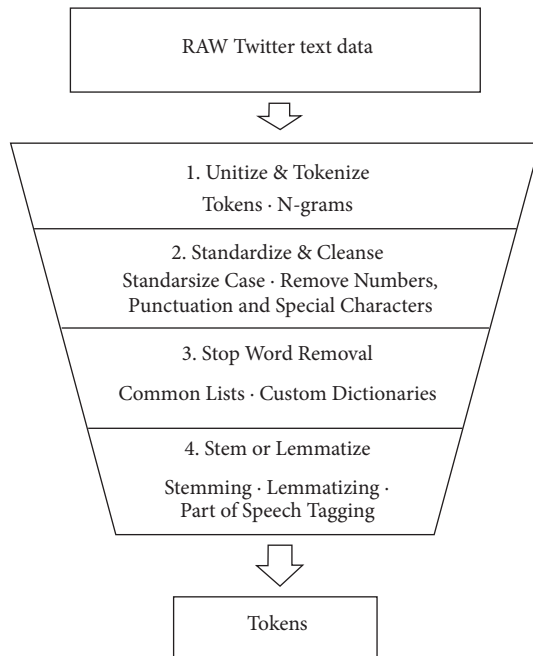


FIGURE 2: Sentimental analysis pipeline.

words to their word stem, base, or root form, generally a written word form, is known as stemming. The stem does not have to be the same as the word's morphological root; it is typically enough that related words map to the same stem, even if that stem is not a valid root in and of itself. To eliminate stop words from a sentence [22], divide your text into words and then check to see if the word is in the NLTK list of stop words.

All things in Twitter are built on the foundation of tweets. "Status updates" is another term for tweets. The Tweet object has a huge list of "root-level" characteristics, including id, created at, and text, among others. Tweet objects also serve as the "parent" object for a number of child objects. The key attributes from the tweets include the following:

- (a) text: the tweet's actual text
- (b) created at: the date the tweet was created
- (c) text: the text of the tweet itself
- (d) favorite count, retweet count: the total number of favourites and retweets favorited
- (e) lang: the language's abbreviation
- (f) id: the tweet identifier
- (g) geo: if available, geo-location data
- (h) user: author's complete profile
- (i) entities: URLs, @-mentions, hashtags, and symbols are examples of entities

Data preparation is highly recommended for a variety of reasons, including datum or database quality, data analysis process, and the capacity to apply related algorithms for removing noisy and missing data and increasing data reliability, as high-quality data models require high-quality

data. Sentiment analysis entailed determining the subjectivity contents of a given piece of material by first pre-processing it to detect stop words and symbols and then checking the subjectivity contents [23]. Machine learning approaches and lexical-based methods are both used to determine the polarity of an opinion's substance. Sentiment classifies information as either favourable, negative, or neutral.

Feature selection is a method for condensing a large number of attributes into a smaller subset with the maximum accuracy [24]. The limitation of overfitting, the improvement of accuracy, and the reduction in training time are all advantages of using this option on the data. Filters and wrappers are two types of feature selection strategies. Filters utilize statistical tests like Infogain, Chi-square, and CFS to determine the best subset of features, whereas wrappers use a learning algorithm to determine the best subset of features.

In the literature, the majority of comprehensive techniques extract unigrams, bigrams, and unigrams and bigrams as three separate features to apply to Naive Bayes, maximum entropy, and support vector machines classifiers, respectively [25]. N -grams features have been frequently employed in twitter sentiment classification to speed up processing speed. In more concrete terms, information gain for classification is a measure of how frequently a feature appears in one class compared to all other classes. A unigram with high information is one that appears frequently in good tweets and rarely in negative tweets. In the next section, we detail the proposed approach through n -gram stack encoder followed by setting the rules and training the neural network for classifying streaming data.

4. Proposed Design

Sentiment analysis is an automated approach of identifying and classifying subjective information in text data. An opinion, a judgment, or a visceral response to a particular issue or element of a product might be included. Polarity detection is the most popular method of sentiment analysis, which involves categorizing statements as positive, negative, or neutral. Sentiment analysis makes sense of human language using natural language processing and machine learning to produce correct results automatically. Connect sentiment analysis tools directly to social networks so that we can keep track of tweets as they come in and obtain real-time insights from social mentions.

We focus on an automatic method for sentiment analysis in order to examine these vast amounts of data. For sentiment analysis on tweets, we train a neural network. A pretrained word embedding follows the network's text input layer. The vectors are then input into a 128-dimensional fully connected layer with 50% dropout and sigmoid activation, followed by a regression output layer.

An n -gram is an n -unit string that is a portion of a bigger string. Characters and entire words can be used as these units. A word is just a string of x characters separated by a space on both sides. Because word borders are also essential information, spaces are usually included when utilizing character n -grams for classification.

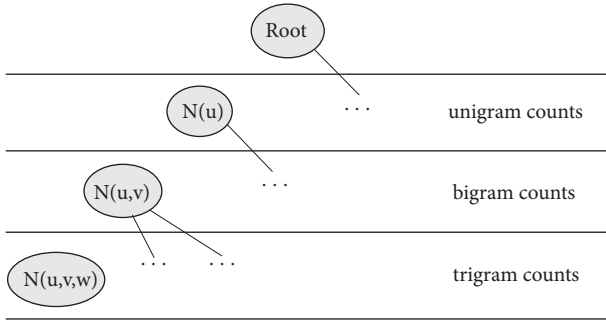


FIGURE 3: N -grams model for generative applications.

An n -gram is a contiguous sequence of n elements from a given sample of text or speech in the fields of computational linguistics and probability as shown in Figure 3. The components might be phonemes, syllables, letters, words, or base pairs, depending on the application. n -grams are often retrieved from a text or audio corpus. N -grams are also known as shingles when the components are words. In statistical natural language processing, n -gram models are commonly utilized. Words are modelled in such a way that each n -gram is made up of n words for parsing.

The lack of explicit representation of long-range dependency in n -gram models is frequently criticized [26]. Because an n -gram model's only stated dependency range is $(n-1)$ characters and because natural languages contain many occurrences of unbounded dependencies, an n -gram model cannot discriminate unbounded dependencies from noise in principle. But in practice, n -gram models have proven to be particularly effective in modelling language data, which is a critical component of modern statistical language applications.

In general, word-level N -grams give "better" outcomes for generative applications, but they come with their own set of problems. An N -Gram grammar is a representation of a N th order Markov language model in which the existence of $N-1$ extra symbols affects the chance of a symbol appearing. Word sequence probabilities are derived by the cooccurrences of words in the corpus, and N -Gram grammars are frequently created using data taken from a huge corpus of text. N -Gram grammars have the advantage of covering a considerably bigger language than would customarily be possible with a corpus-based approach.

When you count the number of n -grams shared by two strings, you get a measure of how similar they are. This metric is impervious to a wide range of textual errors, including misspellings. n -grams are commonly used as features in machine learning to predict the particular qualities of components in a dataset.

All of the tweets in the training set can be transformed to n -grams in the case of hashtag prediction. After that, n -gram frequency counts, for example, might be used to find patterns in the data. Specific hashtags could be linked to specific n -gram frequencies for tweets. A hashtag prediction system of maximum possibilities can be created in this fashion. A method like this might also extract all n -grams inside a specific range.

The amount of unique words identified in the training data determines the number of features in normal circumstances. The program can extract n -grams at both the character and word level. Other settings in this method can be predetermined, such as omitting stop words and non-ASCII letters. The non-ASCII letter should be removed before sentiment analysis to improve the performance.

4.1. Algorithm. We will begin with the uniform model. Because it assigns the same probability to every word in the text, this model will have very low average log likelihoods on the evaluation texts. After that, we combine this uniform model with the unigram model and reevaluate it using the evaluation texts.

The bigram model is then added to the mix. Similarly, each model in this three-model interpolation will have the same interpolation weight of $1/3$. We continue to add greater n -gram models to the mix while maintaining the same mixture weights across models.

Step 1: set the weight of each n -gram model to a number between 0 and 1, so that the total is less than 1.

Step 2: the goal function is differentiated with regard to each n -gram model weight (a_1 to a_5). The gradient of the objective function defined with respect to that of weight is often referred to as this derivative.

Step 3: each model's weight should be increased by a fraction of its gradient from Step 2. Because it influences how quickly the model weights are updated in each step, this fraction is frequently referred to as learning rate.

$$a_j = a_j + \lambda \frac{\partial J}{\partial a_j}, \quad (1)$$

where j = mapping cost, starts from 0 and increases during each iteration until average log likelihood reaches the max value.

Step 4: by subtracting the new n -gram model weights from 1, the uniform model's weight can be adjusted.

Step 5: finally, we repeat steps 2 and 3 until our average log likelihood has reached its maximum value.

4.2. Training the Model. For a given n -gram model, we have the following.

- (1) The probability of each word is decided by the $n-1$ words that are before it. In a trigram model ($n=3$), for example, the probability of each word is determined by the two words that come before it.
- (2) The fraction of times this n -gram appears among all the preceding $(n-1)$ -grams in the training set is used to calculate this probability. To put it another way, training the n -gram model simply entails calculating these conditional probabilities based on the training text.

As a result, we can utilize the average log likelihood as the n -gram model's assessment metric once more. The stronger our n -gram model is, the higher the average probability it assigns to each word in the evaluation text.

More sophisticated methods for predicting the likelihood of a word w given a history h , or the probability of an entire word sequence W , will be required. To determine the probabilities of complete sequences such as $P(w_1, w_2, \dots, w_n)$, we must use the chain rule of probability to decompose the probability:

$$w_{ij}(t+1) = w_{ij} + \eta \frac{\partial \log(p(v))}{\partial w_{ij}}. \quad (2)$$

Now, by applying the chain rule to words, we get

$$\begin{aligned} P(w_{1:n}) &= P(w_1)P(w_2|w_{1:1}) \dots P(w_n|w_{1:n-1}) \\ &= \prod_{k=1}^n P(w_k|w_{1:k-1}). \end{aligned} \quad (3)$$

The chain rule demonstrates the connection between computing a sequence's joint probability and computing a word's conditional probability given prior words. (3) implies that by multiplying a number of conditional probabilities together, we may estimate the joint probability of an entire sequence of words.

The proposed n -gram model is based on the idea that rather than assessing a word's likelihood based on its whole history, we can approximate it using only the latest few words. The bigram model uses only the conditional probability of the preceding word $P(w_n|w_{n-1})$ to approximate the probability of a word given all previous words $P(w_n|w_{1:n-1})$. A Markov assumption states that the probability of a word is exclusively controlled by the preceding word. Markov models are a type of probabilistic model that assumes we can anticipate the likelihood of a future unit without having to look too far back in time. As a result, the general equation for this n -gram approximation of the conditional probability of the following word in a sequence is

$$P(w_n|w_{1:n-1}) \approx P(w_n|w_{n-N+1:n-1}). \quad (4)$$

Maximum likelihood estimation, or MLE, is a simple method for estimating probabilities. By taking counts from a corpus and normalizing them to fall between 0 and 1, we can get the MLE estimate for the parameters of an n -gram model. The program can extract n -grams at both the character and word level. Other settings in this method can be predetermined, such as omitting stop words and non-ASCII letters as well.

Tokenizing strings is the process of converting strings into tokens. Separate strings are detected and labelled with a unique ID in this way. Token separators are frequently perceived as white spaces and punctuation marks. Tokens that appear frequently in the data are given a lower weight than tokens that appear infrequently. Each token represents a feature, which is represented by a feature vector. A multivariate sample is a vector that comprises all of the

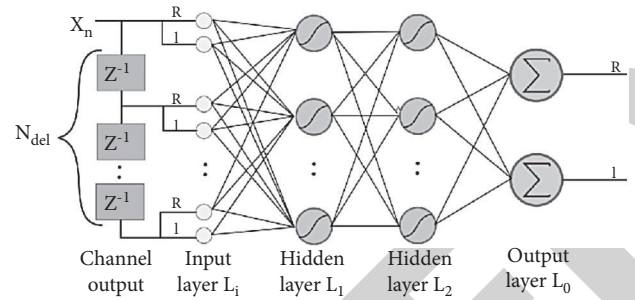


FIGURE 4: DBN learning architecture with two hidden layers and one output layer.

features for a specific document. Thus, vectorization is the generic process of converting textual data into numerical feature vectors that machine learning algorithms can handle. Setting the rules and training the neural network for classifying streaming data forms the second phase of the work and is discussed in detail in the next section.

The primary principle behind utilizing N -grams to generate text is to presume that the last word (x_n) of an n -gram may be deduced from the previous words in the same n -gram ($x_{n-1}, x_{n-2}, \dots, x_1$), which is known as context. So the model's basic premise is that in order to forecast the next word, we do not need to maintain track of the entire phrase; instead, we just need to look back for $n-1$ tokens.

5. Deep Belief Neural Network for Classification

A deep belief neural network (DBN) may learn to probabilistically recreate its inputs when trained on a set of examples without supervision. After that, the layers serve as feature detectors. A DBN can be further taught under supervision to do categorization after completing this learning stage [27]. DBNs are made up of simple, unsupervised networks like restricted Boltzmann machines (RBMs) or autoencoders, with the hidden layer of each subnetwork serving as the visible layer for the next. A generative energy-based undirected model with a "visible" input layer and a "hidden" layer, as well as connections between but not within layers, is known as an RBM. This results in a quick, unsupervised layer-by-layer training approach, in which contrastive divergence is applied to each subnetwork in turn, starting with the "lowest" pair of layers as shown in Figure 4.

The contrastive divergence approach is used to speed up the learning for an RBM, and the fundamental idea is to update all of the hidden units in parallel, starting with visible units, reconstructing visible units from the hidden units, and then updating the hidden units again. Layers of stochastic binary units with weighted connections make up a basic belief network. Furthermore, because the network is an acyclic graph, we can see what kinds of data the belief network believes in at the leaf nodes. The goal of a belief network is to infer the states of unseen stochastic binary units and adjust the weights between them so that the network can provide data that are equivalent to what is

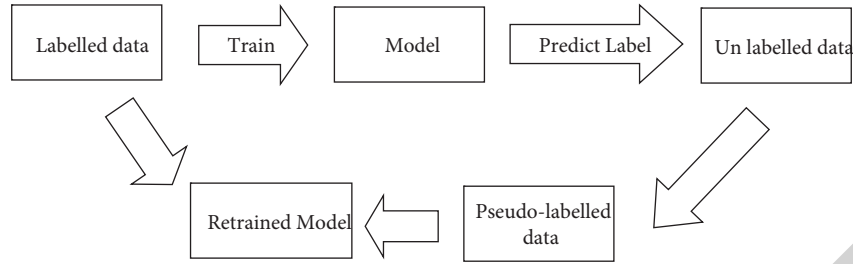


FIGURE 5: Labelling method.

observed. The stochastic binary units in belief networks have a state of 0 or 1, with the probability of reaching 1 controlled by a bias and weighted input from other units. This is represented as

$$w_{ij}(t+1) = w_{ij} + \eta \frac{\partial \log(p(v))}{\partial w_{ij}}. \quad (5)$$

Here $p(v)$ corresponds to the probability of a visible vector, η represents the learning rate, and W is the weight matrix between the hidden (i) and visible layers (j). The contrastive divergence (CD) technique is used to learn a layer of features from visible units as the initial stage in training DBN [28]. The next stage is to treat previously learnt features' activations as visible units and learn features from a second hidden layer. When the learning for the final hidden layer is completed, the entire DBN is trained. DBN may be trained using this simple greedy learning technique. This is because the CD algorithm in the training RBM seeks for the local optimum for each layer, and the following stacked RBM layer takes those optimally trained values and looks for the local optimum again. As each layer is continually educated to achieve the best value, the global optimum is likely to emerge at the end of this method.

We can employ persistent contrastive divergence because the sampling caused by contrastive divergence greatly biases the samples from the model distribution toward the most recent data. New samples are produced at each iteration by sampling conditioned on the most recently sampled hidden states, which are retained between data points, and the model distribution is randomly initialized.

By stacking RBMs and interpreting the hidden layer of the lower RBM as the visible layer of the next layer, DBNs can be created. It has been demonstrated that adding hidden layers and using the previously stated unsupervised learning methods for RBMs will enhance the lower bound on the training data's log-likelihood. Higher layers are more likely to encode abstract features, which are usually extremely useful for classification applications. After that, supervised learning methods can be used to train the DBN's top layer, and error backpropagation can be used to optimize the entire multilayer network for the job. DBNs can also be used to link together distinct sets of data. Preprocessing hierarchies for both inputs may be created independently in this case, and the top levels of both hierarchies can be considered

TABLE 1: Statistics overview of the tweets COVID-19 dataset.

Feature	Total	Unique	Percentage of tweets (%)
Hashtag	3653928	566308	30
Mention	5363449	1251963	40
Entity	11537537	331307	70

as a shared visible layer for a new association layer to be formed on top of them. As a result, DBNs can have tree-like structures as well as single hierarchies.

Data labelling by hand is both costly and time-consuming. We can use pseudo labelling to label our data to speed up the process of labelling it as shown in Figure 5. Pseudo labelling is a semisupervised labelling technique that still requires labelled data. Machine learning methods are used to develop models in pseudo labelling. As a class of unlabeled data, the model was then utilized to construct a pseudo label.

We utilized multinomial Nave Bayes as a model in this study. In this work, we created our sentiment analysis model by performing several experimental model variations. We built the first model variations using deep belief neural network. We also compared the accuracy of the three algorithms by using term frequency-inversed document frequency (TF-IDF) as a characteristic instead of TFIDF. Several deep neural network methods, including convolutional neural network (CNN) and long short-term memory (LSTM), were used to create the second sentiment analysis model variations (LSTM).

The major purpose of this project's exploratory data analysis phase was to become familiar with the data frame's columns and begin developing research ideas. Starting with the phase of importing the data with pandas, we can see that all of the following information is available for each tweet and that includes user name, description, followers, friends, favourites, id, date, text, hashtags, sources, and retweets. The depth of a convolution neural network suffers from overfitting due to the necessity to train a large number of hyper parameters. Dropout regularization is applied to completely linked layers to solve the problem of a large number of concealed units and their connections. Two key metrics, namely the polarity and subjectivity, are calculated. By sorting the data set by polarity scores and then presenting the top rows, we may examine these types of tweets and complete the sentimental analysis.

TABLE 2: Sample tweets classification from the data set.

Sample tweet	Sentiment category
Bright vision, a community hospital, is transferring all patients to create room for stable COVID-19 cases.	Mixed sentiment
Any fellow patriot who celebrates Boris contracting the Corona virus is a complete cunt.	Negative/sad
Twittizens, good morning I wish you a day without coronas.	Positive/joy
Perhaps if I lock my front door, the coronavirus will be kept away.	Anger
In order to infect visitors with malware, hackers create false coronavirus maps.	Fear
My heart hurts so much at the notion of Jacob's Nashville performance being cancelled. Please go away.	Negative/sad

TABLE 3: Classifier accuracy comparison with the proposed method.

N-gram	Type of attribute	Classifier accuracy			
		Proposed DBN (%)	Naïve Bayes (%)	SVM (%)	K-nearest neighbors (%)
Unigram	All twitter data	80.3	79.4	81.9	73.3
	Information gain >0	84.1	86.6	83.6	74.2
	Best 70% on ranking	88.1	88.0	83.2	73.5
Bigram	All twitter data	86.1	75.2	85.8	62.7
	Information gain >0	90.3	89.0	82.8	63.3
	Best 70% on ranking	89.5	83.0	87.8	62.7
1 to 3 gram	All twitter data	86.1	85.7	82.5	68.8
	Information gain >0	90.1	92.5	84.1	66.0
	Best 70% on ranking	89.0	88.3	83.8	66.4

6. Results and Discussion

The easiest approach to assess a language model's performance is to embed it in an application and track how much the application improves. Extrinsic evaluation refers to this type of end-to-end assessment. Extrinsic evaluation is the only way to determine whether a certain component change will actually help with the task at hand. An intrinsic evaluation metric is one that assesses a model's quality without regard to its application. A test set is required for an intrinsic evaluation of a language model. The probabilities of an n -gram model, like many other statistical models in our field, are derived from the training set or training corpus on which it is trained. The performance of an n -gram model on unseen data, referred to as the test set or test corpus, can subsequently be used to assess its quality.

The research focuses on sentiment analysis of Twitter messages utilizing the Python programming language with the Tweepy and TextBlob libraries. Tweepy allows us to find relevant information by using keywords, hashtags, timelines, trends, or geolocation. For searching the tweets with the COVID hashtag, the following command is used.

```
api = tweepy.API(auth, wait_on_rate_limit = True)
#important.
self.tweets = tweepy.Cursor(api.search, q = 'COVID-19,
lang = "en").items(100)
```

The prepared dataset, which included positive, negative, and neutral classes, was used in the experiments. A total of 47,000 tweets were used in the training process, with 11,000 being used in the testing cycle. Different N -gram models were combined with a set of classifiers in the suggested strategy in Table 1.

TABLE 4: Words classification based on different sentiments across time periods.

Sentiment	March 2020	April 2020	May 2020	June 2020
Positive	32430	32437	31572	26507
Negative	34181	31538	37410	20677
Fear	31496	35542	34982	29184
Noncategorized	136916	155750	147115	101254
Total	235023	255267	251079	177622

In order to determine the best settings for the preprocessing approaches and classifiers, we ran a series of experiments to see which alternatives would yield the most accurate results. Section 3 includes a detailed description of the chosen preprocessing methods. Some of the sample tweets are listed below in Table 2 with corresponding sentiment category.

We notice that when we select qualities with an information benefit greater than zero, the number of them is reduced significantly. We chose around 70% of the attributes ranked as more deserving for the second feature extraction, where the attributes are analysed using the random forest algorithm. The suggested classifier accuracy is compared to other techniques in the literature in Table 3.

Table 3 shows how classifier performance varies based on the preprocessing methods used. The behaviour of dataset representations is not consistent. There is no representation that consistently produces better results when compared to others. In general, 1-to-3-grams outperform the other formats; however, they are a close second to unigram.

The attribute selection process enhances classification performance when compared to picking all attributes, according to our findings. This results from the elimination of redundant and irrelevant features from datasets, which

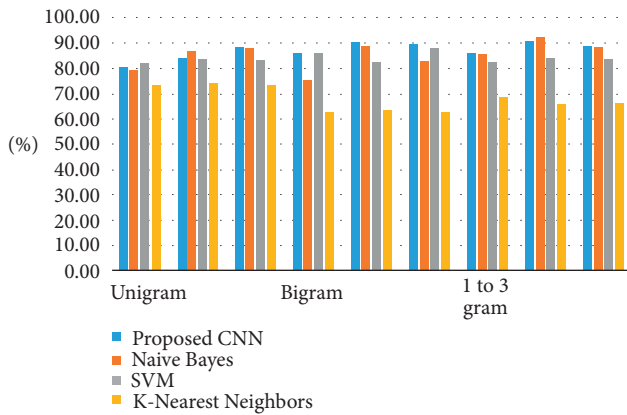


FIGURE 6: Classifier accuracy between proposed and literature methods.

can lead to overfitting by deceptive modelling techniques. Table 4 shows the general community sentiment for the study period, organized by sentiment words. It demonstrates that in March and April, the public had a very favourable attitude. Late May saw a minor decrease, followed by a substantial decrease in June. This could be attributable to an increase in COVID-19 confirmed cases. There were peaks and valleys in negative sentiment.

Some words, such as “coronavirus” and “Wuhan,” have the highest frequency in our database. In the tweets column, there are a variety of #hashtags. However, they are nearly same in all attitudes; therefore, they are unable to provide us with useful information. The efficiency of the proposed algorithm is shown in Figure 6 when compared to other categorization methods in the literature.

It is evident from Figure 6 that the proposed method outweighs the other existing methods in the literature in terms of classifier accuracy. The number of repetitions also affects classification accuracy. As the number of iterations is impacted by the size of the experiment, we chose iteration = 50, 100, 150, 200, and 300 and then manually examined the data throughout each iteration. The results reveal that in the range from 100 to 200 iterations, there is no significant change. As a result, we limited our model to 150 iterations. While sentiment analysis of tweets can reflect public opinion, it cannot express the true impact of the virus, as evidenced by the rapidly increasing number of cases and deaths every day.

7. Conclusion

With the rise in popularity of social networking sites, they have become a strong tool for influencing people and disseminating information to the general public. However, due to the lack of contextual information in the texts, sentiment analysis for brief texts such as Twitter is particularly difficult. As a result, numerous algorithms are always being created in order to achieve the finest sentiment analysis model result. A preliminary phase of text preprocessing and feature extraction is required to complete the classification operation. Because preprocessing activities have an impact on

classification quality, we run a number of experiments on various generated datasets. Deep belief networks are a type of deep architectural network that is constructed from stacks of restricted Boltzmann machines. DBNs may also be utilized for tasks in both an unsupervised and supervised situation, and they take full advantage of outstanding techniques like unsupervised pretraining and fine tuning on a downstream job. The results of our research show that by selecting and representing features correctly, sentiment analysis accuracy can be increased. The produced datasets were analysed in terms of positive and negative attitudes, fear, and trust emotions expressed in the tweets. To avert anarchy and panic, policymakers can adopt public opinion surveillance techniques. For well-intentioned data, Twitter can be used. All governments can benefit from timely awareness of public mood in order to establish an effective strategy for better managing the situation and a communication strategy for disseminating accurate and trustworthy information and engaging the public in the appropriate response activities. This will be our future directions as well. In future, additional twitter data can be considered for experimentation. The impact of COVID-19 on financial sector, employability, and personal life of the individuals may be analysed and prediction will be performed using machine learning techniques.

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.

References

- [1] P. H. Chen, S.-F. Tsao, L. Li, Y. Yang, T. Tisseverasinghe, and Z. A. Butt, “What social media told us in the time of COVID-19: a scoping review,” *The Lancet Digital Health*, vol. 3, 2021.
- [2] T. Vijay, P. Karmakar, A. Chawla, and B. Dhanka, “Sentiment analysis on COVID-19 twitter data,” in *Proceedings of the 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, Jaipur, India, December 2020.
- [3] N. Chintalapudi, G. Battineni, and F. Amenta, “Sentimental analysis of COVID-19 tweets using deep learning models,” *Infectious Disease Reports*, vol. 13, no. 2, pp. 329–339, 2021.
- [4] P. Song, C. Geng, and Z. Li, “Research on Text Classification Based on Convolutional Neural Network,” in *Proceedings of the International Conference on Computer Network, Electronic and Automation (ICCNEA)*, Xi’an, China, September 2019.
- [5] H. Liang, Y. Sun, X. Sun, and Y. Gao, “Text feature extraction based on deep learning: a review,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, 2017.
- [6] S. Das, A. K. Chakraborty, and A. Kumar Kolya, “Sentiment analysis of covid-19 tweets using evolutionary classification-based LSTM model,” in *Advances in Intelligent Systems and Computing*, Singapore, 2021.

Retraction

Retracted: Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] T. Veeramakali, A. Shobanadevi, N. R. Nayak, S. Kumar, S. Singhal, and M. Subramanian, "Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4690936, 8 pages, 2022.

Research Article

Preserving the Privacy of Healthcare Data over Social Networks Using Machine Learning

T. Veeramakali,¹ A. Shobanadevi ,¹ Nihar Ranjan Nayak,² Sumit Kumar,³ Sunita Singhal,⁴ and Manoharan Subramanian ⁵

¹Department of Data Science and Business Systems, School of Computing, SRM Institute of Science and Technology, Kattankulathur, Chennai, India

²Sri Venkateswara College of Engineering Technology, Chittoor, India

³Indian Institute of Management, Kozhikode, India

⁴Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur 303007, India

⁵Department of Computer Science, School of Informatics and Electrical Engineering, Hachalu Hundesa Campus, Ambo University, Ambo, Ethiopia

Correspondence should be addressed to Manoharan Subramanian; manoharan.subramanian@ambou.edu.et

Received 1 February 2022; Accepted 11 April 2022; Published 5 May 2022

Academic Editor: Gabriella Casalino

Copyright © 2022 T. Veeramakali 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.

A key challenge in clinical recommendation systems is the problem of aberrant patient profiles in social networks. As a result of a person's abnormal profile, numerous vests might be used to make fake remarks about them, cyber bullying, or cyber-attacks. Many clinical researchers have done extensive study on this topic. The most recent studies on this topic are summarized, and an overarching framework is provided. When it comes to the methods and datasets that make up the data collection, the feature presentation and algorithm selection layers provide an overview of the various types of algorithm selections available. The categorization and evaluation of diseases and disorders has been one of the major advantages of machine learning in medical. Because it was harder to predict, it rendered it more controllable. It might range from difficult-to-find cancers in the early stages to certain other illnesses spread through the bloodstream. In healthcare, we may pick methods in machine learning depending on reliable outcomes. To do so, we must run the findings through each method. The major issue arises during information training and validation. Because the dataset is so large, eliminating mistakes might be difficult. The providers, other characteristics, various algorithms, data labelling techniques, and assessment criteria are all presented and contrasted in depth. Detecting anomalous users in medical social networks, on the other hand, is a work in progress. The result evaluation layer provides an explanation of how to evaluate and mark up the results of the various algorithm selection layers. Finally, it looks forward to more study in this area.

1. Introduction

With the widespread application of the development of mobile Internet technology, online healthcare social networks have quickly become an essential part of people's network life due to their convenience, flexibility, and rich content. However, the vast user privacy information in healthcare social networks and its vast commercial value have often become the targets of criminals who attempt to commit illegal activities. Among them, abnormal users are one of the standard methods criminals use to attack

healthcare social networks. For example, merchants distort product value orientation for commercial interests [1–3]; criminals use multiple vests to deceive Internet users, steal information, or even online fraud [4–6]. According to statistics, the types of abnormal users also show a variety of forms due to different types of healthcare social networks. As shown in Table 1, they are widely present on different social platforms. For example, as of June 2012, Face book has about 8.7% (8.3×10^7) fake users, while Twitter faces the same problem. About 5% of users are fake users. Some experts believe that this proportion is possible up to 10% [7].

TABLE 1: Types of abnormal users.

Abnormal user types	Main existence platform
Spammer	Twitter, Weibo platform, CNN, Amazon, Yelp
Hoaxes	Wikipedia
Bot	Twitter
Sock-puppet	Wikipedia, Twitter, Facebook, CNN
Troll	Slashdot zoo, Wikipedia
Fake user	Google, Twitter, Amazon
Vandal	Twitter, Wikipedia, TCP

Abnormal users in healthcare social networks have a wide range of existence and severe harm. Related scholars have summarized the research on the detection technology of strange users in healthcare social networks. Sun et al. [8] summarized the status quo and associated technologies of abnormal patient profile and abnormal behaviour detection in healthcare social networks. Song et al. [9] focused on analysing malicious patient profile algorithms and their applications based on features, space, and density. In the study by Hu et al. [10], focusing on four aspects of traditional spam, false comments, spam, and link factories, network characteristics, content characteristics, and behaviour characteristics are extracted. It explains the application of clustering algorithms, classification algorithms, and graph algorithms. For the problem of false information detection, Yuan et al. [11] studied graph-based algorithms. They divided them into subgraph analysis and mining algorithms, label transfer algorithms, and hidden factor decomposition algorithms. The research mentioned above only summarizes the current work from features or algorithms and is not comprehensive enough.

This article summarizes the implementation process of social network abnormal patient profile technology, as shown in Figure 1. The data collection layer introduces data acquisition methods and related data sets; the feature presentation layer explains attributes features, content features, network features, activity features, and additional features; the algorithm selection layer introduces supervised, unsupervised, and graph algorithms; the result evaluation layer explains data labelling methods and method evaluation indicators.

This research is organized as follows. Section 1 describes the introduction, the supervised algorithm is described in Section 2, Section 3 describes the unsupervised learning, the graph algorithm is described in Section 4, Section 5 describes comparison of different algorithms, evaluation parameter is described in Section 6, and finally Section 7 describes the conclusion part.

2. Supervised Algorithm

When the acquired data contain tags, the researchers design numerical features based on the idea of classification and divide users into abnormal users and regular users to detect strange users. Supervised algorithms can be divided into single classification algorithms and integrated classification algorithms:

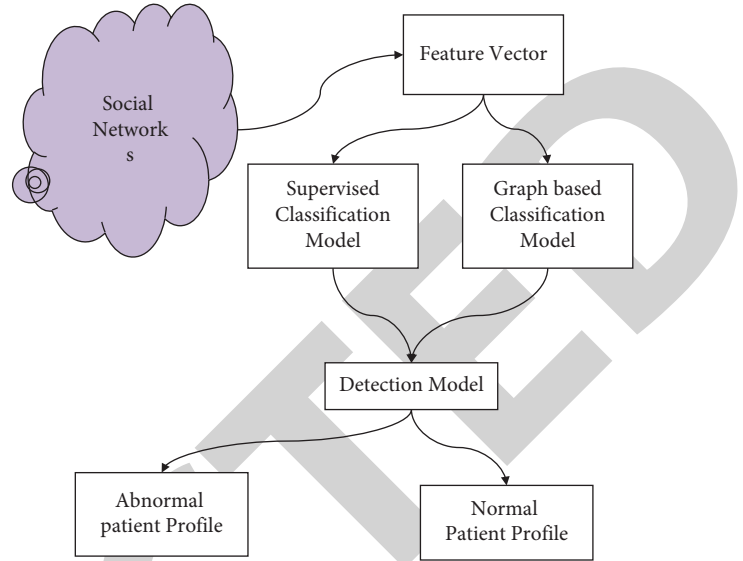


FIGURE 1: Abnormal user architecture of healthcare social networks.

- (1) Single classification algorithm refers to using only one classification algorithm to detect abnormal users. Commonly used is logistic regression, support vector machines, decision trees, and so on. Tara [5] and Qi et al. [7] used logistic regression to detect malicious users in Twitter and phishers in CNN and found that the name language pattern feature is the most significant feature that distinguishes malicious users from normal users. Jiang et al. [12] used logistic regression to detect false reviews in Amazon. According to the classification of products and assessments, it emphasizes detecting 4 more dangerous faulty comment areas. It is found that the characteristics of response activities are the most effective for the problem, and Zhu [2] found that the overall score deviation is essential for detection. Fake comments have no effect. Meng et al. [13] used support vector machines to detect Wikipedia's vests and found the 30 features that contributed the most to the problem through experiments. Zhang et al. [14] used support vector machines to detect abnormal behaviors of network users. Tables 2 and 3 show the comparison of detection characteristics and classification of detection algorithm, respectively.

The spammers in Twitter introduce the parameter J into the support vector machine to give the prior distribution of spammers, adjust J to balance the improvement in accuracy or recall, and select through gain and chi-square tests. There are several distinguishing characteristics. For example, Venkatesan et al. [15] used decision trees to detect cultural attackers in Wikipedia, sacrificed part of the accuracy requirements to achieve a higher recall rate, and won the 2010 PAN competition. Wang et al. [16] used decision trees and Bayesian network algorithms to detect false news about Hurricane Sandy on

TABLE 2: Comparison of detection characteristics.

Detection feature	Features	The essential	Feature evaluation
Attribute characteristics	Using artificial design methods, it is easy to bypass attackers. The algorithm design is simple; the efficiency is low, the accuracy rate is relatively low, the data level is trim, and it has strict privacy protection.	Breakthrough privacy protection	Uncommonly used
Content characteristics	Natural language processing method is adopted, which is easy to be bypassed by attackers. In addition, the algorithm design is complicated, the efficiency is low, the accuracy rate is relatively low, the data level is significant, and the privacy protection is slight.	Design complex algorithms and reasonable language models	Commonly used
Network characteristics	Adopting complex network processing methods, not easy to be bypassed by attackers, simple algorithm design, low efficiency, relatively low accuracy rate, significant data level, and no privacy protection.	Master the global structure	Mainstream
Activity characteristics	Using behavioural pattern analysis and processing methods, it is not easy to bypass attackers; the algorithm design is simple, the efficiency is high, the accuracy rate is high, the data level is significant, and the privacy protection is slight.	Select the most distinguishable activity information	Mainstream
Auxiliary features	Using time-series model analysis, it is not easy to bypass attackers; the algorithm design is complex, the efficiency is high, the accuracy rate is high, the data level is trim, and it has slight privacy protection.	Effective use of time dimension information	Popular

TABLE 3: Classification of detection algorithms.

	Algorithm	Thought	Algorithm key issues
Supervised algorithm	Single classification algorithm	Classification	Select the key feature with the greatest degree of discrimination
	Integrated classification algorithm		
Unsupervised algorithm	Decomposition mining from top to bottom	Clustering	Choose a reasonable similarity index
	Cluster mining from bottom to top		
Graph algorithm	Spectral decomposition Random walk	Graph outlier detection	Processing large-scale sparse graph data

Twitter and found that the effect of decision trees is better. In addition, the contribution of text features is relatively significant. When transaction data (super clever agreement) technique is combined with traditional relational database strategies, data security, authenticity, time management, and other aspects of data regime are significantly improved. Logistic regression was employed by Jiang et al. [12] to detect fake reviews on Amazon. It stresses on four more risky defective remark areas, according with categorization of items and evaluations.

- (2) Integrated classification algorithm integrates multiple single classification algorithms to obtain higher accuracy, such as random forest, Adaboost, etc. For example, Kanhere et al. [17] used the random forest algorithm to detect abnormal users in discussion communities such as CNN. They found that the longer the sample time, the worse the prediction accuracy of the method, which confirmed that changes in user behaviour quickly lead to abnormal users. Wang et al. [18] used six classification algorithms, including random forest and Bayesian network, to detect the vests in Wikipedia. They found that the best features were reply

frequency, increased bytes, and average contribution through experiments. Noh et al. [19] used the Adaboost method which detects the political navy in Twitter, uses the chi-square test to give the 10 most contributing characteristics, and analyzes the characteristics of the discovered political navy. Shalash et al. [20] used support vector machines, random forests, and Adaboost methods to detect deception in healthcare social networks. The Adaboost way is more effective, and this newly defined indicator is more effective. Bhanumurthy et al. [21] integrated Bayesian, NSNB, Winnow, and other algorithms into linear joint algorithms and obtained the effectiveness of each algorithm for the problem by optimizing the weights.

3. Unsupervised Algorithm

When the sample data do not contain labels or contains few titles, based on the idea of clustering, researchers propose to use unsupervised learning algorithms to solve the problem of abnormal patient profile. Unsupervised algorithms are divided into decomposition mining from top to bottom and cluster mining from bottom to top.

- (1) *Decomposition Mining Algorithm from Top to Bottom*. When there is no label for the sample data, the researchers detect abnormal user groups by decomposing the social network graph. Perez et al. [22] constructed an SVN network based on topic similarity, deleted a part of edges based on text feature similarity to form an SPN network, clustered and mined abnormal users' group communities based on the similarity of modulus, and gave the accuracy of the method. The TIA algorithm [23] initializes normal users and malicious users according to different centrality value boundaries, then takes various decomposition diagram operations according to different attack modes, and continuously updates malicious users and regular user groups to achieve the purpose of predicting malicious users in the Slashdot network. The D-CUBE method [24] decomposes the relationship tensor by deleting the attribute value dimension with the largest cardinality or density, until an abnormal group is left at the end, and iteratively obtains multiple deviant user groups. This method uses a distributed algorithm. It is suitable for large-scale graph data format. The ND-SYNC method [25] is directly based on the RTFRAUD way for community discovery of the constructed feature space, using the deviation of internal and external synchronization to detect group anomalies and find Fake users in Twitter. There are various attributes of the data set to decide the user is abnormal or normal such as labelling method. In order to evaluate aberrant user detection systems, you must first learn how to label data. The labelling results are not persuasive despite the fact that his characteristics are simple and easy to implement; however, the data base has a high accuracy rate. It is difficult to control from the web page.
- (2) *Bottom-Up Cluster Mining Algorithm*. When a part of the data sample labels are available, the researchers use the similarity and known label samples to cluster the graph structure to solve the problem of abnormal user group detection. The Copy Catch method [26] mainly constructs the time matrix of healthcare social networks, clusters to maximize the number of strange users in the core of TNBC, detects abnormal user attack groups in Face book, and provides proof of stability and convergence. He et al. [27] constructed similar groups based on the MD5 similarity of the text and the same clustering of the URL pointing to the target and then judged whether each group is a fake user group through the distribution coverage of counterfeit users and the time burst. Lin et al. [28] found a method called Eigen-Spoke's new model and used the model's score to cluster samples until the model no longer increased to detect the social network and user groups. Jain et al. [29] initialized a small number of robots in Twitter and clustered them using the similarity of text information until a sufficient number of robots were found. The Catch-Sync

method [30] redefines synchronization and normality.

4. Graph Algorithm

Graph algorithms are becoming increasingly popular for spotting unusual users due to the increasing importance of network architecture and activity factors found in graph data. Among graph-based algorithms, spectral decomposition and random walks are two of the most commonly used techniques.

According to this approach, processing results in a characteristic matrix may be corrupted in order to create alternative groups. Scholars have worked hard to use spectrum-based decomposition methods to solve the problem of patients with aberrant profiles. If you look at [25], for example, the author created a hierarchical tree structure, combined the content matrix with a sparse representation, and then utilised the spectral decomposition simulation approach to iterate the optimal weights in order to discover the deceiver further. According to [29], the author used both a content matrix and a random walk network matrix to get the final result. Spammers on Twitter were tracked down using the spectral decomposition approach. Following [31], a new method was devised by author that integrated emotional information into both content and adjacency matrixes. However, there have been contributions from other researchers. As an example, author in [32] used a threshold and a seed to generate a kid.

Using the subgraph, the figure decomposes it into a smaller subspace in order to determine whether or not the user in the subspace is a phony YouTube user. The FEMA method [33] decomposes the three-dimensional tensor at different times to get the mapping matrix and the core tensor according to specified regularization criteria in response to the growth in the dimension of impact, notably the time component. It also increases your chances of being a strange user. To better identify assaults, the author in [17] employed SVD matrix decomposition to rebuild the degree of network nodes and imposed a restriction on the degree to better hide them from detection.

In the random walk algorithm, the node transition path is calculated, the unknown node's relationship to the known node or the transition probability is determined, and if the unknown node is abnormal, it is determined. Currently, the system is used to identify vest accounts in healthcare social networks. Bhanumurthy and Anne [33] technique employs a modified random walk algorithm to compute the transfer path of the node. If the node path and the path of the standard node cross, it is determined to be a regular user. However, each assault path may sustain at most $O(n \log n)$ vests node. If the end edge of the node path and the last edge of the standard node path coincide, then the user is considered a normal user and the number of tolerated vest nodes is reduced to $O(\log n)$ for each attack path [34]. The two approaches described above, however, can only identify one node as being inefficient in each cycle. This is how the author of [35] identifies multiple vest accounts quickly by taking

TABLE 4: Comparison of detection algorithms.

Detection algorithm	Advantage	Shortcoming
Supervised algorithm	(1) High accuracy	(1) Need to include label data
	(2) Fast detection speed and high efficiency	(2) Need to train in advance
	(3) Mature algorithm design and mature deployment technology	(3) Need to select distinguishing features
	(4) Good real-time	(4) The selected features are easy to be bypassed by attackers, and the unknown mode detection effect is poor
Unsupervised algorithm	(1) Only need not contain label data	(1) Low accuracy
	(2) No need to train in advance	(2) The algorithm design is complicated, and the efficiency is low
	(3) Effective detection of unknown patterns	(3) Poor real-time performance
Graph algorithm	(1) Only graph data are needed	(1) Low accuracy and poor real-time performance
	(2) No need to train in advance	(2) The theoretical assumptions are complicated, and the reality is untenable
	(3) Effective detection of unknown patterns	(3) The algorithm design is complex, and the efficiency is low
		(4) There are different differences in social networks

TABLE 5: Method evaluation indicators.

Methodological review Estimated index	Formal definition
Accuracy	$A = TP + TN / TP + TN + FP + FN$
Precision	$P = TP / TP + FP$
Recall	$R = TP / TP + FN$
F1-score	$F = 2PR / P + R$
MCC	$MCC = TP * TN - FP * FN / \sqrt{(TP + FP)(TP + FN)(TN + FN)(TN + FN)}$

random walks from a normal node and then performing a similar operation on a node, using the same bar to determine whether it is a vest or not. Then, using the discovered vest node and the same principle to identify multiple vest accounts, the author can quickly identify multiple vest accounts. Assigning credibility levels to other nodes using the Sybil Rank algorithm [4] involves using the random walk algorithm. According to the standardized and degree value findings, the nodes with a lower credibility value are placed at the bottom of the list. It is a suspicious node. Markov random fields are used in the Sybil Belief approach [13] for detecting vesting accounts. To begin, a random value is assigned to each node in the network to determine if it is a normal or vest node. Then, using a Markov random field, each node's posterior probability is determined. To put it another way, there is a 50% chance that the node is not abnormal.

5. Comparison of Different Algorithms

Various detection algorithms have their advantages and disadvantages and have their application scenarios, as shown in Table 4, which lists the advantages and disadvantages of different detection algorithms.

5.1. Method Evaluation Layer. After selecting the appropriate feature representation and algorithm selection, researchers need to evaluate the effect of the method to a certain extent and need to obtain data annotations and method evaluation indicators.

5.2. Data Labelling Method. How to label data is a prerequisite for evaluating abnormal user detection methods. Although his features are simple and easy to implement, the labelling results are not convincing; however, the blacklist has a high accuracy rate. It is not easy to manage from the website.

6. Evaluation Parameter

Generally, the commonly used method evaluation indicators include accuracy, recall, precision, F1-score, and ROC (AUC) curve index. The researchers propose using unsupervised learning methods to overcome the problem of anomalous patient profiles when the sample data do not have labels as few titles, based on the principle of clustering. The finding analysis layer explains how to assess and mark up the outcomes of the several methods selection stages. The ROC curve is a curve drawn based on the confusion matrix's actual rate and false positive rate values as the coordinate axes, and AUC represents the area under the ROC curve. These indicators are based on a confusion matrix, and the definitions of various indicators are shown in Table 5.

This paper presents comparative analysis over supervised machine learning (SVM) and graph technique to identify abnormal patient profiles for reliable healthcare data over two healthcare dataset, i.e., Pubmed [36] and Medhelp [37]. It determines that the SVM leads over graph-based technique (GBT) and acquires 89%–92% accuracy, whereas SVM gains only 81%–84% accuracy over Pubmed (PM) and Medhelp (MH) dataset, respectively, as shown in Table 6 and Figure 2.

TABLE 6: Comparison of graph-based technique and SVM.

Technique	Accuracy		Precision		Recall		F1-score	
	PM	MH	PM	MH	PM	MH	PM	MH
Graph-based technique	81.56	84.24	74.12	76.52	65.64	69.84	75.52	76.24
SVM	89.32	92.14	78.28	81.86	75.52	76.52	86.61	88.34

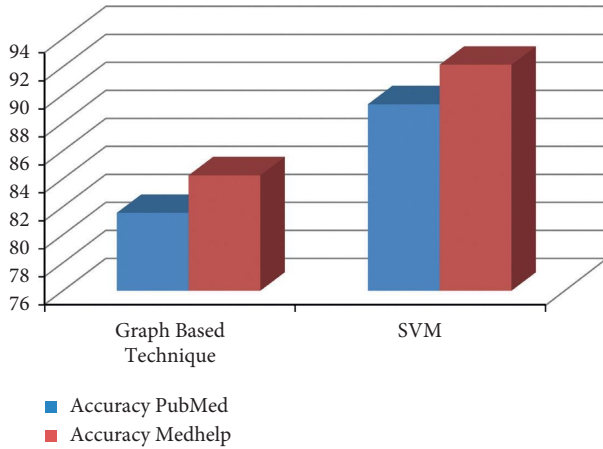


FIGURE 2: Accuracy of SVM and graph-based technique.

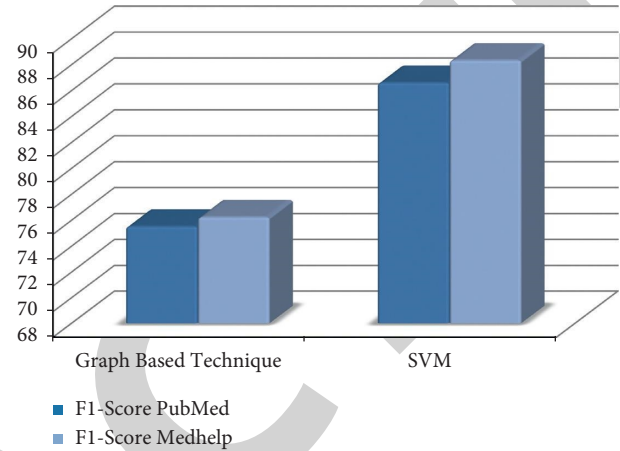


FIGURE 5: F1-score of SVM and graph-based technique.

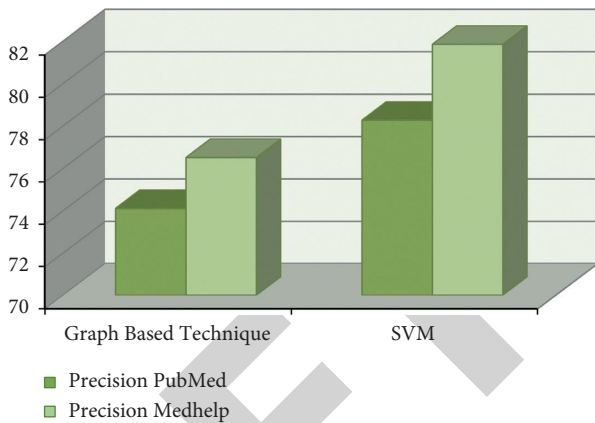


FIGURE 3: Precision of SVM and graph-based technique.

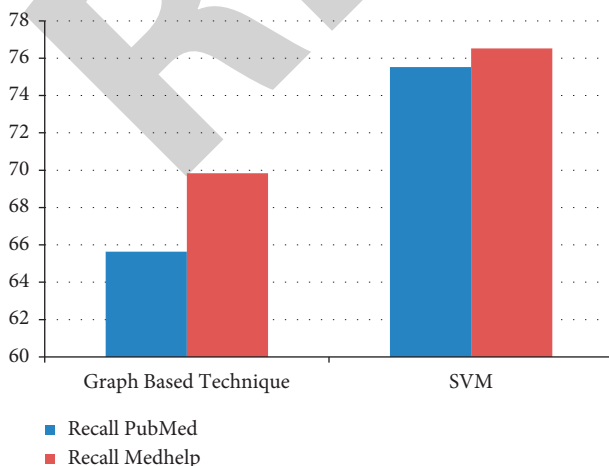


FIGURE 4: Recall of SVM and graph-based technique.

However, SVM gains 78.28%–81.86% and GBT acquires 74.12%–76.52 precision over Pubmed (PM) and Medhelp (MH) dataset, respectively, as shown in Table 6 and Figure 3.

However, SVM gains 65.64%–69.84% and GBT acquires 75.52%–76.52% recall over Pubmed (PM) and Medhelp (MH) dataset, respectively, as shown in Table 6 and Figure 4.

However, whereas SVM gains 75.52%–76.24% and GBT acquires 86.61%–88.34% F1-score over Pubmed (PM) and Medhelp (MH) dataset, respectively, as shown in Table 6 and Figure 5.

7. Conclusion

With the gradual increase in the influence of healthcare social networks, more and more malicious users are focusing their attacks on healthcare social networks. Among them, the harm of abnormal users to healthcare social networks seriously threatens the information security of healthcare social networks and even the safety of users' property and life. To this end, in response to the problem of abnormal patient profile in healthcare social networks, this paper proposes an overall architecture. The processing architecture of the problem is explained separately from the data collection layer, feature representation layer, algorithm selection layer, and result evaluation layer and different data. The conclusion assessment layer explains how to assess and sign up the outcomes of the several algorithm selection stages. Finally, additional research in this area is anticipated. One of the future scopes of this research is that it will play a crucial role for comparing the effectiveness of supervised machine learning (SVM) and graph techniques for identifying aberrant patient profiles in two healthcare datasets. As the impact of healthcare social networks grows, more bad individuals are concentrating their attacks on them. Unusual

users' harm to medical social networks is one of them, and it poses a severe threat to the data security of healthcare social networks, as well as the protection of users' property and lives. The sources, other features, different algorithms, different data labelling methods, and evaluation criteria are summarized and compared in detail. However, detecting abnormal users in healthcare social networks is an evolving process.

Data Availability

The data are available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] H.-c. Shin, "Abnormal detection based on user feedback for abstracted pedestrian video," in *Proceedings of the 2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1036–1038, Jeju, South Korea, October 2019.
- [2] H. Zhu and R. Gan, "Implementation system of network user abnormal behavior detection algorithm based on data layering," in *Proceedings of the 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 1400–1403, Palladam, India, November 2021.
- [3] M. Wang, L. Qiu, and X. Wang, "GDMS: a geospatial data mining system for abnormal event detection and visualization," in *Proceedings of the 2019 20th IEEE International Conference on Mobile Data Management (MDM)*, pp. 355–356, Hong Kong, China, June 2019.
- [4] K. T. Akhter Md Hasib, I. Chowdhury, S. Sakib et al., "Electronic health record monitoring system and data security using blockchain technology," *Security and Communication Networks*, vol. 2022, Article ID 2366632, 15 pages, 2022.
- [5] K. Tara, A. K. Sarkar, M. A. G. Khan, and J. R. Mou, "Detection of cardiac disorder using MATLAB based graphical user interface (GUI)," in *Proceedings of the 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, pp. 440–443, Dhaka, Bangladesh, December 2017.
- [6] J. Zhang, J. Ou, C. Ding, and W. Shi, "An abnormal behavior detection based on deep learning," in *Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, pp. 61–65, Guangzhou, China, October 2018.
- [7] L. T. Qi, H. P. Huang, P. Wang, and R. C. Wang, "Abnormal item detection based on time window merging for recommender systems," in *Proceedings of the 2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE)*, pp. 252–259, New York, NY, USA, August 2018.
- [8] C. Gandhi, S. S. Ahmad, A. Mehbodniya et al., "Biosensor-assisted method for abdominal syndrome classification using machine learning algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–14, Article ID 4454226, 2022.
- [9] R. Song and F. Liu, "Real-time anomaly traffic monitoring based on dynamic k-NN cumulative-distance abnormal detection algorithm," in *Proceedings of the 2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems*, pp. 187–192, Shenzhen, November 2014.
- [10] S. Hu, X. Jia, and Y. Fu, "Research on abnormal behavior detection of online examination based on image information," in *Proceedings of the 2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 88–91, Hangzhou, China, August 2018.
- [11] Y. Yuan, K. Ji, R. Sun, K. Ma, Z. Chen, and L. Wang, "An integration method of classifiers for abnormal phone detection," in *Proceedings of the 2019 6th International Conference on Behavioral, Economic and Socio-Cultural Computing (BESCC)*, pp. 1–6, Beijing, China, October 2019.
- [12] G. Jiang, H. Chen, C. Ungureanu, and K. Yoshihira, "Multi-resolution abnormal trace detection using varied-length N-grams and automata," in *Proceedings of the Second International Conference on Autonomic Computing (ICAC'05)*, pp. 111–122, Seattle, WA, USA, June 2005.
- [13] B. Meng, W. Andi, X. Jian, and Z. Fucai, "DDOS attack detection system based on analysis of users' behaviors for application layer," in *Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 596–599, Guangzhou, China, July 2017.
- [14] Y. Zhang, C. Liu, and H. Qin, "Artificial immunity-based anomaly detection of network user behavior," in *Proceedings of the 2013 Ninth International Conference on Natural Computation (ICNC)*, pp. 644–648, Shenyang, China, July 2013.
- [15] V. Rakhshan, A. H. Okano, Z. Huang, G. Castelnuovo, and A. F. Baptista, "Biomedical applications of computer vision using artificial intelligence," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9843574, 2 pages, 2022.
- [16] J. Wang and B. Cui, "Network behavior abnormal detection for electricity management system based on long short-term memory," in *Proceedings of the 2018 IEEE International Conference on Energy Internet (ICEI)*, pp. 264–268, Beijing, China, May 2018.
- [17] P. Kanhere and H. K. Khanuja, "A Methodology for outlier detection in audit logs for financial transactions," in *Proceedings of the 2015 International Conference on Computing Communication Control and Automation*, pp. 837–840, Pune, India, February 2015.
- [18] J. Wang, C. Liu, J. Fang, T. Bai, and J. Wang, "Detection of abnormal data of gateway energy meter based on user dynamic behavior mining," in *Proceedings of the 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, pp. 3033–3038, Wuhan, China, October 2020.
- [19] D. Gupta, B. Bajpai, G. Dhiman, M. Soni, S. Gomathi, and D. Mane, "Review of ECG arrhythmia classification using deep neural network," *Materials Today Proceedings*, vol. 2022, pp. 2214–7853, 2021.
- [20] W. M. Shalash, A. A. AlZahrani, and S. H. Al-Nufaii, "Crowd detection management system," in *Proceedings of the 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*, pp. 1–8, Riyadh, Saudi Arabia, May 2019.
- [21] M. Y. Bhanumurthy and K. Anne, "An automated detection and segmentation of tumor in brain MRI using artificial intelligence," in *Proceedings of the 2014 IEEE International*

Retraction

Retracted: Identification and Classification of Depressed Mental State for End-User over Social Media

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] A. Kumar, A. Thakare, M. Bhende, A. K. Sinha, A. C. Alguno, and Y. P. Kumar, "Identification and Classification of Depressed Mental State for End-User over Social Media," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8755922, 10 pages, 2022.

Research Article

Identification and Classification of Depressed Mental State for End-User over Social Media

Akhilesh Kumar ¹, **Anuradha Thakare** ², **Manisha Bhende** ³, **Amit Kumar Sinha** ⁴,
Arnold C. Alguno ⁵ and **Yekula Prasanna Kumar** ⁶

¹Department of Information Technology, Gaya College, Gaya, Bihar, India

²Pimpri Chinchwad College of Engineering, Pune, India

³Marathwada Mitra Mandal's Institute of Technology, Pune, India

⁴Mechanical Engineering Department, Shri Mata Vaishno Devi University Katra, J&K-182320, India

⁵Department of Physics, Mindanao State University-Iligan Institute of Technology, Tibanga Highway, Iligan City, Philippines

⁶Department of Mining Engineering, College of Engineering and Technology, Bule Hora University, 144 Oromia Region, Blue Hora, Ethiopia

Correspondence should be addressed to Yekula Prasanna Kumar; prasannaky@bhu.edu.et

Received 26 January 2022; Revised 5 March 2022; Accepted 11 March 2022; Published 21 April 2022

Academic Editor: Gabriella Casalino

Copyright © 2022 Akhilesh Kumar 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 researching social network data and depression, it is often necessary to manually label depressed and non-depressed users, which is time-consuming and labor-intensive. The aim of this study is that it explores the relationship between social network data and depression. It can also contribute to detecting and identifying depression. Through collecting and analyzing college students' microblog social data, a preliminary screening algorithm for college students' suspected depression microblogs based on depression keywords, and semantic expansion is researched; a comprehensive lexical grammar was proposed. This research provided has a preliminary screening method based on depression keywords and semantic expansion for college students' suspected depression microblogs, with a screening accuracy. This method forms a depression keyword table by constructing the basic keyword table and the semantic expansion based on the word embedding learning model Word2Vec. Finally, the word table is used to calculate the semantic similarity of the tested microblogs and then identify whether it is a suspected depression microblog. The experimental results on the microblog dataset of college students show that the comprehensive lexical method is better than the SDS questionnaire segmentation method and the expert lexical method in terms of screening accuracy; the comprehensive lexical approach can quickly and automatically screen out a tiny proportion of suspected doubts from a large number of college students' microblogs. Depression Weibo can reduce the workload of experts' annotation, improve annotation efficiency, and provide a suitable data processing basis for the subsequent accurate identification (classification problem) of patients with depression.

1. Introduction

Statistics show that there are more than 340 million depression patients in various countries globally and 10 million to 20 million people are suicidal every year [1]. According to the statistics of the Chinese Ministry of Health [2], as of 2012, there were at least 30 million medical records for depression in my country. As a unique group with less social experience, low psychological endurance, and multiple

responsibilities for the future family and society, college students have a significantly higher incidence of depression than other groups [3]. The proposed work plays a key role to explore the relationship between social network data and depression. It also contributes to detecting and identifying depression through collecting and analyzing college student's microblog social data.

A survey [4] shows that the penetration rate of Weibo among college students is as high as 90%. Their personality

characteristics drive the behavior of users using Weibo, and personality characteristics can be used as one of the clues to infer the psychological characteristics of Weibo users. The Chinese name for “microblog” is “Weibo.” Weibo is a microblogging platform similar to Twitter that is claimed to become the most popular in China. Weibo users express their psychological characteristics such as their views, thoughts, and emotions by publishing online texts. Therefore, the behavior characteristics of users using Weibo and the semantic characteristics of Weibo content may be used to characterize the psychological characteristics of Weibo users. Through in-depth mining and analysis of Weibo users’ network texts and logs, the psychological characteristics of users over some time can be obtained, which in turn provides the possibility for analyzing users’ mental health status, including depression.

The present article has been planned into various sections. The present section deals with introducing the concept related to depression and its detection. Section 2 puts light on the discussion of related research work. Section 3 illustrates the analysis of online social behavior of depression group. Section 4 describes the preliminary screening algorithm of college students suspected depression microblog based on depression keywords and semantic expansion. Section 5 describes the Weibo features for depression. The experimental result analysis is described in Section 6, and finally Section 7 portrays the conclusion and possible future works based on the proposed framework.

2. Related Work

Psychological analysis with the help of social network data has gradually become a research hotspot. The current research is mainly in two directions. The first research is to explore the relationship between social network data and depression. For example, researchers at the Missouri University of Science and Technology studied the association between Internet usage patterns and depression among college students [5]. The artificial intelligence (AI) can transform medicines and healthcare in profound ways: automation may be used to assess diagnosis patient information, like electrocardiogram, neurology, or x-ray pictures, in order to diagnose illnesses at a preliminary phase based on limited modifications.

The authors in [6] collected data on users with and without depression from Twitter and used the least-squares method to perform regression analysis on the collected data; they counted the time users posted Twitter and analyzed the two types of users. The time difference in posting tweet and the Pearson correlation coefficient method was used to analyze the degree of correlation between user characteristics and depression. The authors in [7, 8] used Facebook data to detect the depression tendency of adolescents. They identified the depressed user through SM rank algorithm, and the source of this data was obtained by collecting and analyzing college student’s microblog social data a preliminary screening algorithm for college students’ suspected depression microblogs based on depression keywords and semantic expansion.

They analyzed the depression status of college students with the help of psychological diagnostic criteria and Facebook homepage information.

The second research direction is to use social network data to detect and identify depression. For example, the authors in [9, 10] obtained a large amount of Twitter data, used the CES-D scale to get the user’s depression state label, and analyzed the user’s social network behavior data for feature extraction to construct a depression detection model, proving that Twitter data can be used to detect whether a user has depression. The authors in [11] used data from multiple blog platforms such as Yahoo Japan and Livedoor, combined with Japanese-specific language features for feature extraction. They used machine learning methods to build a depression detection model, proving that blog data can be used to detect depressed users. The authors in [12] found that by analyzing the linguistic content and text structure features of blog texts it is possible to identify the emotional state of network users. The authors in [13] found that it is also possible to identify the emotional state of network users by analyzing the short text content of blogs.

The depression detection and recognition model proposed in the second research belongs to the computer field’s classification problem. It is also necessary to manually label the depressed and non-depressed users for training and test set construction. Since manual annotation is time-consuming and labor-intensive, based on the analysis of online social behaviors of depression groups, including network behavior, text semantics (words and topics), etc., this paper proposes a preliminary screening algorithm for suspected depression microblogs.

3. Analysis of Online Social Behavior of Depression Group

3.1. Data Sources

3.1.1. Weibo Data of Depression Groups. Sina Weibo user “Zuofan,” a depression patient, committed suicide on March 18, 2012, after leaving the last note on Weibo, which had a considerable impact. There are more than one million comments under the last words of this Weibo, and it is still increasing. Therefore, there is no lack of many depressed users in the thread expressing negative emotions. This article obtained the Weibo thread of “Zuofan,” searched for the depression Weibo in the line, and used the Depression Weibo to find depression users. To determine depression microblogs, this paper invited a total of 6 experts engaged in psychology-related work in different industries to mark depression microblogs in the microblogs obtained above independently. In the end, the six experts unanimously identified it as a depression microblog.

Furthermore, a user who publishes more than four depression microblogs at different times is identified as a depression user. Finally, the depression group sample dataset and the depression group Weibo sample dataset are formed [14]. The sample dataset consists of 8,081 depressed users and 90,568 microblogs (including 40,035 depression and 50,533 non-depressed microblogs) published by these

users [15]. The obtained microblogs were published in 2014. Until 2018, the acquisition time is January 2019.

3.1.2. Weibo Data of Ordinary College Students. As the college student group concerned in this paper, the microblog data of 53,573 ordinary college students in 8 colleges and universities in the capital were obtained. Considering the research ethics, neither college names nor user names were collected. Instead, a comparative analysis was carried out and used to screen datasets of suspected depression microblogs. This research acquired the “Zuofan” Weibo thread, explored the line for depression Weibo, and utilized the Depression Weibo to locate depression individuals. To identify depression microblogs, this article asked a total of 6 specialists from various sectors who work in psychological to individually mark depression microblogs in the microblogs acquired above. Finally, the six specialists all agreed that it was a depressive microblog [16]. The microblogs were published from November 12, 2018, to December 12, 2018, and acquired in January 2019, with a total of 701,827 microblogs. The microblog data of 8 colleges and universities in the capital are shown in Table 1.

3.2. Analysis of Weibo Network Behavior of Depressed Groups. To study how people’s post microblogs will change under the influence of depression, this paper uses the microblog sample datasets of the depression group and ordinary college students described in Section 3.1 to compare the depression group and regular college students. It should be noted that the timestamps of 130,210 microblog data of 3 colleges and universities in the ordinary college student group are damaged. The time cannot be accurate to “hours;” this part of the microblog data is discarded, the typical college students used in this section. There are 571,617 group microblog samples.

3.2.1. Analysis of Weibo Posting Behavior of Different Groups. The relationship between the frequency and change rate of Weibo posting and time between the depression group and the general college student group is compared, as shown in Figure 1. The horizontal axis is 24 hours a day, and the frequency of microblog posts on the vertical axis refers to the ratio of the number of seats in a certain period to the total number of seats in the day.

It can be seen from Figure 1 that the general trend of the microblog posting frequency of the depression group and the college student group over time is the same. However, the frequency of posting on Weibo and posting change differs between the depressive group and the general college student group. From 21:00 to about 8:00 the next day, the depression group posted more frequently, and the posting frequency was higher throughout the day. On the one hand, it shows that users with depression are more active than ordinary college students at night and in the early morning. On the other hand, it also shows that users with depression are also their most active periods at night and in the early morning. It can be seen that the activity of the depression

TABLE 1: Microblog data of 8 colleges and universities in the capital.

School	Number of posts	User number
1	117 844	9858
2	231 004	16855
3	64822	5656
4	61802	5418
5	120281	5506
6	56987	627
7	13484	2618
8	19893	7302

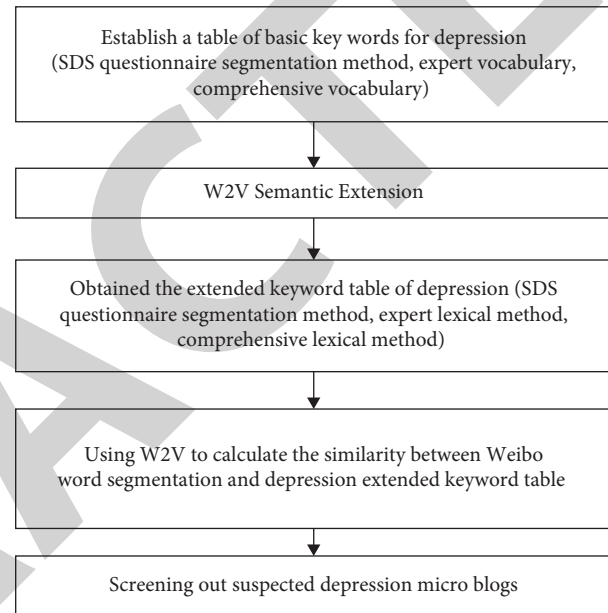


FIGURE 1: Identify algorithm steps.

group shows a prominent phenomenon of “low day and night high.”

As shown in Figure 1, college students’ change rate of postings has apparent fluctuations. Combined with the time distribution of the change rate of postings, the fluctuations occur during getting out of class, lunch, lunch break, and dinner, indicating that the postings of college students are more affected by the actual work and rest. For obvious reasons, the site reflects that the biological clock of ordinary college students is more regulars. On the contrary, the posting change rate of the depression group is relatively flat in the daytime, except for the apparent changes in the early morning and at night, which reflects that the group pays too much attention to themselves, does not want to do things, disregards diet, and even lives passively in life. This phenomenon reflects the characteristics of the depression group accompanied by decreased volitional activity.

3.3. Characteristic Analysis of Group Words for Depression. Studies have shown that words with high frequency in documents, namely, high-frequency words, represent the focus of documents to a certain extent. This paper counts and analyzes the high-frequency words and characteristics of

Weibo posted by depression groups and ordinary college students and understands the focus of the two types of users. As shown in Table 2, this paper lists the top 20 high-frequency words on Weibo posted by the depression group and the college student group, respectively.

“Language Exploration and Word Techniques” (LIWC) are widely used to study the relationship between word analysis and psychological characteristics. This paper uses the Simplified Chinese version of the “Language Exploration and Word Technology” (SC-LIWC) tool [17] to analyze the word characteristics of the depression group as follows:

- (a) Depressed groups use the first-person singular pronoun “self” most frequently in Weibo texts. The above phenomenon shows that the self-awareness or self-perception of the depressive group is too firm, and they are more immersed in their world in social life and are reluctant to connect with other people.
- (b) Depressed groups use the exact word “really” with a high frequency in Weibo texts. This shows that patients with depression are more likely to go to extremes in their views of the world and are more likely to see the world in a “black or white” concept.
- (c) Depressed groups also use negative words “no,” “do not want,” and “but” in Weibo texts more frequently. This phenomenon shows that this group often has negative emotions in social life, and then they are more likely to deny themselves, view the world negatively, and treat life negatively.
- (d) In addition, depressive groups also widely use function words and filler words with no actual meaning in Weibo texts. This phenomenon shows that the group has the problem of imprecise and unclear thinking, and then it reflects the psychological characteristics of the group such as hesitation and contradiction.

In contrast, the words frequently used by ordinary college students in Weibo texts are mostly social process words (“reply,” “repost”), positive emotion words (“hahaha,” “like”), and proper nouns (“Weibo” “Zhu Yilong,” “Chaohua,” “Bu Fan,” “You Changjing,” “Mickey”), and so on. This shows that ordinary college students are more concerned about the outside world and social hotspots, and they are more connected with the outside world and interact with others in social life. These characteristics reflect the group’s upbeat, optimistic, and other psychological factors.

3.4. Topic Analysis of Depression Groups. Studies have shown that the text’s topic content reflects the text’s central idea to a certain extent [18]. Therefore, this paper will extract the topic content of Weibo posted by depression groups and analyze its implicit main idea. We use Linear Discriminant Analysis (LDA) [18] to model the topic of the depression group and the number of issues is set to 20. The Latent Dirichlet Allocation (LDA) method is an autonomous learning method that tries to classify a set of measurements

into a number of separate groups. The most frequent use of LDA is to find a user-specified number of topics from those articles that are in a textual collection. It is also a creative statistic framework that enables unidentified entities to describe why certain sections of the dataset are similar. The results showed that most topics reflected the four aspects of the psycho-emotional disorder, somatic disorder, psychomotor disorder, and psychological disorder in patients with depression. The above content is highly consistent with the four dimensions of depression assessment questionnaires such as SDS. This paper selects five topics for analysis, as follows. The five topics and key words of selected depression groups are shown in Table 2.

- (a) Topic 1. Lonely topic: depressed patients are lonely and feel that they have no one to talk to. A psycho-affective disorder can be classified as depression.
- (b) Topic 2. Choose a topic of death anxiety. This may be related to patients with severe depression whom the disease has tortured for a long time, hoping to be relieved but afraid of death, so they prefer to end their lives suddenly in an accident. This is a typical psychological symptom of patients with severe depression, classified as a psychological disorder of depression.
- (c) Topic 3. I hate my topic. Feeling that oneself is a waste, garbage, useless, and dispensable are typical symptoms of patients with depression, which can be classified as psychomotor disorders.
- (d) Topic 4. The topic of sleep disorders. Insomnia is a specific symptom of depression and can be classified as a somatic disorder of depression.
- (e) Topic 5. A topic that encourages you to persevere. This may be related to the positive side of depression patients in the process of fighting depression and receiving treatment, cheering, and promoting themselves. In addition to topic 5, some cases also reflect positive issues such as depression treatment and social support for depressed patients. However, these topics are not reflected in the evaluation questionnaire. Therefore, it can be regarded as a microblog text different from traditional depression, such as questionnaire evaluation—the difference in detection.

4. Preliminary Screening Algorithm of College Students’ Suspected Depression Microblog Based on Depression Keywords and Semantic Expansion

This algorithm first establishes the basic keyword table of depression and then uses the Word2Vec tool to expand the vocabulary further to obtain the extended keyword table for depression. The flow of the algorithm is shown in Figure 1. To find the best algorithm, this study used three different methods to establish the depression basic keyword table and the corresponding depression extended keyword table for comparison. The current method plays a crucial role to

TABLE 2: 5 topics and key words of selected depression groups.

Topic ID	Topic heading
Topic 1	No friends, lonely, no one to chat, need a place to go out, there are people around
Topic 2	Fear of death, despair, relief, accident, face the terrible future, dare not now
Topic 3	You and others are disgusting, useless, garbage, you only want to disgust
Topic 4	Today I cannot sleep till tomorrow, I cannot eat at night
Topic 5	Live hard, hope, come on, persist in pain, continue to live, see, live

explore the link between social network data and depression. The proposed research is more accurate than existing approaches like the SDS questionnaire segmentation method and the expert lexical method in terms of accuracy. The comprehensive lexical approach can quickly and automatically screen out a tiny proportion of suspected doubts from a large number of college students' microblogs.

4.1. Generation of Essential Keywords for Depression

Method 1. SDS questionnaire segmentation method refers to using the "jieba" text segmentation tool to segment the Depression Self-Rating Scale (SDS) and the segmentation result as the primary keyword table for depression. First, half of the items representing positive emotions in the SDS scale were converted into objects representing negative emotions. Next, all things were divided into words. Words such as subject and modal particles were removed to obtain a vocabulary consisting of 47 words (such as feeling, mood, depression, depression, morning, mood, crying, etc.).

Method 2. The expert vocabulary refers to brainstorming strategies by several experts to conduct brainstorming based on the four dimensions of psycho-emotional disorders, physical disorders, psychomotor disorders, and psycho-behavioral disorders by using research experience and obtaining a primary keyword list for depression. It consists of 238 words (e.g., low mood, depression, depression, and sullenness, insomnia, waking up quickly, nightmares, loneliness, heavy day and night, etc.).

Method 3. Synthetic lexicon, the depression essential keyword list of the synthetic glossary, is the same as the expert lexicon. Then the essential dictionary is expanded according to the method in Section 3.2 to form the depression expanded keyword list.

4.2. *Generation of Word2Vec Semantic Expansion and Depression Expansion Keyword Table.* The method of Word2Vec semantic expansion is as follows: calculate the cosine similarity between each word of the depression basic keyword table in the previous step and all the words in the dictionary, and take the top 10 words with the most significant similarity as the synonym of the word. In this way, each word in the primary keyword table can be screened for ten words close to it, duplicated, and then words unrelated to depression or words that do not meet the experimental requirements, such as English words and codes, are manually removed. Depression Extended Keyword List: the

Python synonyms package was called during the experiment, and synonyms used the word vectors trained by Wiki data-corporus to generate a synonym table.

It should be noted that the expanded keyword list of depression obtained by comprehensive lexical grammar is a word list generated by adding drug names related to depression based on expanded keywords of depression in expert lexical grammar. This study lists the chemical names and trade names of all drugs currently on the market to treat depression, such as agomelatine, amoxapine, Bristol-Myers Squibb, Prozac, phenelzine, Malak, imipramine, trazodone, etc. 74 kinds. After expanding the three essential word lists, 392, 474, and 548 keyword lists for depression were obtained.

4.3. *Similarity Analysis.* Taking the microblog data of ordinary college students in Section 2 as a sample, screening for depression first, we need to perform preprocessing on Weibo, such as deleting modal words, word segmentation, etc., and then perform similarity analysis. The similarity calculation method divides each microblog into A_1, A_2, A_m and each word segment. The word vector K_1, K_2, K_m of the concave extended keyword table calculates the cosine similarity one by one. The cosine similarity value is used to correlate word segmentation and depression. For example, in the microblog word segmentation A_1 , the cosine similarity between A_1 and K_1, K_2, \dots, K_n and other n -words is obtained in turn, and the largest one of the n cosine similarities is taken as the correlation between the word and depression. The screening criteria for suspected depression microblogs are as follows: take the average of the top 3 microblogs with the highest similarity between microblogs A_1, A_2, \dots, A_m and the depression keyword table (after many manual tests, 3 are better). If the average value is higher than 95%, Weibo is considered to be related to depression. If the number of word segments in Weibo is less than 3, calculate the average of all words.

5. Weibo Features for Depression

Since the behavior of Weibo users is mainly affected by two factors, their cognitive level and network environment, the influence of users in this paper consists of two aspects: the initial value of the user's power and the influence of the user's behavior at different times. The initial value of the user's mark is mainly calculated by the number of users' followers and the number of microblogs posted by the user before the topic spreads. The behavioral impact of the user at time t is calculated through the user interaction network at

time t . The user interaction network at time t is divided into users at time t .

Existing research on microblog communication usually uses users as nodes and attention among users and fans as edges to build a user relationship network. However, reducing the impact of redundant node information (such as zombie nodes) on research does not directly rely on attention. Instead, we build a network with fan information but rely on users' forwarding, comments, and likes in different moments of Weibo topics to build a network. To facilitate the description of the problem, here are the following definitions.

Definition 1 (user forwarding relationship network at time t in microblog topic). The user forwarding relationship network is represented by a two-tuple $H_{1t} = (V_{1t}, F_{1t}, X_{1t})$ where H_{1t} represents the user forwarding relationship network at time t in the microblog topic, V_{1t} represents all users participating in the case at time t , F_{1t} represents the edge set in the network, and each edge in the edge set represents the existence of the existence between two users.

Forwarding relationship, X_{1t} , is the forwarding weight of the directed edge.

Definition 2 (the network of user comments and likes in the microblog topic at time t). The network of user comments and likes is represented by a two-tuple $H_{2t} = (V_{2t}, F_{2t}, X_{2t})$ where H_{2t} represents the user comments of the microblog topic at time t , V_{2t} represents all users participating in the case at time t , F_{2t} represents the edge set in the network, each edge in the edge set represents a comment, or like of the relationship between two users, X_{2t} is the comment of the directed edge like weight.

Definition 3 (user influence). The user influence in the microblog topic is represented by a quadruple $J = (H, V, J_t, f)$, where H includes H_{1t} and H_{2t} , U represents the set of users participating in the topic, which indicates the user's influence at time t , and f is the mapping relationship of each element in the quintuple I . The definition of the influence of the user w_i is as follows:

$$J(w_i) = f(J_t(w_i)), \quad w_i \in V. \quad (1)$$

5.1. Swarm Model Definition. In the Swarm model, the agent interacts with other individuals entirely according to its judgment in the environment where the group has no control centre, thereby influencing the whole. Like the agent in the Swarm model, any user in the microblog topic can participate in the functions of posting microblogs, following interactions, commenting, and liking and the interaction between users will also affect the behavior of individuals. According to the similarity between user interaction in Weibo topic and agent communication in Swarm model, the Swarm model is integrated into the Weibo topic user influence evaluation algorithm. The key is to flexibly combine

users to publish Weibo, forward, and comment in agent movement and compliments.

Based on the calculation method proposed by Specter and Klein in the literature [12], according to the research content of this paper, the physics of the Swarm model is redefined here.

Meaning, before giving definitions, first understand the following two concepts.

In the microblog topic, the w_i user's neighborhood user node at time t refers to the user set that has a more significant impact on the user w_i . It is the set of users who forward the user w_i microblog. The neighborhood users of the user w_i are different at different times. It can be obtained from the statistics of the user forwarding relationship network at the moment t of the microblog topic.

User nodes around the w_i user at time t in the microblog topic refer to the set of users that impacts the user w_i , but the impact is not very large. It is a set of users who comment and like the user w_i Weibo. At different times, the users around the user w_i are other and based on the microblog topic. The user at time t is statistically derived from the network of comments and likes.

Definition 4. (Swarm model for Weibo topics)

- (1) In the Swarm model, Y_1 represents the mean vector of an agent pointing to all agents far away from its ranged. In Weibo, the neighborhood user nodes of the user w_i are used for calculation, and the analysis is carried out through the user forwarding relational network, calculated as follows:

$$Y_{1t} = \sum_{y \in V} J_{t-1}(y) * (X_{1t}(yw_i)), \quad (2)$$

$$X_{1t}(yw_i) = \frac{Q_t(yw_i)}{U(yw_i)}$$

Among them, V is the set of users participating in the microblog topic, $J_{t-1}(y)$ is the user influence of user y at the last moment, $(X_{1t}(yw_i))$ is the contribution t of user y to user w_i Weibo forwarding, $Q_t(yw_i)$ is the number of microblogs forwarded by user y to user w_i at time t , and $U(yw_i)$ is the total number of microblogs delivered by user v to all users in the topic at time t .

- (2) Y_2 is the vector in the Swarm model that points to the centre of the simulated world and is represented by the average of the top 20% of user influence in the last iteration of Weibo user influence evaluation. The calculation is as follows:

$$Y_{2t} = \frac{\sum_{y \in \text{Top}} J_{t-1}(y)}{n}. \quad (3)$$

Among them, Top is the top 20% user set at the last moment, and n is the number of Top users.

- (3) Y_3 is the average velocity vector of all agents around a particular agent. It is expressed by the moderate

influence of user nodes around the user w_i on the user. The calculation formula is as follows:

$$Y_{3t} = \sum_{y \in V} J_{t-1}(y) * (X_{2t}(yw_i)), \quad (4)$$

$$X_{2t}(yw_i) = \frac{C_t(yw_i)}{U_{2t}(yw_i)}$$

Among them, $X_{2t}(yw_i)$ is the contribution of user y microblog comments and likes to user w_i at time t , $C_t(yw_i)$ is the number of comments and likes of user y to user w_i at time t , and $U_{2t}(yw_i)$ represents the total number of comments and likes of user y at time t .

- (4) Y_4 is the vector that an agent points to the centre formed by all the agents around it. The calculation formula given in this paper is as follows:

$$Y_{4t} = \frac{\sum_{y \in V} J_{t-1}(y)}{N}. \quad (5)$$

Among them, N is the number of all users in the topic.

- (5) Y_5 is a random unit-length vector, which is not considered in calculating Weibo user behavior influence.

The influence formula of user w_i at time t is expressed as

$$J_t w_i = b1 * Y_{1t} + b2 * Y_{2t} + b3 * Y_{3t} + b4 * Y_{4t}. \quad (6)$$

6. Experimental Result Analysis

6.1. Screening Effectiveness Analysis. Since the expert lexicon and the comprehensive lexical method are only proper nouns without drugs, there is no need to compare the three algorithms in pairs to determine the optimal way. Instead, the performance on the blog can be reached and then compared with the microblogs screened out by the comprehensive lexical medicine proper nouns compared with the expert lexical ones. The comprehensive lexical method is different from existing methods in term of accuracy; the proposed method can quickly and automatically screen out a tiny portion of suspected doubts from large number of college students microblogs. This paper randomly selects 2% of the microblog screening results of the two algorithms of the SDS questionnaire segmentation method and comprehensive verbal method and submits it to experts to determine whether the screening results are valid. Table 3 shows the expert evaluation results of the linguistic segmentation method and comprehensive lexical method of the SDS questionnaire.

The following five microblogs are all screened in the whole language rather than expert speeches based on comparing vocabulary and specialist vocabulary.

- (a) Is there any food that can clear away heat, reduce fire, and detoxify? Huang Lianpo has been too irritable recently.

- (b) Facts have proved that depression does not equal unhappiness.

- (c) Harvard Brain Scientist: if you are not good at studying and feel depressed, a more effective way than taking medicine is. . .

- (d) I am also recently often depressed and uncontrollable.

- (e) Fentanyl that became popular overnight, the black swan has come again, are Renfu Pharmaceuticals and Renhua Pharmaceuticals at risk? On Sunday morning, fentanyl was unexpectedly the most critical market focus for December. Suppose previous protests against fentanyl in the US and Canada failed to attract domestic investors' attention. In that case, the fentanyl released today will undoubtedly be prevalent again and again. From the perspective of the North American market, the ferocity and danger of fentanyl have been elevated to the "opium" war.

It can be seen from the above five items that the comprehensive lexical method adds a certain amount of noise to the expert verb form (such as items (1) and (5), items (2), (3), and (4)) with negative emotion.

To sum up, comparing the three algorithms, the comprehensive vocabulary method performed the best in screening suspected depression microblogs. Compared with the expert lexicon, although a certain amount of noise will be introduced after introducing the drug name dimension, which increases the misrecognition rate, it is possible to screen out suspected depression microblogs missed by the expert lexicon. Furthermore, after expanding the scope of the sample set, we can directly screen out the relevant microblogs that meet the depression treatment drugs in the microblogs.

6.2. Comparison of Impact Coverage. The user depression coverage of top k is compared, and the experimental results are shown in Figure 2 with Table 4. The results show that the SM rank algorithm can effectively find users with high depression rank in Weibo topics. The impact coverage rate of depression analysis is shown in Table 4.

6.3. Depression Comparison in Different Periods. This paper proposes that the SM rank algorithm can calculate the depression of different users in different periods, where the period is set to days. The period can be set to a shorter or longer time (such as two h , six h) according to different datasets in one month; Figures 3 and 4 show the daily depression curves of the top 5 users on other datasets, respectively. It can be seen that the depression of each user is distinct at different times; the daily depression of the top five users in gene editing is shown in Figure 3, where user 1 and user 4 have a significant depression on the first day, but their impact decreases in the next two days, indicating that the user is the topic of the microblog. The initiators of user 2, user 3, and user 5 had more depression on the second day, and they were the leading communicators of the topic. The

TABLE 3: Expert evaluation results of screening algorithm.

Evaluation item	SDS questionnaire segmentation method	Synthetic lexical
The total number of Weibo identified as suspected depression	90771	20731
Randomly select 2% of the number of Weibo	1900	415
Experts determine the number of Weibo with obvious negative emotions	98	275
Proportion (%)	5.088	68.70

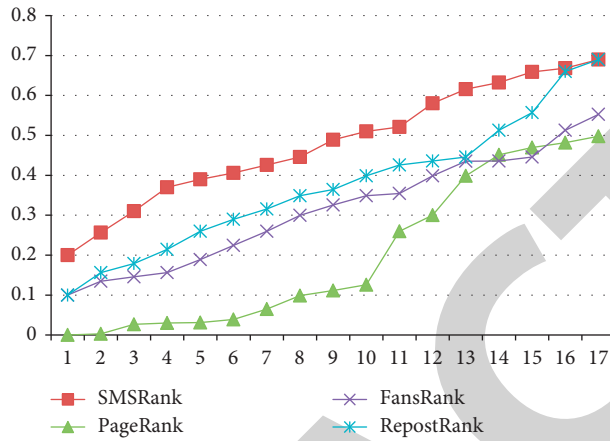


FIGURE 2: Impact coverage rates of depression analysis.

TABLE 4: Impact coverage rates of depression analysis.

Serial number	SMS rank	Page rank	Fans rank	Repost rank
1	0.2	0	0.1	0.1
2	0.2566	0.003	0.1349	0.1562
3	0.31	0.0268	0.1457	0.1789
4	0.3698	0.0298	0.1562	0.2145
5	0.3897	0.0311	0.1889	0.2598
6	0.4059	0.0388	0.2245	0.2897
7	0.4256	0.065	0.2598	0.3157
8	0.4455	0.0989	0.2997	0.3489
9	0.4892	0.1115	0.3257	0.3645
10	0.51	0.1256	0.3489	0.3985
11	0.5213	0.2596	0.3545	0.4258
12	0.5805	0.2999	0.3985	0.4356
13	0.6156	0.3986	0.435	0.4459
14	0.6325	0.4515	0.4356	0.5126
15	0.6589	0.4698	0.4459	0.5569
16	0.6685	0.4817	0.5126	0.6599
17	0.6899	0.4978	0.5526	0.6899

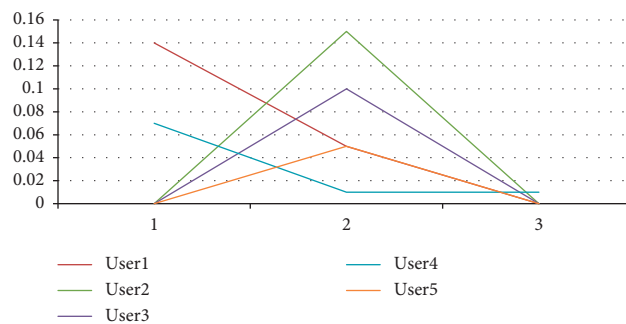


FIGURE 3: Daily depressions of the top 5 users in "gene editing."

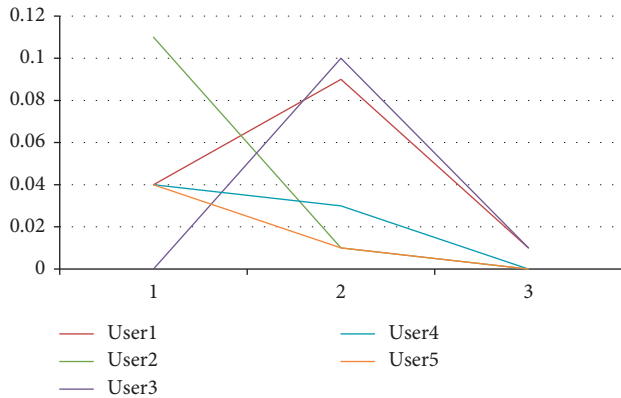


FIGURE 4: Daily depressions of the top 5 users in "food safety."

daily depressions of the top five users in food safety are shown in Figure 4, where user 2, user 1, user 4, and user 5 had a more significant depression on the first day, indicating that they played an essential role in initiating Weibo topics. In contrast, user 1 and user 3 depression level d is in the middle. If it is larger, it means that they are the primary communicators of this Weibo topic; on the last day, when the popularity of the event decreases and is about to die, the depression of all users decreases. The above conclusion is similar to the situation in the entire network.

7. Conclusion

The depression group sample dataset and the depression group microblog sample dataset were constructed, and the microblog data of ordinary college students were collected. Based on this, we analyzed and summarized the characteristics of social network behaviors such as network behavior and text semantics (words and topics) of depression groups publishing Weibo. Based on these characteristics, expert wisdom was synthesized. This research is very useful to explore the relationship between social network data and depression. It also detected and identified depression through collecting and analyzing college students' microblog social data. The Word2Vec tool was used to establish an extended keyword list for depression, which supported the study to propose a preliminary screening algorithm for microblogs suspected of depression. The experimental outcome on the microblog dataset of college students shows that the comprehensive lexical method is better than the SDS questionnaire segmentation method in terms of screening accuracy, and the screening accuracy rate is 65.7%. This paper presents a preliminary screening algorithm for college students' suspected depression microblogs, based on depression keywords and semantic expansion. Although the accuracy rate is not very high, it can quickly screen microblogs with depressive emotions from many college students' microblogs, reduce the workload of experts' labeling, improve the labeling efficiency, and further accurately identify subsequent depression patients (classification problem) to provide a sound basis for data processing. To improve the screening accuracy, the first recognition error and the second recognition error were analyzed and

discussed, respectively. Optimizing the algorithm combined with semantic context analysis will be proposed in the future [19].

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research work was self-funded.

References

- [1] S. Tokuno, G. Tsumatori, S. Shono et al., "Usage of emotion recognition in military health care," in *Proceedings of the 2011 Defense Science Research Conference and Expo (DSR)*, pp. 1–5, San Francisco, CA, USA, July 2011.
- [2] P. Wang, Y. Sun, W. Meng et al., "The method of how to predict weibo users' recovery experience on the weekend based on weibo big data," *IEEE Access*, vol. 8, pp. 194072–194081, 2020.
- [3] E. N. Clark, M. Sejersten, P. Clemmensen, and P. W. Macfarlane, "Evaluating enhancing the acute myocardial infarction criteria in the Glasgow electrocardiogram analysis program by including ST depression," in *Proceedings of the 2010 Computing in Cardiology*, pp. 29–32, Belfast, United Kingdom, September 2010.
- [4] X. Ya-ping, L. Wen-zhong, Y. Meng, X.-F. Xu, D. Hong-mei, and H. Xian-ping, "Notice of Retraction: research on the newly-founded undergraduate college students' psychological health assessment and analysis," in *Proceedings of the 2011 International Conference on Electronics and Optoelectronics*, pp. V1-472–V1-475, Dalian, China, July 2011.
- [5] W. Sen, B. Zhang, T. Li, H. Lv, S. Hu, and Y. Peng, "Evaluation method of user comprehensive influence based on analytic hierarchy process," in *Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA)*, pp. 348–355, Guangzhou, China, May 2021.
- [6] D. Schicchi, G. Pilato, and G. Lo Bosco, "Deep neural attention-based model for the evaluation of Italian sentences complexity," in *Proceedings of the 2020 IEEE 14th International Conference on Semantic Computing (ICSC)*, San Diego, CA, USA, February 2020.
- [7] H. Wu and Q. Hu, "Personal social relations research based on Weibo location check-in data," in *Proceedings of the 2014 22nd International Conference on Geoinformatics*, pp. 1–6, Dallas, TX, USA, November 2014.
- [8] G. Yin, F. Jiang, S. Cheng, X. Li, and X. He, "AUTrust: A practical trust measurement for adjacent users in social networks," in *Proceedings of the 2012 Second International Conference on Cloud and Green Computing*, pp. 360–367, Xiangtan, China, November 2012.
- [9] K. K. Arun, V. R. Navaneeth, S. Prabhu, M. Ramesh kumar, and M. Giriraj, "Experimental investigation of turning process parameter under several cutting conditions for duplex steels for minimization of cutting temperature," *Materials Today Proceedings*, 2022.

Retraction

Retracted: Stress-Relieving Video Game and Its Effects: A POMS Case Study

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] A. Ajmal, H. Aldabbas, R. Amin, S. Ibrar, B. Alouffi, and M. Gheisari, "Stress-Relieving Video Game and Its Effects: A POMS Case Study," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4239536, 11 pages, 2022.

Research Article

Stress-Relieving Video Game and Its Effects: A POMS Case Study

Abdullah Ajmal,¹ Hamza Aldabbas ,² Rashid Amin ,¹ Sundas Ibrar,¹ Bader Alouffi ,³
and Mehdi Gheisari ^{4,5}

¹Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan

²Prince Abdullah Bin Ghazi Faculty of Information and Communication Technology, Al-Balqa Applied University, Al-Salt, Jordan

³Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁴Department of Computer Science and Technology, Harbin Institute of Technology, Shenzhen, China

⁵Department of Computer Science, Islamic Azad University, Tehran, Iran

Correspondence should be addressed to Rashid Amin; rashid4nw@gmail.com

Received 9 December 2021; Revised 1 March 2022; Accepted 11 March 2022; Published 20 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Abdullah Ajmal 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.

Stress is the response or a change in our bodies to environmental factors like challenges or demands that are physical and emotional. The main cause of stress is illnesses and it is gaining more interest, a hot topic for many researchers. Stress can be brought about by a wide range of normal life occasions that are hard to avoid. Stress generally refers to two things: first, the psychological perception of pressure and the body's response to it. On the other hand, it involves multiple systems, from metabolism to muscles to memory. Many methods and tools are being developed to reduce stress in humans. Stress can be a short-term issue or a long-term problem, depending on what changes in your life. The emphasis of this article is to reduce the effects of stress by developing a stress-releasing game and verifying its results through the Profile of Mood States (POMS) and POMS-2 survey. Games are associated with stress levels; hence, parameters like sounds, visuals, and colors associated with reducing stress are used to develop a game for the stress reduction in the players. The survey research aims to determine that the purpose-built game will affect the player's stress level using a reliable psychological survey paper. The survey collected a variety of information from its participants over six months. Different aspects of a person's psychology and reactions are recorded in this scenario by calculating the mean, standard deviation, degree of freedom, zero-error, and probability-value%. The POMS and POMS-2 results are obtained from the custom-built game, and these are found to be effective in reducing stress.

1. Introduction

Profile of Mood States (POMS) is a recognized test first formulated by McNair et al. (1971) [1]. The survey includes 65 questions to represent the mood state of the participant. The POMS is a mechanism for measuring mood. The POMS and POMS-2 test cannot be used in clinical psychology, psychotherapy, and sports science. The POMS and POMS-2 represent the exclusively appropriate mechanism to measure mood-related symptoms properly. In addition, the results point to exact indications related to the place of work-related

well-being. The previous revision points out the psychometric results for the POMS and POMS-2 tests. Its multidimensionality is beneficial for research. POMS and its 2nd edition are ready to provide data to evaluate workplace-related stresses and do not need much effort [2]. POMS classify anger as an emotional situation that ranges from minor impatience feelings until the anger is linked to stimuli from the autonomous nervous system. POMS and POMS-2 define mental disturbance as confusion, insecurity of interest and emotions, depression reduction, and deprived self-image. The fatigue issue clarifies the sense of physical, mental

tiredness, and tension factor or feelings up to the time of worry, anxiety, vigor issue, as differentiated by enthusiasm, temper, and physical energy. COVID-19 has caused significant distress around the globe. It is not limited to adults only, but stress is increasingly affecting children of all age groups. Countries worldwide implemented strict precautions on their citizens in an attempt to control the spread. Most stressing is quarantine. The survey is evaluated and the selection of the psychoanalysis component will depend on the purpose of assessing everyone at some viewpoints with dissatisfaction. There is more than one cause, which includes various circumstances such as losing a loved one, work, and even the weather, which causes a damp and troubling condition known as regular full of feeling clutter disorders. The Profile of Mood States 2nd edition was published in 2012 by the Multi-Health System (MHS) to assess transient feelings and mood among individuals aged 13 years and above. It was designed to cover the age between 13 and 17 years old and adults aged 18 years and above to measure their effective moods and emotions. To enhance their mood, humans will generally discover social embrace or do exercise; these activities can be different for everyone. According to one estimate, 4.4% of the global population suffer from stress and 3.6% from anxiety disorder, whereas depression is more common, at 7.5%, with anxiety being the sixth-highest disorder, and both are estimated to increase. Hence, therapies like cognitive behavioral therapy (CBT) are essential treatments to investigate treatment [3]; those who do not face the disorder longer respond to either pharmacological or psychological interventions, which are frequently encouraged as fast remedies for anxiety and depression. Such approaches provide more options to consider client preferences, treatment accessibility, charge for enlarging devotion, and suitability.

Other methods like digital involvement, including online, computerized apps, and serious games, ensure reducing depression and anxiety symptoms. This study was performed to observe a player's mental stress while playing a stress-releasing game specifically built to release the player's stress levels. According to the National Institute of Mental Health, about 20.9 million or 9.5% population in the USA aged 18 years or older face temperament disorders. In total, 14.8 million adults between the ages of 15 and 44 have depression concerns [4]. Moreover, 1.5% and 3.3 million of the population in the USA are affected by Dysthymic Disorder. At the beginning, the middle ages have disorder for just about 30 years (National Institute of Mental Health) [5]. Stress is usually familiar because of a person's state when too much is predictable, regarding severe pressure or strain, and when they are hardly capable of dealing with excessive and wide exterior demands. Stress is taken as ecological demand as it affects the human organism and if not resisted, it would be unsafe. Casual video games (CVG) have around 200 million players globally. The culture and ages provide a vast diversity in the electronic gaming communities online and offline playing the games on various platforms like PC and handheld cellular phones. These can be further elaborated by the Casual Games Association's 2008 report, which stated that the most opened application on Windows XP was

Microsoft Solitaire. According to the 2006 JWT Intelligence report, the CVG or web games business will grow to \$55 billion by 2009. The world's most stressful countries yearly are shown in Table 1, Table 2, and Table 3.

The idea of the study is to conclude that a video game may improve mood and reduce stress using appropriate measurements. CVG is related to deliberation in its effects on physiological stimulation, suspected confidence, and harm mood. The mere exemption is an augmented helpful influence intended for mood. The outcome supported the prediction that CVG can be a suitable way to reduce stress in the subjects. The results of this study evaluated the positive effects of CVG on stress levels compared to other methods like yoga and meditation. CVG induces an increase of alpha signals in the right frontal part of the brain and reduces it in the left frontal part. The rise of alpha in the left hemisphere is linked to negative behavior like depression, bad mood, and avoidance. At the same time, the increase of alpha in the right hemisphere improves the overall mood and makes a person more engaging. These relations of alpha waves in the right and left parts of the brain are discussed in [6]. The most used scale for measuring positive and negative effects on mood is the Positive and Negative Affect Schedule (PANAS). The PANAS in several languages has exposed first-rate psychometric possessions in the all-purpose clinical samples and population, including women with fibromyalgia, drug addicts, and samples of forensic nature. PANAS has not been examined on depressive and anxiety disorders and other clinical samples. Furthermore, the advent of treatments based on the Internet has expanded into a diverse range of similar scales that can be performed online, such as online surveys and feedback forms. We are conducting this survey research to eliminate stress using the POMS and POMS-2 tests [7]. Our goal is based on this survey research that will work for people's moods in a few minutes. We have to manage anxiety, stress, depression, fatigue, confusion, unhappiness, and vigor. With this test, we can keep ourselves relaxed to a great extent. Our test will significantly reduce stress. The following are the key contributions of this article:

- (i) Evaluation of the POMS and POMS-2 survey is performed to reduce stress among those who are using purpose-built games.
- (ii) The effects of casual gaming on stress are highlighted by performing a survey in which the measurement of Total Mood Disturbance (TMD).
- (iii) The change in stress is evaluated by performing surveys and purpose-built games. The level of stress in several ways is reduced, including calculating mean, standard deviation, zero-scale, and probability-value.

The rest of the article is organized as follows. Section 2 presents the literature review, and Section 3 offers the problem statement. The proposed solution is discussed in Section 4, and the results are included in Sections 5 and 6. We concluded the article in Section 7.

TABLE 1: world's most stressful countries (2022).

Ranks	Countries
1	Nigeria
2	South Africa
3	El Salvador
4	Mongolia
5	Guatemala
6	Colombia
7	Pakistan
8	Jamaica
9	Macedonia
10	Bolivia

TABLE 2: Stress level in 2021.

Percentage% (%)	Countries
59	Greece
58	Philippines
57	Tanzania
55	Albania
55	Iran
55	Sri Lanka
55	USA
53	Uganda
52	Costa Rica
63	Pakistan

TABLE 3: Stress level in 2020.

Percentage% (%)	Countries
53	Iraq
51	Lebanon
51	Peru
55	Albania
50	Egypt
53	USA
47	Tunisia
46	Iran
76	China
74	Pakistan

2. Literature Review

Video games are becoming popular and more interactive day by day [8]. Due to the effects of video games on mood, several neurologists are becoming more interested in them. Because of their nature, video games take part in a generous responsibility in feeling and significantly put into individual behavior and cognitive functions [9]. According to Reinecke, the contributors required by the game decrease the capability of players to think over demanding events. The physiological relaxation provided by video games differs widely between games and is not always an accurate representation of stress [10].

Reinecke's (2009) study presented an idea for the revival of stress regarding four distinct parts, including psychological disinterest, physiological relaxation, mastery experiences, and apparent control [11]. The four major aspects

are used to reduce stress hypothetically to show that video games reduce stress and aid in healing. For this study, the game was intended to determine stress reduction. It builds ongoing investigation by joining physiological events to observe the probable stress-reducing effects of video gaming. According to the "Casual Games Association" (2013), over 200 million people take part in CVGworldwide [12]. The association aims to look at stress and reduce the effects of casual games compared to guided relaxation. PANAS Mood scale test and Dundee Stress State Questionnaire (DSSQ) were used for the dimensions and a stress-inducing task [13]. CVGs were always efficient as a form of stress relief and were extra productive to engage the contributors. CVGs are taken as meditation and to reduce stress. It was completed using suitable physiological measures to resolve changes in physiological stimulation and mood. In previous studies, one unique thing is to highlight different games such as "Bejeweled 2," "Personal Zen," "Plants vs. Zombies," "Bubble Shooter," "Peggle," and "Bookworm Adventures" with their positive impacts on how these games can reduce our stress [14].

Additionally, when people put their all into their treatment plan, their outcomes frequently are improved, such as reduced anxiety and stress-related symptoms [15]. Previous studies use the "Rich get Richer" model to explain that the less worried individuals manage to play World of Warcraft (WoW) to improve their offline lives [16]. Encouraging amazingly pushed players more increment the stretch enduring in their lives by playing dangerously the online amusement inside which they are required protect from their offline inconveniences which for greatly focused people amplifies to some degree than soothing their enduring. It is important to investigate how in-game encounters and exercises shape mental diversion, leading to risky play among more focused individuals [17]. They pursued their participant observation ethnographic inquiry with 28 semistructured interviews, resulting in increased stress. The previous reading recognized more than 900 studies, from which duplicates were detached and leftover abstracts read. Around 21 articles were potentially relevant to the succeeding environment of our theories related to stretch and tricky online gaming. Perceived Stress Scale (PSS) is used in this survey, which defines stress as respondents' superficial intensity of control over their lives, with a higher score representing a high perceived stress [18]. Recently published articles exploring the thought of utilizing a CVG and pointing to the impacts of CVG's on the signs of uneasiness, misery, and stretch were then selected based on the consideration and prohibition criteria [19].

Using cardiac coherence, stress was accurately measured. Cardiac coherence improves as the rhythm of the heart is managed. The stress pilot biofeedback device was used to increase cardiac coherence (CC). Contributors were randomly assigned to appoint in amusement by violent and peaceful game for 20 minutes while cardiac reliability was recorded. Participants' emotional changes were evaluated on different survey papers for psychological analysis of mood, with the POMS test being the most successful. POMS measures six different scales: "Tension, Depression, Anger,

Vigor, Fatigue, and Confusion.” POMS test inside reliability has been reported at about 90 or above. POMS test dependability lies between 68 and 74 for all factors. The previous research presented the point that CVG can be an effective means of reducing stress. In the article, purpose-built stress-reducing video gaming is analyzed using POMS and POMS-2 [20].

3. Problems’ Statement and Proposed Solution

Stress is the body’s standard reaction when changes occur, resulting in physical, enthusiastic, and mental reactions. Moreover, emotional and physical signs typically in the student population, including headaches, fatigue, sadness, anxiety, and the incapability toward control, can intensify by stress. It reveals a 58% increase in stress [21]. Stress causes the body to overflow with hormones that prepare its systems to pass up or tackle danger. Stress can control all aspects of your life, consisting of your emotions, behaviors, thinking ability, and physical health. The fundamental cause of stress in a youngster’s life comes from work, exercise, companions, and family. There are two structures. The principal structure is called acute stress, which lasts for a short period. The causes of stress come with both internal and external effects. It depends on the person, but the internal effects include anxiety, irritability, and nervousness. External effects include breathing faster, sweating, muscles tensing, dry mouth, keener senses, lack of energy, headaches, and sickness. A stressed person tends to have trouble paying attention and have trouble recalling facts. Each day stress builds up in a teenager’s life, which can be negative as this carries into their adulthood. We are conducting this survey research to eliminate stress because of using the POMS and POMS-2 tests. POMS is a psychological test prepared by McNair et al. (1971)[22]. The POMS test is considered to measure fixed psychological traits. The survey consists of 65 statements to explain the feelings of people. The POMS test can be reviewed swiftly due to the ease of the experiment. A long time ago, the POMS test had six different extents of mood swings. The measurement of the POMS test provides a fast, cheap method of reviewing transient, changeable active mood states. POMS test is an ideal mechanism for measuring and monitoring moods. It allows fast estimation of “transient, fluctuating feelings, and enduring effect states.” The Profile of Mood States 2nd Edition–Adult (POMS 2–A) is a self-report assessment of mood that is adaptable to capturing transient and fluctuating feelings or relatively enduring states in adults aged 18+ years. When used in combination with other information, results from the POMS 2–A can help to better understand an individual and guide intervention decisions.

There are several ways to reduce stress by doing physical exercise as it is a great way to alleviate stress. Exercise allows us to get up and go and lowers blood pressure while increasing our ability to deal with stress. Therapies can also reduce the level of stress, including yoga, meditation, and cognitive behavioral therapy (CBT). POMS and POMS-2 are effective tests because of the simplicity and ease of participant understanding. Both are considered as short tools

through good quality psychometric properties that can be performed on healthy and unhealthy people. Our goal is based on this survey research on people’s mood management. Using the POMS and its 2nd Edition of POMS, we have to manage anxiety, stress, depression, fatigue, confusion, unhappiness, and vigor.

Figure 1 shows the methodology of the POMS test where TMD is calculated in a controlled environment and then once after the target exercise has been performed.

4. Methodology

For the survey, a game was developed using Unity Engine Development Tool using C# language. The assets used in-game were designed and developed by the author keeping in mind the colors, textures, and lightning that are typically used for releasing stress. The game does not have any hardness levels as it is not a competitive game and focuses on a stress-free environment throughout the game, but it does have three different scenarios of infinite maps (never-ending maps like subway surfer).

For analysis, POMS and POMS-2 survey was performed; a rating scale of psychological nature was used to calculate different states of mood. It is an authenticated test formulated by McNair in the McNair et al. (1971) [23]. POMS-2 is developed by MHS. The survey consists of 60–65 questions that evaluate a person’s mood. POMS and its 2nd edition require the participant to choose the best possible statement about the state mood they are feeling at that moment. POMS has a pool of 65 different questions divided into six main categories (“tension, depression, anger, vigor, fatigue, and confusion”) and is a “Last Week” or “Right Now” management. Internal dependability was reported as 90 or above. For all variables, test reliability is recorded among 68 and 74 [24].

Figure 2 shows the purpose-built game played for this survey to evaluate the participants’ stress levels before and after the gameplay. POMS and its 2nd edition were used previously for four main research areas: controlled drug evaluation, psychoanalysis, and therapy, different responses to affecting conditions, and simultaneous validity coefficients and other POMS test show a relationship. For each of the 65 questions, there are five options given as “Not at all” for 1 and up to “Extremely” for 5. The participant will select the option on the mood he/she is experiencing at that moment. The TDM score of each test is calculated by adding up the categories of “tension, depression, fatigue, confusion, and anger” and then subtracting the total score of the vigor. The tension, depression, fatigue, confusion, and anger categories are weighted as negative in the TDM score, while the vigor is weighted positively. TMD of the POMS and POMS-2 provides an accurate description of the participant’s mood. The POMS test and its 2nd edition have always been reported independently for the five or six factors, which is also followed in research. POMS and POMS-2 also provide reliable results on the levels of state of anxiety in the participants. Many modifications have been made in POMS and POMS-2 for different environment settings and world regions and provide fewer time-consuming surveys.

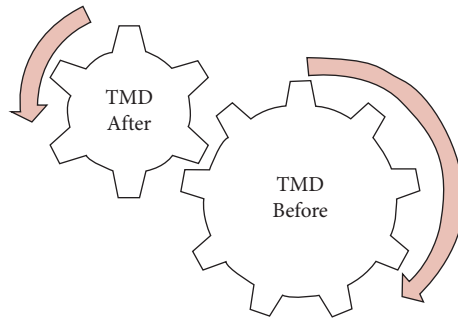


FIGURE 1: POMS strategy.



FIGURE 2: Stress-releasing game.

5. Development and Properties of Game

5.1. Participants. People between 18 and 30 years old and adults are selected for this survey; the participant's mean age was 23 years.

5.2. Administration. For 5 to 10 minutes, the paper-and-pencil self-report measure requires accomplishment users. The psychometric assessment's purpose is to reduce stress. People filled the survey papers before and after accordingly to reduce the level of stress.

5.3. Reliability and Validity. Multiple studies examining several patient groups and the study investigating the structure of the level established significant support for the majority of the POMS seven variables and the subscale fatigue-inertia was established to have exclusive honesty.

Figure 3 defines the psychological effects on the human being, both physically and mentally. These show psychological effects on tension, anger, fatigue, confusion, and vigor are the main building block that causes stress.

5.4. Scoring. It involves respondents pointing out that every point describes their frame of mind in an overload before and after using a 5-point scale ranging from "Not at all" to "Extremely." POMS is obtainable in an immediate scoring format, in which the participant's answers are automatically conveyed through the scoring template.

5.5. Sample Size. The "sample size formula is calculated as a sample size of students," as follows:

$$S = A2 * (c) * \frac{(1 - c)}{a2}, \quad (1)$$

where $A = A$ is the value ("1.96 for 95% confidence level"), P is the percentage in the decimal value (".5 used for sample size needed"), C is the confidence (for example, ".04 = ± 4").

5.6. Levels of POMS. The POMS test has six different extents of mood swings for more than a period. The POMS test includes "Tension or Anxiety, Anger or Confusion, Vigor, Fatigue, and Depression," whose division is shown in Figure 4. The POMS test comprises a 5-point scale ranging from "Not at all" to "Extremely," which is used to estimate their mood states.

It is commonly used for the background of psychotherapy, medicine, and sports science. The POMS and POMS-2 represent the appropriate mechanism to assess mood-related symptoms properly. The POMS and POMS-2 have to cover up independent proportions and somewhat with similar proportions. The study indicates some intersections of the applied instruments concerning psychological well-being.

5.7. Total Mood Disturbance (TMD). Adding up the scores for "Tension, Depression, Anger, Fatigue, and Confusion" minus the score for "Vigor" gives the TMD as follows:

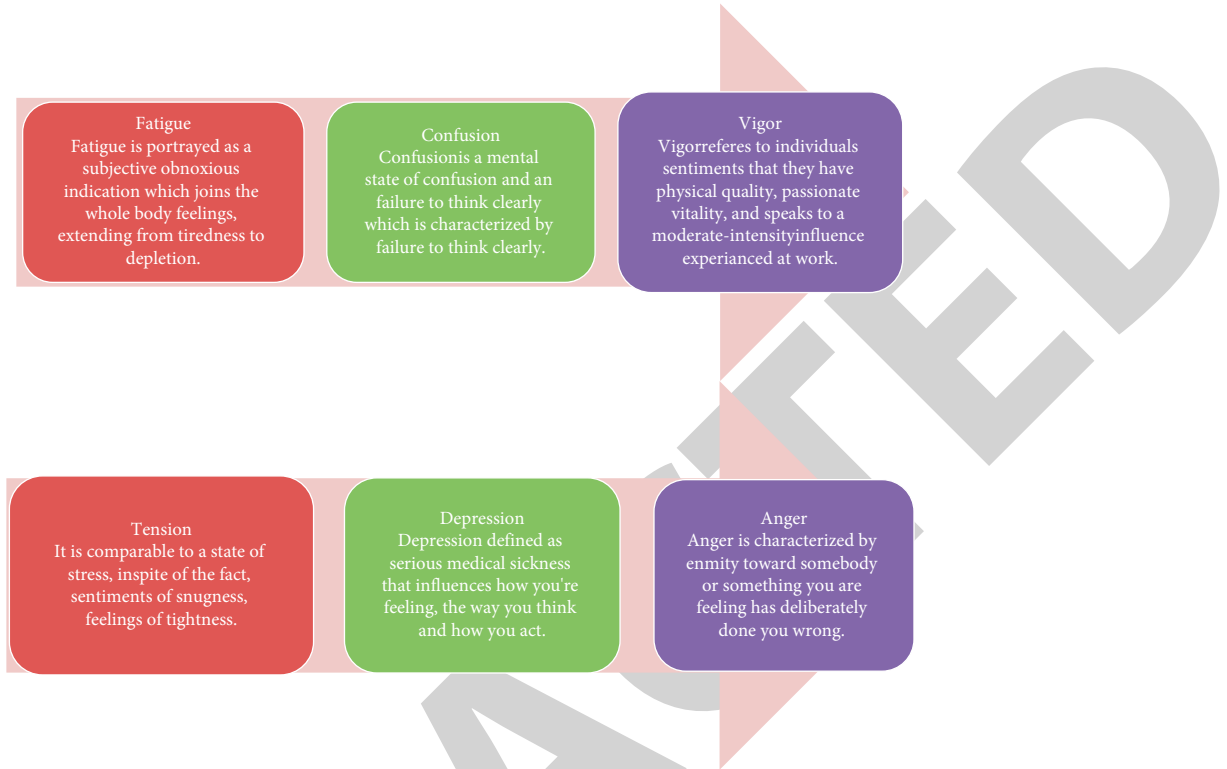


FIGURE 3: Psychological effects.

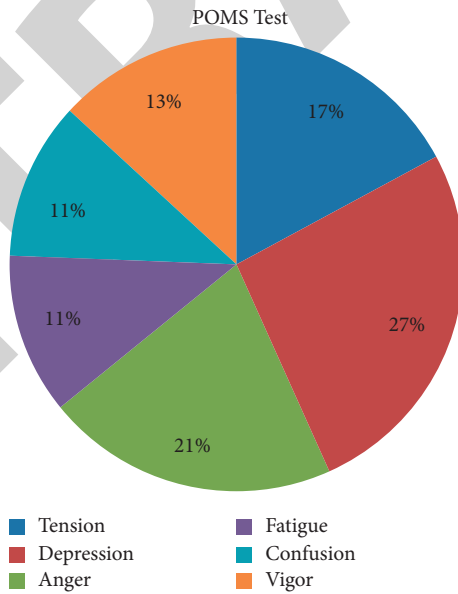


FIGURE 4: Pie chart (TMD).

$$TMD = ("Tension + Depression + Anger + Fatigue + Confusion") - "Vigor". \tag{2}$$

5.8. Score.

- The overall score for “Tension” is calculated by adding up the scores for (Tense, Shaky, On Edge,

Panicky, Relaxed, Uneasy, Restless, Nervous, and Anxious).

- (i) The overall score for “Depression” is calculated by adding up the scores for (Unhappy, Sorry for things

Done, Sad, Blue, Hopeless, Unworthy, Discouraged, Lonely, Miserable, Gloomy, Desperate, Helpless, Worthless, Terrified, and Guilty).

- (ii) The overall score for “Anger” is calculated by adding up the scores for (“Anger, Peeved, Grouchy, Spiteful, Annoyed, Resentful, Bitter, Ready to Fight, Rebellious, Deceived, Furious, and Bad Tempered”).
- (iii) The overall score for “Fatigue” is calculated by adding the scores (Worn Out, Listless, Fatigued, Exhausted, Sluggish, Weary, and Bushed).
- (iv) The overall score for “Confusion” is calculated by adding the score for (Confused, Unable to Concentrate, Muddled, Bewildered, Efficient, Forgetful, and Uncertain about things).
- (v) The overall score for “Vigor” is calculated by adding up the scores for (Lively, Active, Energetic, Cheerful, Alert, Full of Peep, Carefree or Vigorous).

5.9. Interpreting of POMS-2. In this section, the information applies to interpreting all T-scores presented in this report and interpreting changes in T-scores. Responses on the POMS 2–A are combined to produce a TMD score and scores on six mood clusters: Anger-Hostility (AH), Confusion-Bewilderment (CB), Depression-Dejection (DD), Fatigue-Inertia (FI), Tension-Anxiety (TA), and Vigor-Activity (VA). A scaled score is also calculated for Friendliness (F). TMD is determined by summing the Negative Mood State scales and subtracting VA (a Positive Mood State scale). Friendliness is considered separately as a mood state that may influence the severity of mood disturbance through interpersonal functioning. We evaluated probability-value% and zero-error for the POMS-2 test.

6. Results

The survey aims to determine the stress level using a reliable psychological survey paper. The survey collected a variety of information from its participants over six months. A total of 30 participants participated in the survey. The survey asked participants what they felt would be a reasonable explanation for gaming from a menu of items where they could choose more than one answer. This survey allocates the fast evaluation of temporary and changeable emotions as long-term affect states. POMS tests having psychological states can be evaluated rapidly due to the test’s ease, which is another noteworthy evaluation feature. The analysis of the participant based on their sex is shown in Table 4.

Table 5 shows the collected data (before and after) from survey statements regarding overall respondents. It clearly shows that the stress level among the respondents has been reduced after gameplay. The analysis data indicate that the video related to reducing stress can be a helpful instrument for quick concern among the participant and reduce the equal level of stress.

TABLE 4: The state of the art in previous studies.

Years	State of the art
2020	Trier Social Stress Test
2020	Perceived Stress Scale
2018	Ardell Wellness Stress Test
2015	Delphi’s method
2014	Kessler Perceived Stress Scale
2018	Acceptance and Commitment Therapy (ACT)

6.1. Mean. The mean is an arithmetic average of the dataset. The mean is calculated by adding up all the numbers together and dividing by the number of items in the set. The formula for calculating the mean is mentioned in as follows:

$$\mu = \frac{\sum}{N}. \quad (3)$$

$$(1) \text{ Calculating Mean Before} = \frac{2477}{30} = 82, \quad (3a)$$

$$(2) \text{ Calculating Mean After} = \frac{1244}{30} = 41. \quad (3b)$$

Table 6 shows the calculation of mean (Before and after). It offers a clear decline in the TMD in the controlled environment and after playing the game.

Table 7 shows the mean TMD (before and after) and it shows a clear decline in the TMD in the environment and after playing the game.

Figure 5 shows the mean represented graphically to show a clear decline in the TMD in the environment and controlled after playing the game.

6.2. Statistical Analysis of Standard Deviation. The standard deviation (SD) in statistical analysis is a measurement of numbers and how it is spread. It is represented by (Greek letter) σ symbol. The SD is a statistic measuring the dispersion of the dataset comparative to its mean and is considered the square root of variance. The formula for SD is mentioned as follows:

$$D = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}. \quad (4)$$

(1) S = sample standard deviation.

(2) \sum = sum of total numbers.

(3) \bar{x} = sum of total numbers.

(4) \bar{X} = sample mean.

6.3. Statistical Analysis of Population Standard Deviation. The population standard deviation is defined as the measure of the spread (variability) of the score regarding the variable. It symbolizes the sum of the “squared” deviations of the scores from their population mean:

TABLE 5: Data collected from survey statements (before and after).

Category	Before (%)	Category	After (%)
Extremely	60	Extremely	20
Fairly	60	Fairly	30
Moderate	80	Moderate	40
A little	40	A little	70
Not at all	20	Not at all	80

TABLE 6: Calculation of mean (before and after).

S.N	Mean (before)	Mean (after) $\mu = \frac{\sum}{N}$
1	62	44
2	132	55
3	105	86
4	100	66
5	101	98
6	84	47
7	124	49
8	133	90
9	76	49
10	102	19
11	95	67
12	44	33
13	94	83
14	74	14
15	121	28
16	76	21
17	102	6
18	91	39
19	104	26
20	78	11
21	62	22
22	36	9
23	40	20
24	29	26
25	63	32
26	67	19
27	79	55
28	61	14
29	89	70
30	53	46
$= 2,477/30, \mu = 82$		$= 1,244/30, \mu = 41$

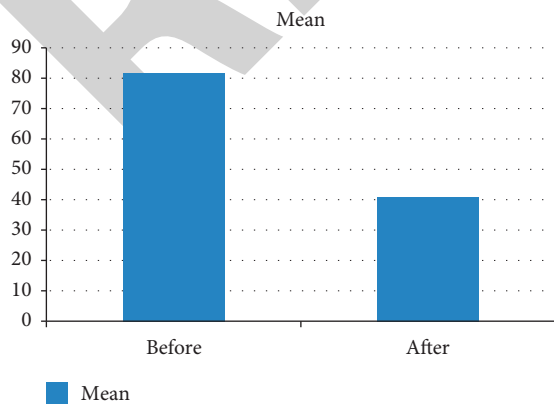


FIGURE 5: Mean graph (before and after).

$$\sigma = \sqrt{N \sum x^2 - \frac{(\sum x)^2}{N^2}}, \quad (4a)$$

where σ is the population standard deviation, N is the size of the population, and X^2 means each value from the population.

$$\text{Standard Deviation (Before)} = 226,$$

$$801 = 6135529, \quad (4b)$$

$$S = \pm 27.71,$$

$$\text{The Population Standard Deviation (After)} = \pm 27.25, \quad (4c)$$

$$\begin{aligned} \text{Standard Deviation} &= \sum x^2 = 71294, \\ &= \sum x^2 = 1547536, \quad (4d) \\ \sigma &= 26.06, \end{aligned}$$

$$\text{Population Standard Deviation (After)} =$$

$$\sigma = \frac{\sqrt{N \sum x^2 - (\sum x)^2}}{N^2} \sigma = 25.63. \quad (4f)$$

Table 8 shows the calculation for standard deviation and population standard deviation (before and after). It shows a clear decline in TMD in the environment and controlled after playing the game.

Table 9 shows the mean, standard deviation, and population standard deviation. It clearly shows that TMD has been reduced after playing the game.

6.4. *Degree of Freedom (Df)*. The degree of freedom is the self-sufficient value that statistical analysis can estimate and indicates the number of dependent values that vary in an analysis lacking breaking any constraints. Typically, the degree of freedom equals the sample size minus the number of parameters while calculating an analysis. The formula for calculating the degree of freedom is mentioned as follows:

$$D_f = N - 1. \sqrt{2}, \quad (5)$$

where N is the number of values in the dataset (sample size), calculations of degree of freedom (before + after), and degree of freedom (before, after):

$$\begin{aligned} D_f &= 30 - 1, \\ D_f &= 29. \end{aligned} \quad (5a)$$

6.5. *Zero-Score*. Zero-score is the standard score that provides an idea of mean data, which is far from the mean. But more theoretically, zero-score is defined as the measure of numerous SDs below and above the population mean. To compute Z-score through other data available like the

TABLE 7: Mean TMD.

Overall POMS changes	Mean
Before	82
After	41

TABLE 8: Calculation of standard deviation (before and after).

N	SD (before)		SD (after)	
	\bar{X}	X^2	\bar{X}	X^2
1	62	3844	44	1936
2	132	17424	55	3025
3	105	11025	86	7396
4	100	10000	66	4356
5	101	10201	98	9604
6	84	7056	47	2209
7	124	15376	49	2401
8	133	17689	90	8100
9	76	5776	49	2401
10	102	10404	19	361
11	95	9025	67	4489
12	44	1936	33	1089
13	94	8836	83	6889
14	74	5476	14	196
15	121	14641	28	784
16	76	5776	21	441
17	102	10404	6	36
18	91	8281	39	1521
19	104	10816	26	676
20	78	6084	11	121
21	62	3844	22	484
22	36	1296	9	81
23	40	1600	20	400
24	29	841	26	676
25	63	3969	32	1024
26	67	4489	19	361
27	79	6241	55	3025
28	61	3721	14	196
29	89	7921	70	4900
30	53	2809	46	2116
	$= (2477) 2 = 6135529$	$= 226,801, S = \pm 27.71 \sigma = \pm 27.25$	$= (1244) 2 = 1547536$	$= 71294 S = \pm 26.06 \sigma = 25.63$

TABLE 9: Calculation of Standard Deviation (Before and after).

Mean	$\mu = \frac{\sum}{N}$	SD	$S D = \sqrt{\sum (x - \bar{x})^2 / n - 1}$	SD population	$\sigma = \sqrt{N \sum x^2 - \frac{(\sum x)^2}{N^2}}$
Before	82	Before	± 27.71	Before	± 27.25
After	41	After	± 26.06	After	± 25.63

observed value, mean of the sample, and SD, equation (6) denoted the measurement for zero-score.

The zero-score formula is as follows:

$$Z = x - \frac{\mu}{\sigma}, \tag{6a}$$

where Z is the standard score and X is the observed value, mean of the sample. The standard deviation of the sample calculations of zero-score for “before” is as follows:

$$Z = 232.52. \tag{6b}$$

Zero-score for “after” is as follows:

$$Z = 46.93. \tag{6c}$$

6.6. *Probability-Value%*. P-value% age is defined as the probability of receiving the outcome at least as concentrated as the experimental results of a statistical hypothesis test. P-value% is observed by a difference that could have occurred presently by random chance. As a result, the lesser the p-value, the larger the statistical inference of the observed difference-value calculation for the Z-table, which is used to calculate the p-value. Suppose results in a p-value used for a Z-score of negative “1.304.”

Z-table needs to take the positive “1.304” into account, which is the upper right tail. Z-table computes the true p-value to obtain a positive p-value for a positive Z-score by multiplying with “0.0968,” “2,” and “0.1936.” This would be a p-value of 19.36%. P-value for a negative Z-score (before) is

TABLE 10: Mean and SD and Population SD (before and after).

POMS overall changes	Mean	Standard deviation (SD)	Population standard deviation (σ)	Z-scores (distribution of X)	P-values (%)
Before	82	± 27.71	± 27.25	232.52	21.44
After	41	± 26.06	± 25.63	46.93	6

$Z = 232.52$." The outcome regarding p -value of 0.01072 and 10.72% for a negative Z -score is 10.72%. For P -value for a positive Z -score (before), the outcome regarding the p -value of .01072 is multiplied by 2 to get 21.44%. P -values (before and after) are in percentage using zero-score. P -value for negative Z -score (after) is $Z = 46.93$, the results in a p -value of .00003 3%; for a negative Z -score is 3%. The results in a p -value of 3% or 3% are multiplied by 2, to get 6%.

Table 10 shows the overall changes using POMS on different measures, including mean, standard deviation, population standard deviation, Z -score, and P -values. It clearly shows the stress level has been reduced to a greater extent. It shows a clear decline in the TMD in the controlled environment and after playing the game.

7. Conclusion

This article presented the POMS and POMS-2 used for psychological testing. The POMS and POMS-2 provide understanding effects of games on a person's stress levels and can provide insight into the development and production of games. Different aspects of a person's psychology and reactions are recorded in this scenario. As our game was developed to be a stress releaser, we evaluated stress levels. It was made clear by the survey that our game reduced a certain amount of stress and improved the overall mood of the participants. In future research, we can evaluate the same game through different surveys and methods, like electroencephalography (EEG), which is used to measure brain signals and how they behave concerning stress. Other methods used can also be probed to evaluate stress and how it behaves in our game.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

Acknowledgments

This research was supported by the Taif University Researchers Supporting Project (no. TURSP-2020/314), Taif University, Taif, Saudi Arabia.

References

- [1] C. A. Anderson and B. J. Bushman, "Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-analytic review of the scientific literature," *Psychological Science*, vol. 12, no. 5, pp. 353–359, 2001.
- [2] E. Andrade, C. Arce, J. Torrado, J. Garrido, C. De Francisco, and I. Arce, "Factor structure and invariance of the POMS mood state questionnaire in Spanish," *Spanish Journal of Psychology*, vol. 13, no. 1, pp. 444–452, 2010.
- [3] R. Busching and B. Krahé, "Charging neutral cues with aggressive meaning through violent video game play," *Societies*, vol. 3, no. 4, pp. 445–456, 2013.
- [4] L. E. Carlson, M. Speca, K. D. Patel, and E. Goodey, "Mindfulness-based stress reduction in relation to quality of life, mood, symptoms of stress, and immune parameters in breast and prostate cancer outpatients," *Psychosomatic Medicine*, vol. 65, no. 4, pp. 571–581, 2003.
- [5] R. Amin, M. A. Al Ghamdi, S. H. Almotiri, and M. Alruily, "Healthcare techniques through deep learning: issues, challenges and opportunities," *IEEE Access*, vol. 9, pp. 98523–98541, 2021.
- [6] T.-J. Chou and C.-C. Ting, "The role of flow experience in cyber-game addiction," *CyberPsychology and Behavior*, vol. 6, no. 6, pp. 663–675, 2003.
- [7] S. Cohen, D. Janicki-Deverts, and G. E. Miller, "Psychological stress and disease," *JAMA*, vol. 298, no. 14, pp. 1685–1687, 2007.
- [8] S. L. Curran, M. A. Andrykowski, and J. L. Studts, "Short form of the profile of mood states (POMS-SF): psychometric information," *Psychological Assessment*, vol. 7, no. 1, pp. 80–83, 1995.
- [9] P. Gonçalves, F. Benevenuto, and M. Cha, "Panas-t: a psychometric scale for measuring sentiments on twitter," 2013, <https://arxiv.org/abs/1308.1857>.
- [10] Y. Liu, Q. Gao, and L. Ma, "Taking micro-breaks at work: effects of watching funny short-form videos on subjective experience, physiological stress, and task performance," in *Proceedings of the International Conference on Human-Computer Interaction*, vol. 12772, pp. 183–200, Málaga, Spain, September 2021.
- [11] K. Hattori, M. Asamoto, M. Otsuji et al., "Quantitative evaluation of stress in Japanese anesthesiology residents based on heart rate variability and psychological testing," *Journal of Clinical Monitoring and Computing*, vol. 34, pp. 371–377, 2020.
- [12] C. V. Russoniello, K. O'Brien, and J. M. Parks, "The effectiveness of casual video games in improving mood and decreasing stress," *Journal of CyberTherapy & Rehabilitation*, vol. 2, pp. 53–66, 2009.
- [13] A. Díaz-García, A. González-Robles, S. Mor et al., "Positive and negative affect schedule (PANAS): psychometric properties of the online Spanish version in a clinical sample with emotional disorders," *BMC Psychiatry*, vol. 20, pp. 1–13, 2020.
- [14] V. Desai, A. Gupta, L. Andersen, B. Ronnestrand, and M. Wong, "Stress-reducing effects of playing a casual video game among undergraduate students," *Trends in Psychology*, vol. 29, pp. 1–17, 2021.
- [15] M. Morfeld, C. Petersen, A. Krüger-Bödeker, S. Von Mackensen, and M. Bullinger, "The assessment of mood at workplace - psychometric analyses of the revised Profile of

Retraction

Retracted: Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] A. Elfaki, A. M. AlQarni, A. A. AlGhamdi, M. A. AlShammari, F. Nasir, and R. Alabdulqader, "Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6530019, 6 pages, 2022.

Research Article

Arabic Validity of the (CARE) Measure for Improving Medical and Mental Health Services

Abdelaziz Elfaki ¹, Amani M AlQarni,² Amal A. AlGhamdi,² Malak A. AlShammari,² Farheen Nasir,³ and Rana Alabdulqader⁴

¹Department of Psychiatry, College of Medicine, Imam Abdurrahman Bin Faisal University, Dammam, Saudi Arabia

²Department of Family and Community Medicine, College of Medicine, Imam Abdurrahman Bin Faisal University, Dammam, Saudi Arabia

³Department of Basic Science, King Saud Bin Abdulaziz University for Health Science, AlAhsa, Saudi Arabia

⁴Medical Intern, King Fahd Hospital of the University, AlKhobar, Saudi Arabia

Correspondence should be addressed to Abdelaziz Elfaki; aelfaki@healthgates.net

Received 3 February 2022; Revised 20 February 2022; Accepted 8 March 2022; Published 18 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Abdelaziz Elfaki 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 of the Study. To construct then examine the internal consistency, reliability and validity of the Arabic consultation and relational empathy (CARE) measurement tool. *Design and Methodology.* The CARE measurement tool was translated into Arabic version and examined on 1245 patients of a primary health care center in the eastern province of Saudi Arabia. *Results.* The majority of the item's responses showed high level of satisfaction. The coronach alpha of our study that examines the Arabic version of the CARE measurement tool 10 items was 0.96 showing an excellent internal consistency. The Kaiser-Meyer-Olkin measure was 0.96 indicating the adequacy of the data for factor analysis and the Bartlett test of sphericity shows ($\chi^2(45) = 8743.126$, $p < 0.001$) indicating the adequacy of the correlation matrix for analysis. *Conclusion.* The Arabic version of CARE Measure seems to be consistent and reliable in the primary health care setting.

1. Introduction

Scientifically physician empathy is considered as an important component of consultations in primary health care and is crucial to the physician patient relationship. Physician empathy is defined clinically as the “socio-emotional competence of a physician to be able to understand the patient’s situation, perspective and feelings, to communicate that understanding and check its accuracy, and to act on that understanding with the patient in a helpful (therapeutic) way” [1]. Patient perceptions of physician empathy has significant relation with patients’ satisfaction with their physicians and clinicians, interpersonal trust, and compliance with physicians’ recommendations as well as an added benefit to physicians in improving their diagnostic accuracy [2, 3]. Being an empathetic treating clinician, enables patients to cope better with their chronic illnesses both on

subjective and objective parameters like reduced anxiety, better quality of life and clinical performance [4].

Improving clinical empathy is of interest to medical educators and health care authorities; finding an effective and convenient self-rated tool to be rated by patients is crucial to allow patient centered measures [5].

Many measurement tools of clinical empathy have been developed e.g. the EUROPEP questionnaire [6]. However, none of the scales are specific to family medicine clinics and suitable to be self-rated by patients without the need to be administered by physicians [7]. The consultation and relational empathy measurement tool (CARE) is broadly used in its ENGLISH version, it provides a specific, reliable, and patient centered estimate of perceived family physicians’ clinical empathy [8].

Mercer and Reynolds have developed the CARE scale. It includes a 10-item self-administered measure involving four

major domains assessed by the patients themselves to report an accurate conception of their experience. These are emotional, concerned with the patient's own experience, ethical, cognitive, and behavioral domains [1].

The CARE tool has been evaluated several times in the literature to prove its psychometric appropriateness and has been applied successfully in many high quality published research [8–10].

This tool, as stated previously, has been widely used in its original English language. No attempt to translate it to Arabic and use it has ever been made before. In our study, we aimed at examining the internal consistency and construct of the Arabic translated version of the CARE measurement tool.

2. Study Approval and Data Collection Methods

This is a cross sectional study that was done on 1245 patients of Imam Abdurrahman bin Faisal University's family medicine polyclinics for the period of June first, 2021–July first, 2021.

All patients during the data collection period were approached by our data collectors. The consenters then asked to fill up an anonymous noncompulsory self-administered electronic survey. The exclusion criteria was non-Arabic speakers and age less than 18 years old.

The research project was approved by the Ethics council at Imam Abdurrahman bin Faisal University (IAU), it complied with the policies and procedures of the institutional review board at IAU and granted approval with an IRB Number (IRB-2020-01-402).

3. The CARE Measure Description

The consultation and relational empathy (CARE) measurement tool was developed by Dr Stewart Mercer and colleagues to predict patient's perception of his/her physician's empathy.

It is a 10-item, with a 5-point Likert scale ranging from excellent [5] to poor [1] for each item. Up to two 'Not Applicable' responses or missing values are allowed, and will be replaced with the remaining items' average score. Responses exceeding two 'Not Applicable' responses or missing values are discarded from the analysis [8].

4. Translation

The (CARE) measure is originally in English language, and the English version was translated and back translated into Arabic by two independent professional translators. These versions were then discussed by a three bilingual consultant clinical psychologist. Pilot study was done on 32 patients to evaluate the understanding of the study participants to the language of the questionnaire, further amendments were done before the final version was released to be used for the actual study.

5. Analysis

The analysis was done using STATA software version 17. For the EFA and CFA the data set was randomly separated into

two equal samples as the EFA was run on the first subsample whilst the CFA was run on the second one.

5.1. Exploratory Factor Analysis. Before running the EFA the data was examined for suitability of analysis using Kaiser–Meyer–Olkin measure and Bartlett test of sphericity. The estimation method of EFA was done using principal factor analysis and no rotation was administered. For choosing the factor number we depended on the retained Kaiser's criterion of Eigenvalue (>1) and scree plot. The significant factor loading was set to be >0.3 .

5.2. Confirmatory Factor Analysis. Maximum likelihood ratio estimation was chosen as the estimation method of the CFA. The root mean squared error of approximation (RMSEA; $RMSEA < 0.05$), Tucker-Lewis index (TLI; $TLI > 0.95$), comparative fit index (CFI; $CFI > 0.95$), and the chi-square test ($p < 0.05$) were chosen to evaluate the model fit. Path modification was considered, based on the Lagrange test results.

6. Results

The total number of involved participants $n = 1245$. There was a wide variation amongst participants as the mean age of the study participants was 32.39 (SD = 13.89; min. = 18; max. = 85) and most of them were in their 1920s ($n = 503$, 40.37%). Male participants ($n = 713$, 57.27%) were more than female participants and majority of this study participants were Saudis ($n = 1046$, 84.02%). In addition, large number of participants were having college degree or higher ($n = 901$, 72.37%); however, many of them were unemployed ($n = 772$, 62.01%) as 93% ($n = 718$) of them were housewives (Table 1).

As seen in (Table 2), the majority of the item's responses showed high level of satisfaction as "letting you telling your story" item had the highest level of satisfaction ($n = 934$, 75%) chose "excellent" as their response. On the other side, the item "making plan action with you" had the lowest satisfaction as around $n = 33$, 2.54% reported "fair" or "poor" responses.

The coronach alpha of the of the CARE Measure tool 10 items was 0.96 showing an excellent internal consistency. The Kaiser-Meyer-Olkin measure was 0.96 indicating the adequacy of the data for factor analysis and the Bartlett test of sphericity shows ($\chi^2(45) = 8743.126$, $p < 0.001$) indicating the adequacy of the correlation matrix for analysis. Table 3 shows the polychoric correlation matrix summary that shows a good correlation and a common variance between the 10 items. The number of factors that were as guided by Kaiser's criterion and the scree plot (Figure 1) was one factor. The factor loading retained from the EFA is summarized in (Table 4) and it is seen there that the factor loading that is (>0.3) was only seen in relation to factor 1.

The CFA that was run on the other half of the data to confirm the EFA analysis results showed an excellent model fitting statistics as following; $\chi^2(35) = 359.464$, $p < 0.001$; $RMSEA = 0.122$ (CI = 0.111 to 0.134); $CFI = 0.959$; $TLI = 0.948$ which indicate the suitability of this run latent

TABLE 1: Characteristics of study participants ($n = 1245$).

	N = 1245	%
Age groups		
<20 years	69	5.54
20–29 years	503	40.37
30–39 years	301	24.16
40–49 years	168	13.48
50–59 years	132	10.59
60–69 years	57	4.57
70–79 years	12	0.96
≥80 years	4	0.32
Gender		
Female	532	42.73
Male	713	57.27
Nationality		
Saudi	1046	84.02
Non-Saudi	199	15.98
Partnership		
Without partner	632	50.76
With partner	613	49.24
Educational level		
High school and lower	344	27.63
College and higher	901	72.37
Employment status		
Employed	473	37.99
Nonemployed	772	62.01
Patient or patient caregiver		
Patient	939	75.42
Caregiver	306	24.58

TABLE 2: The distribution of participant’s responses ($n = 1245$) of CARE questionnaire questions.

Items	Responses											
	Poor		Fair		Good		Very good		Excellent		Not applicable	
	n	%	n	%	n	%	n	%	n	%	n	%
1. Making you feel at ease	5	0.40	17	1.37	85	6.83	214	17.19	912	73.25	12	0.96
2. Letting you tell your “story”	4	0.32	13	1.04	80	6.43	195	15.66	934	75.02	19	1.53
3. Really listening	6	0.48	18	1.45	79	6.35	239	19.20	903	72.53	00	0.00
4. Being interested in you as a whole person	8	0.64	15	1.20	84	6.75	216	17.35	905	72.69	17	1.37
5. Fully understanding your concerns	9	0.72	23	1.85	89	7.15	199	15.98	901	72.37	24	1.93
6. Showing care and compassion	6	0.48	19	1.53	83	6.67	208	16.71	909	73.01	20	1.61
7. Being positive	6	0.48	8	0.64	94	7.55	205	16.47	907	72.85	25	2.01
8. Explaining things clearly	8	0.64	16	1.29	93	7.47	212	17.03	894	71.81	22	1.77
9. Helping you to take control	8	0.64	19	1.53	87	6.99	214	17.19	895	71.89	22	1.77
10. Making a plan of action with you	9	0.72	22	1.77	90	7.23	204	16.39	898	72.13	22	1.77

TABLE 3: The polychoric correlation matrix of the 10 items of CARE measurement tool ($n = 1245$).

	1	2	3	4	5	6	7	8	9	10
1 Making you feel at ease	1.00									
2 Letting you tell your “story”	0.92	1.00								
3 Really listening	0.74	0.70	1.00							
4 Being interested in you as a whole person	0.88	0.93	0.74	1.00						
5 Fully understanding your concerns	0.87	0.93	0.74	0.94	1.00					
6 Showing care and compassion	0.86	0.90	0.72	0.94	0.94	1.00				
7 Being positive	0.90	0.93	0.69	0.96	0.94	0.96	1.00			
8 Explaining things clearly	0.85	0.91	0.70	0.91	0.92	0.95	0.94	1.00		
9 Helping you to take control	0.87	0.89	0.69	0.91	0.90	0.94	0.95	0.95	1.00	
10 Making a plan of action with you	0.88	0.90	0.73	0.91	0.90	0.92	0.93	0.92	0.95	1

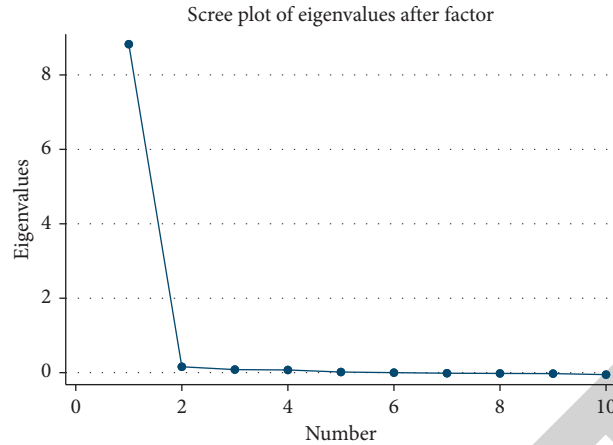


FIGURE 1: scree plot for exploratory factor analysis that was done on the first half of the sample ($n = 623$).

TABLE 4: The factor loadings of the 10 items of the CARE measurement tool that was gain from the exploratory factor analysis that was run on the first half of the sample ($n = 623$).

Number	Items	Factor 1	Factor 2	Factor 3	Uniqueness
1	Making you feel at ease.	0.92	0.18	0.02	0.12
2	Letting you tell your “story”	0.95	0.09	-0.10	0.07
3	Really listening	0.75	0.23	0.13	0.37
4	Being interested in you as a whole person	0.97	0.04	-0.09	0.06
5	Fully understanding your concerns	0.96	0.05	-0.08	0.07
6	Showing care and compassion	0.97	-0.10	-0.03	0.05
7	Being positive	0.98	-0.09	-0.09	0.03
8	Explaining things clearly	0.96	-0.14	0.03	0.06
9	Helping you to take control	0.96	-0.16	0.11	0.04
10	Making a plan of action with you	0.95	-0.04	0.14	0.07

model. The estimation results of the CFA are summarized in Figure 2.

In summary, the Arabic translated version of CARE Measurement tool showed an excellent internal consistency and its internal construct showed only one latent factor that was explained by all the 10 items with an excellent CFA fitting model statistics.

Table 2 summarises the distribution of the participants responses to the 10 items of the CARE measurement tool ($n = 124$).

Extraction method = principal factor analysis; rotation method = unrotated, loading larger than 0.4 is in bold.

7. Discussion

We translated and validated in this study the original English version of the CARE Measurement tool into Arabic language in a primary health care facility (PHC) of a university hospital in Eastern province, Saudi Arabia. Table 1 is for the demographic information showing total 1245 participants were involved which included both patients 75.42% and caregivers 24.58%. The mean age range of the participants was 32.39; male female ratio was 713 : 532. 84.02% of the participants were Saudi nationals while 15.98% belonged to other nationalities. 72.37% of the participants had college or higher-level degrees but 62.01% were unemployed. In Table 2 all the 10 items of the CARE measure show high level of

satisfactory responses by the participants. A study on the check of the validity and reliability of the CARE measure showed high satisfaction responses by the patients where it was not affected by the demographic differences [11].

In the Arabic CARE Measure, the number of “not applicable” was very low in almost all the items. The lowest was 0% in the third item, stated as “really listening” and the highest was 2.01% in item seven, stated as “being positive”. For the rest of the three items, namely no. 8, 9 and 10, “not applicable” was selected 1.77% of the time. Items 8, 9, and 10 were stated as “explaining things clearly”, “helping you to take control” and “making a plan of action with you”, respectively and all were based on the cognitive and behavioral aspects of empathy. One reason for this can be the concept of shared responsibility with the patients rather than sole responsibility of the physicians as indicated in the Swedish version CARE [12] measure. Moreover the number of responses in the poor category was very low as majority had responded in the very good and excellent categories.

The two statistical techniques which were utilized to measure the validity and reliability of the Arabic version of (CARE) measurement tool were Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA which was done on the first half of the sample and as summarized in Tables 2 and 4 shows that the factor loading (>0.3) was only seen in relation to factor 1.

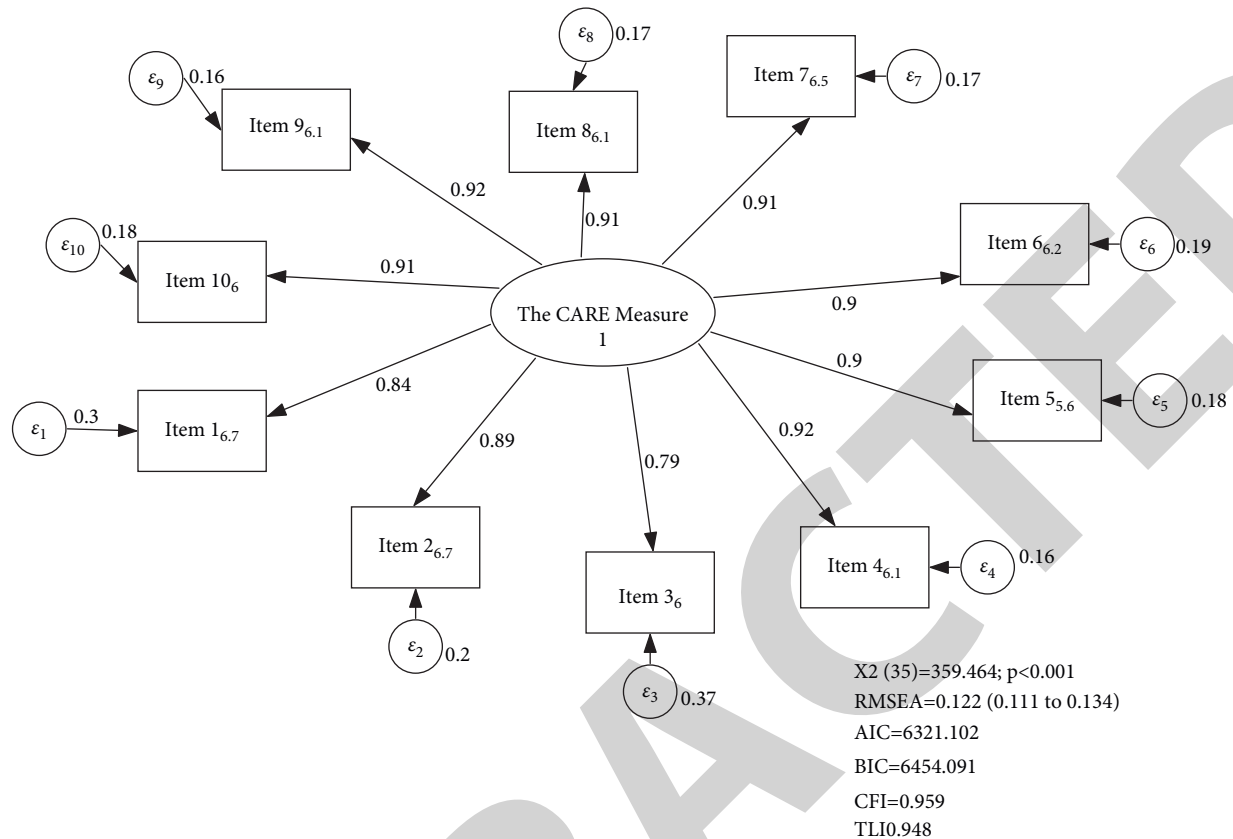


FIGURE 2: Confirmatory factor analysis; estimation method ML that was run on half of the participants ($n=622$). The latent variable is represented with an oval shape whilst each item is represented with a square. The small circles represent the estimation error whilst the arrow represents the regression line.

Cronbach's alpha with a score of 0.96 showed an excellent internal consistency and CFA which was run on the second half of the sample both have confirmed internal structure consistency of the Arabic CARE Measurement tool and consensus with the original, Chinese, and Japanese versions of the CARE Measure [13, 14]. [1–3] As a result, external and internal validity of the Arabic CARE Measurement tool is sustained.

Table 2 shows polychoric correlation matrix of all the 10 items with each other. All the items show strong above average correlation except for item no. 3 whose correlation with the rest of the 9 items lies in the average range minimum $r=0.69$ to maximum $r=0.73$.

One of the study strength points that it was conducted on a large sample, and there was no missing data as all the 1245 responses were available for all the 10 items of the scale. Similar findings were discovered with extremely low “missing values” and “not applicable” data in a research on the patients of secondary care in Scotland [11]. This study indicated CARE measure to be considered as highly relevant with internal and structural reliability plus face and concurrent validity.

In a study CARE measure was also used for validity and reliability check on primary care nurses and the results showed that it has high face and construct validity and internal reliability for the nurses working in the primary care

of the patients [15]. Another study done on the rehabilitation patients in the Southern England showed that CARE fulfilled the strict standards for internal construct validity [16].

8. Limitations

The study was conducted in a single primary health care in eastern province, Saudi Arabia, and is limited to university employees, hospital employees, students, and their families. Therefore, conducting the same study on different primary health care centers, different specialized clinics in different provinces of Saudi Arabia will aid in improving overall patients' satisfaction.

9. Recommendations

It is recommended to develop a CARE measure to be used for the virtual clinics, as after the Pandemic use of the virtual clinic facilities have also become quite common.

10. Conclusion

The original CARE Measurement measure was translated and validated in Arabic, and the study was done in a primary health care context. The Arabic format of the CARE Measurement instrument appears to be consistent, trustworthy, and valid in

Retraction

Retracted: Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] M. Ji, W. Xie, M. Zhao et al., "Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6722321, 15 pages, 2022.

Research Article

Probabilistic Prediction of Nonadherence to Psychiatric Disorder Medication from Mental Health Forum Data: Developing and Validating Bayesian Machine Learning Classifiers

Meng Ji ¹, Wenxiu Xie ², Mengdan Zhao ¹, Xiaobo Qian ³, Chi-Yin Chow ²,
Kam-Yiu Lam ², Jun Yan ⁴, and Tianyong Hao ³

¹School of Languages and Cultures, University of Sydney, Sydney, Australia

²Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong, China

³School of Computer Science, South China Normal University, Guangzhou, Guangdong, China

⁴AI Lab, Yidu Cloud (Beijing) Technology Co. Ltd., Beijing, China

Correspondence should be addressed to Tianyong Hao; haoty@m.scnu.edu.cn

Received 18 January 2022; Revised 16 February 2022; Accepted 19 March 2022; Published 15 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Meng Ji 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. Medication nonadherence represents a major burden on national health systems. According to the World Health Organization, increasing medication adherence may have a greater impact on public health than any improvement in specific medical treatments. More research is needed to better predict populations at risk of medication nonadherence. **Objective.** To develop clinically informative, easy-to-interpret machine learning classifiers to predict people with psychiatric disorders at risk of medication nonadherence based on the syntactic and structural features of written posts on health forums. **Methods.** All data were collected from posts between 2016 and 2021 on mental health forum, administered by Together 4 Change, a long-running not-for-profit organisation based in Oxford, UK. The original social media data were annotated using the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) system. Through applying multiple feature optimisation techniques, we developed a best-performing model using relevance vector machine (RVM) for the probabilistic prediction of medication nonadherence among online mental health forum discussants. **Results.** The best-performing RVM model reached a mean AUC of 0.762, accuracy of 0.763, sensitivity of 0.779, and specificity of 0.742 on the testing dataset. It outperformed competing classifiers with more complex feature sets with statistically significant improvement in sensitivity and specificity, after adjusting the alpha levels with Benjamini–Hochberg correction procedure. **Discussion.** We used the forest plot of multiple logistic regression to explore the association between written post features in the best-performing RVM model and the binary outcome of medication adherence among online post contributors with psychiatric disorders. We found that increased quantities of 3 syntactic complexity features were negatively associated with psychiatric medication adherence: “dobj_stddev” (standard deviation of dependents per direct object of nonpronouns) (OR, 1.486, 95% CI, 1.202–1.838, $P < 0.001$), “cl_av_deps” (dependents per clause) (OR, 1.597, 95% CI, 1.202–2.122, $P, 0.001$), and “VP_T” (verb phrases per T-unit) (OR, 2.23, 95% CI, 1.211–4.104, $P, 0.010$). Finally, we illustrated the clinical use of the classifier with Bayes’ monograph which gives the posterior odds and their 95% CI of positive (nonadherence) versus negative (adherence) cases as predicted by the best-performing classifier. The odds ratio of the posterior probability of positive cases was 3.9, which means that around 10 in every 13 psychiatric patients with a positive result as predicted by our model were following their medication regime. The odds ratio of the posterior probability of true negative cases was 0.4, meaning that around 10 in every 14 psychiatric patients with a negative test result after screening by our classifier were not adhering to their medications. **Conclusion.** Psychiatric medication nonadherence is a large and increasing burden on national health systems. Using Bayesian machine learning techniques and publicly accessible online health forum data, our study illustrates the viability of developing cost-effective, informative decision aids to support the monitoring and prediction of patients at risk of medication nonadherence.

1. Introduction

Medication nonadherence represents a major burden on national health systems around the world. According to the World Health Organization, increasing adherence may have a far greater impact on the health of the population than any improvement in specific medical treatments [1]. The widespread prevalence of medication nonadherence among populations with psychiatric disorders is well known. Systematic reviews show that if nonadherence was defined as taking medication at least 75% of the time, the mean rate of medication nonadherence among people with schizophrenia was 50% [2]. Nonadherence in antidepressants was between 13% and 52% over the course of a lifetime depending on the adherence reporting methods used, and medication nonadherence in bipolar disorder was estimated to be present in 25%–45% of patients with this psychiatric disorder [3–6]. Marcum, Sevick, and Handler summarised 6 representative medication nonadherence phenotypes based on underlying behavioural patterns and barriers to medication adherence at the patient level: (1) lack of understanding and knowledge of the consequences of medication nonadherence; (2) lack of cognitive ability to process and implement complex medication management; (3) lack of vigilance; (4) beliefs that costs outweigh medication benefits; (5) conflicting normative beliefs about medication; and (6) nonbelief in the therapeutic efficacy of medication [7]. These mental nonadherence phenotypes highlighted explanatory factors such as health literacy, education, socioeconomic status, cognitive abilities, and reasoning patterns of nonadhering patients [8–11]. By contrast, researchers studying medication adherence presented evidence which demonstrates the impact of a variety of factors found to be positively associated with medication adherence: health locus of control (belief that health is in one's own control), health literacy, language, cultural backgrounds, and so on [12–22].

Few studies have addressed the two issues in an integral fashion, that is, what kinds of factors may be used to explain and forecast the binary medication adherence outcomes among patients with different psychiatric disorders. Few studies have attempted to establish the interaction and collective impact of these hypothesized external factors for an integrated explanation and predication of patient behaviours. Research shows that explanatory variables of statistical significance are not necessarily of high predictivity [23–27]. This means that factors identified in case-control studies as statistically significant variables do not consequently support the prediction of whether an individual would follow medication regimes. Machine learning is rising as a highly effective analytical technique to solve complex, practical research problems as medication adherence. Machine learning tools can provide cost-effective decision aids complementary to existing diagnostic procedures, quantitative and qualitative methods to clinicians for them for more accuracy, and informed medical decisions to better help their patients [28, 29]. Different from statistics, machine learning does not assume absence of multicollinearity or higher-order interaction among factors. This allows us to leverage existing knowledge across disciplinary boundaries

to develop and interpret machine algorithms which are developed to predict a certain outcome of interest with high precision, accuracy, and practical diagnostic utility. Moreover, the categorization of long texts is still a challenging task due to the high dimensionality of the feature space that causes inefficiency of the machine learning process [30]. Existing research mostly applies keywords extraction to reduce the dimension of the feature space in both long text and image classification tasks [30–33]. Some studies attempted to improve automatic text embeddings representation by applying complex ensemble models to improve the efficiency and performance of machine learning algorithms [28, 34–36]. However, complex ensemble methods are more difficult to generalise. Unlike such previous studies, our study explored a set of high-level syntactic and grammatical features to reduce the dimensionality of feature space and developed a succinct predictive model of high performance with sparse Bayesian models that are more generalisable.

2. Methods

2.1. Research Design. Our study aimed to address the prediction of the binary outcome of medication adherence from a perspective which is distinct from previous studies focusing on patients' background information. That is, instead of gathering information on patients' demographic attributes, health literacy, educational attainment, medication refill records, and so on, we developed machine learning classifiers with quantitative patients posts on a mental health forum which has been in existence for over 20 years. The machine learning classifiers developed can predict the odds of an individual adhering to medication regime or not based on the writing styles (syntactic sophistication and complexity measures) of her/his posts. We interpreted the optimised features included in the best performing Bayesian machine learning model using multiple regression analysis. This helped us to explore and understand the association between patients language style and their health behaviour patterns. The novel Bayesian models developed led to discovering written features of social media data which were positively or negatively associated with medication adherence outcomes among patients with distinct psychiatric disorders.

2.2. Data Collection and Labelling Strategy. Our machine learning models were developed with patient written materials on their medication patterns. Social media data were annotated with high-dimensional features of syntactic complexity and sophistication to predict the odds of medication nonadherence. The source of the data was mental health forum, administered by Together 4 Change, a long running not-for-profit organisation promoting mental health based in Oxford, UK. The forum is structured into five large blocks: mental health experiences, mental health therapies and treatment, and self-help, mental health support forums, recovery, support & help, and local health forums. Within the block of mental health therapies and

treatment, there is a dedicated section on psychiatric drugs and medications forum, which provides the main source of patient discussion data for our study.

We manually screened for posts which satisfied the two following criteria: (1) The content clearly indicates whether the post contributor has been following prescribed psychiatric medications or has interrupted and never resumed his/her medication for various reasons. We eliminated posts which had less relevant contents such as introducing new drugs, peer support, seeking for information for families, friends, or simply expressions of personal emotions without discussing one's medication adherence history. (2) The length of the post contains at least one independent clause (containing at least one subject + verb + object construct). This was to facilitate the annotation of health forum data with English syntactic analysis tools (see Annotation). Posts which contained separate words without clearly logical relations were removed.

The outcome measure, that is, whether a patient is following her/his medication regime, was established through detailed content analysis by human annotators as university researchers. We analysed and labelled posts as negative cases if posts clearly mentioned that the individual was taking medication; posts were classified and tagged as positive cases if the qualitative content analysis showed that the individual interrupted and never resumed medication despite the health consequences experienced. In total, we collected around 500 eligible post items. We divided the full dataset into 70% training dataset and 30% testing dataset. Within the training dataset (352 posts), 172 were from nonadhering patients and 180 were from adhering patients. In the testing dataset (152 posts), 66 were from nonadhering patients and 86 were from adhering patients.

2.3. Annotation of Mental Health Forum Data with Linguistic Features. In selecting natural language annotation tools, we identified the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) as a suitable system. It was developed by Kristopher Kyle at University of Oregon [37–39]. The system provides automatic annotation of English written materials using 4 large sets of linguistic measures: clause complexity, noun phrase complexity, syntactic sophistication, and syntactic complexity. Within each large measure, there are between 9 and 190 features which quantitatively assess the structural and syntactic characteristics of written materials. For example, within syntactic sophistication, there are features which measure the joint probability of a verb and a construction combination (feature tag: average approximate collostructional strength) and lexical diversity (feature tags: main verb lemma type-token ratio, construction type-token ratio, and lemma construction combination type-token ratio). 132 features were developed to measure noun phrase complexity such as standard deviations of dependents per direct object (feature tag: `dobj_NN_stdev`), standard deviations of dependents per passive nominal subject (feature tag: `nsubj_pass_stdev`), and (nonclausal) adverbial modifiers per nominal subject (no pronouns) (feature tag:

`advmod_nsubj_deps_NN_struct`). Originally designed to measure syntactic development in the writing of English learners, the TAASSC system provides a convenience tool to evaluate the lexical, logical, and structural features and patterns of the post data we collected from people with psychiatric disorders. Higher logical, structural, and syntactic complexity is indicative of a more complex reasoning and thinking style. Machine learning classifiers which utilise linguistic complexity features could help us verify and revisit existing knowledge and theories on medication non-adherence, for example, whether nonadherence was due to lack of cognitive ability to process complex medication management procedures or rather conflicting beliefs about the benefits, costs, efficacy, and consequences of medications.

2.4. Feature Optimisation

2.4.1. Classifier Optimisation with Zero Importance Feature Elimination (RZE). Given the high-dimensional nature of the multiple feature sets used in our study, we first applied a Python feature selection tool known as feature-selector to identify and remove features of zero importance (<https://github.com/WillKoehrsen/feature-selector>). This method uses a gradient boosting machine (implemented in the LightGBM library) as the base estimator to learn the importance of the feature. This feature optimisation procedure applies 10-fold validation to reduce variances and biases for estimating feature importance. Moreover, this method leverages an early stopping technique to prevent overfitting to training data. In our study, to balance asymmetric classification errors, for example, classifiers with high sensitivity but low precision, or vice versa, we specified macro F1 as the evaluation metric when training the model to automatically learn feature importance. The resulting optimised feature sets can improve the overall performance of the model in terms of both prediction sensitivity and specificity. Through zero importance feature elimination, none feature was identified as being of zero importance from the syntactic complexity (SCA) feature set (11 in total); 55 features were eliminated due to zero importance from the syntactic sophistication (SS) feature set (135 features in total); 59 features were trimmed for having zero importance from the noun phrase complexity (NPC) feature set (73 in total); next, we removed 6 features from the clause complexity (CC) feature set (12 features in total) which did not improve the model macro F1. Lastly, we applied the macro F1 based importance estimation technique on the combined feature sets of SCA, SS, NPC, and C. 125 features were eliminated as zero importance features.

2.4.2. Recursive Feature Elimination with Support Vector Machine (SVM_RFE). Following classifier optimisation with zero importance feature elimination, we performed recursive feature elimination with support vector machine (SVM_RFE) to further reduce the dimensions of features [40, 41]. An optimised feature number was reached when the minimal cross-validation classification error (CVCE) was

identified through grid search. Figure 1(a) shows that the SVM_RFE reduced the syntactic complexity feature set from 14 to 11 (CVCE = 0.466); Figure 1(b) shows that the syntactic sophistication feature set was reduced from 135 to 5 (CVCE = 0.440); Figure 1(c) shows that the noun phrase complexity feature set was reduced from 73 to 3 (CVCE = 0.415); Figure 1(d) shows the clause complexity feature set was reduced from 26 to 12 (CVCE = 0.406). Lastly, we performed the joint optimisation of all 4 large feature sets which reduced the full feature set (243) to 38 features (CVCE = 0.403). The details of the final optimised features are shown in Table 1.

2.5. Bayesian Machine Learning Classifiers. We used relevance vector machine (RVM) to develop the prediction models on the following considerations. First is model generalisation. RVM is a sparse classifier which has a highly effective mechanism to avoid overfitting issues with relatively small, high-quality datasets like ours. RVM models are known to have good generalisation, which is due to a sparse model dependent only on a small number of kernel functions [42, 43]. Second is model adjustability or flexibility. RVM is a typical Bayesian classifier which produces probabilistic prediction or the posterior probability of a class membership, whereas most supervised machine learning techniques can only return a hard binary prediction which is not very informative in many practical settings. Bayesian models allow more intuitive interpretation of the prediction outcomes. In our study, predictions based on non-Bayesian models can only tell us whether an individual is an adhering patient or not. RVM models, by contrast, assign different probabilities of medication nonadherence to patients based on their unique writing and reasoning styles. This can effectively help us to identify people who were classified as adhering patients (with an assigned probability below a certain threshold level) but were at the same time at high risk of falling out existing medication regimes, based on the structural complexity features of their posts. RVM models can also rate nonadhering patients (with an assigned probability equal to or above a certain threshold level) in terms of their tendency to convert to adhering patients, so that health organisations can accordingly develop personalised interventions to optimise their resource use and patient treatment outcome. Based on these important advantages, we decided to use RVMs to enable more informative decision-making for mental health professionals.

3. Results

Tables 2–5 compare the performance of RVMs with different feature sets on the training and testing datasets. For each feature set, we compared the original TAASSC feature set, the optimised feature set through zero importance feature elimination (RZE), and the optimised feature set through RZE and recursive feature elimination with SVM as base estimator (SVM_RFE). The only exception is syntactic complexity (SCA) in Table 2. There was no feature eliminated in the RZE procedure, so we compared the full feature

set and the optimised feature set using SVM_RFE. As additional classifier performance boosting strategies, we applied 3 feature normalisation techniques on each feature set: min-max normalisation, L_2 normalisation, and Z-score normalisation. The results revealed that there was an overall tendency of performance improvement on both the training and the testing datasets, as we enhanced feature optimisation by using RZE and SVM_RFE successively. This finding was consistent across the 4 large feature sets measuring syntactic complexity, sophistication: syntactic complexity (SCA) (Table 2), syntactic sophistication (SS) (Table 3), noun phrase complexity (NPC) (Table 4), and clause complexity (CC) (Table 5). Feature normalisation had mixed impact on the model overall performance but helped improve asymmetric classification errors, for example, those with imbalanced model sensitivity and specificity. However, none of the optimised feature sets in Tables 2–5 exhibited both an overall good performance and a balanced sensitivity-specificity pair above an acceptable threshold level. As a result, we followed with a combination of the 4 feature sets and optimised features using both RZE and SVM_RFE procedures. We found the best model was the double-optimised feature set ALL RFE 38 with min-max feature normalisation. Table 6 shows that it achieved on the test data an overall AUC of 0.710, accuracy of 0.658, sensitivity of 0.686, and specificity of 0.621.

Apart from the joint optimisation of all features, we also performed pairwise combination of separately optimised feature sets, that is, clause complexity (CC), noun phrase complexity (NPC), syntactic sophistication (SS), and syntactic complexity (SCA), in search of better models. In Table 7, F1 is RVM models which combined the optimised SCA11 and the optimised NPC3 feature sets. We boosted the model with 3 different normalisation techniques, min-max, L_2 , and Z-score, as shown in F2, F3, and F4. The results show that pairwise combination of separately optimised features (sensitivity of 0.628 and specificity of 0.515) balanced the asymmetric classification errors of SCA (11) (sensitivity of 0.244 and specificity of 0.879) and NPC (3) (sensitivity of 0.512 and specificity of 0.621) and improved the overall performance of the model in terms of AUC (F1, 0.631, SCA11, 0.568, NPC3, 0.616) and classification accuracy (F1, 0.579, SCA11, 0.520, NPC3, 0.559). Min-max normalisation boosted the performance of F1 model, as the model AUC and accuracy increased to 0.657 and 0.618, respectively. The same pattern was observed with the combination of two separately optimised models SCA11 and CC12 in model F5. The overall performance of F5 (AUC: 0.568, accuracy: 0.526) improved over both SCA11 (AUC: 0.568, accuracy: 0.520) and CC12 (AUC: 0.559, accuracy: 0.513). Min-max normalisation significantly boosted the performance of F5, as the model AUC and accuracy increased to 0.665 and 0.638, respectively. The 3 high-performing models identified in the pairwise combination of separately optimised feature sets were F6, F14, and F18. We used these competing models for comparison with our best-performing model F46 shown in Table 8.

Table 8 shows the combination of three or four separately optimised feature sets, SCA, NPC, CC, and SS. Overall, the performance of these models improved significantly over

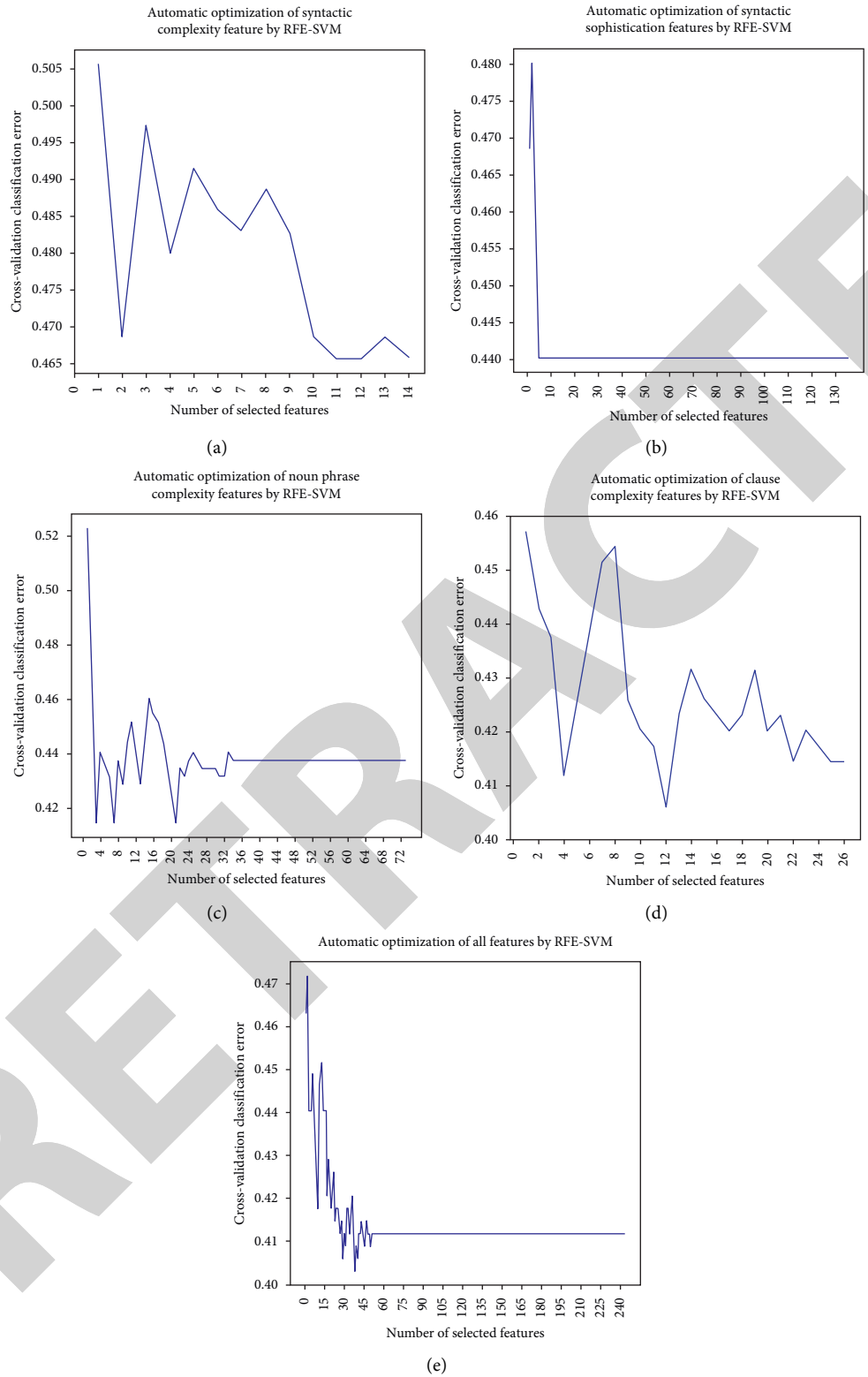


FIGURE 1: Automatic feature selection recursive feature elimination with SVM as the base estimator.

those of individually optimised features (Tables 2–6) and the combination of two optimised feature sets. The two high-performing models that emerged at this stage were models F38 and F42. In the following fine-tuning of model F41

which combined all 4 separately optimised feature sets, SCA11, NPC3, CC12, and SS5, we removed feature “MLT” (mean length of T-unit) from SCA11 and “all_av_approx_collexeme_stddev” (standard deviation of average

TABLE 1: Optimised features through zero importance feature elimination (RZE) and recursive feature elimination with support vector machine (SVM_RFE).

Category	Features	Notation	Feature
Syntactic complexity analyzer	11	SCA11	MLT, MLC, C_S, VP_T, C_T, T_S, CT_T, CP_T, CP_C, CN_T, CN_C
	10	SCA10	MLC, C_S, VP_T, C_T, T_S, CT_T, CP_T, CP_C, CN_T, CN_C
Syntactic sophistication	5	SS5	all_av_construction_freq_stdev, all_av_lemma_freq_stdev, all_av_lemma_freq_type, acad_av_approx_collexeme all_av_approx_collexeme_stdev
	4	SS4	all_av_construction_freq_stdev, all_av_lemma_freq_stdev, all_av_lemma_freq_type, acad_av_approx_collexeme
Noun phrase complexity	3	NPC3	dojb_stdev, advmod_pobj_deps_NN_struct, nsubj_NN_stdev
Clause complexity	12	CC12	aux_per_cl, ccomp_per_cl, nsubjpass_per_cl, prepc_per_cl, nsubj_per_cl, mark_per_cl, ncomp_per_cl, cl_av_deps, cc_per_cl, prep_per_cl, csubj_per_cl, dep_per_cl, CN_C, CT_T, acad_av_construction_freq_stdev, acad_av_lemma_freq_stdev, advmod_pobj_deps_struct, all_av_construction_freq_log, amod_pobj_deps_struct, aux_per_cl, auxpass_per_cl, av_ncomp_deps, av_nominal_deps_NN, cc_per_cl, ccomp_per_cl, conj_and_all_nominal_deps_struct, conj_and_pobj_deps_NN_struct, conj_or_all_nominal_deps_struct, csubj_per_cl, dep_per_cl, det_pobj_deps_NN_struct, dojb_NN_stdev, dojb_stdev, fic_av_delta_p_const_cue_stdev, fic_av_lemma_construction_freq_log, mark_per_cl, nn_all_nominal_deps_struct, nn_dobj_deps_NN_struct, nsubj_NN_stdev, nsubj_per_cl, nsubj_stdev, nsubjpass_per_cl, poss_dobj_deps_struct, prep_per_cl, prep_pobj_deps_NN_struct, prepc_per_cl, prt_per_cl, rcmmod_dobj_deps_NN_struct, tmod_per_cl, xcomp_per_cl
ALL	38	ALL 38	

TABLE 2: Performance of RVM classifiers with syntactic complexity features (no zero importance feature).

RVM	Training data			Testing data			
	AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
SCA full 14	0.468	0.077	0.500	0.566	1.000	0.000	0.361
SCA full 14 with min-max	0.532	0.080	0.514	0.526	0.756	0.227	0.469
SCA full 14 with L_2	0.454	0.089	0.474	0.566	1.000	0.000	0.361
SCA full 14 with Z-score	0.467	0.087	0.513	0.513	0.674	0.303	0.481
SCA RFE 11	0.443	0.047	0.568	0.520	0.244	0.879	0.490
SCA RFE 11 with min-max	0.522	0.107	0.589	0.572	0.674	0.439	0.556
SCA RFE 11 with L_2	0.512	0.106	0.596	0.572	0.663	0.455	0.558
SCA RFE 11 with Z-score	0.464	0.060	0.569	0.559	0.628	0.470	0.549

TABLE 3: Performance of RVM classifiers with syntactic sophistication features (eliminated 55 zero importance features).

RVM	Training data			Testing data			
	AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
SS full 190	0.514	0.050	0.529	0.566	1.000	0.000	0.361
SS full 190 with min-max	0.462	0.076	0.500	0.566	1.000	0.000	0.361
SS full 190 with L_2	0.487	0.046	0.534	0.566	1.000	0.000	0.361
SS full 190 with Z-score	0.522	0.044	0.650	0.632	0.651	0.606	0.628
SS RZF 135	0.518	0.054	0.533	0.566	1.000	0.000	0.361
SS RZF 135 with min-max	0.445	0.103	0.523	0.566	1.000	0.000	0.361
SS RZF 135 with L_2	0.483	0.042	0.541	0.566	1.000	0.000	0.361
SS RZF 135 with Z-score	0.486	0.050	0.628	0.605	0.651	0.546	0.598
SS RFE 5	0.519	0.054	0.533	0.566	1.000	0.000	0.361
SS RFE 5 with min-max	0.469	0.032	0.484	0.566	1.000	0.000	0.361
SS RFE 5 with L_2	0.476	0.049	0.634	0.540	0.767	0.242	0.484
SS RFE 5 with Z-score	0.500	0.043	0.550	0.566	0.919	0.106	0.440

approximate collostructional strength) from SS5. This led to model F45 which contained as few as 29 features (see Table 9 for final features included in model F45). Min-max optimisation further boosted the performance of model F45 on

the testing data, increasing the mean AUC from 0.530 to 0.760 and classification accuracy from 0.566 to 0.763. Normalisation also balanced the sensitivity specificity pair of the model, moderating sensitivity from 1 to 0.779 and

TABLE 4: Performance of RVM classifiers with noun phrase complexity features (eliminated 59 zero importance features).

RVM	Training data			Testing data			
	AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
NPC full 132	0.589	0.049	0.608	0.579	0.581	0.576	0.576
NPC full 132 with min-max	0.553	0.062	0.602	0.566	0.570	0.561	0.563
NPC full 132 with L_2	0.615	0.049	0.603	0.566	0.581	0.546	0.562
NPC full 132 with Z -score	0.555	0.038	0.598	0.559	0.581	0.530	0.555
NPC RZF 73	0.612	0.045	0.614	0.533	0.512	0.561	0.532
NPC RZF 73 with min-max	0.603	0.091	0.614	0.553	0.593	0.500	0.546
NPC RZF 73 with L_2	0.595	0.017	0.611	0.566	0.570	0.561	0.563
NPC RZF 73 with Z -score	0.597	0.026	0.574	0.540	0.547	0.530	0.537
NPC RFE 3	0.634	0.028	0.616	0.559	0.512	0.621	0.559
NPC RFE 3 with min-max	0.643	0.033	0.621	0.586	0.512	0.682	0.586
NPC RFE 3 with L_2	0.660	0.039	0.611	0.605	0.547	0.682	0.605
NPC RFE 3 with Z -score	0.636	0.029	0.618	0.586	0.547	0.636	0.585

TABLE 5: Performance of RVM classifiers with clause complexity features (eliminated 6 zero importance features).

RVM	Training data			Testing data			
	AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
CC full 32	0.594	0.060	0.547	0.540	0.558	0.515	0.536
CC full 32 with min-max	0.563	0.059	0.548	0.540	0.581	0.485	0.533
CC full 32 with L_2	0.605	0.047	0.543	0.526	0.547	0.500	0.522
CC full 32 with Z -score	0.580	0.047	0.532	0.546	0.558	0.530	0.543
CC RZF 26	0.604	0.069	0.552	0.559	0.581	0.530	0.555
CC RZF 26 with min-max	0.566	0.067	0.577	0.540	0.570	0.500	0.534
CC RZF 26 with L_2	0.602	0.051	0.544	0.526	0.547	0.500	0.522
CC RZF 26 with Z -score	0.569	0.047	0.570	0.605	0.651	0.546	0.598
CC RFE 12	0.623	0.048	0.559	0.513	0.512	0.515	0.511
CC RFE 12 with min-max	0.625	0.047	0.585	0.572	0.581	0.561	0.569
CC RFE 12 with L_2	0.624	0.041	0.560	0.540	0.523	0.561	0.538
CC RFE 12 with Z -score	0.590	0.022	0.597	0.586	0.605	0.561	0.582

TABLE 6: Performance of RVM classifiers with all (SCA + SS + NPC + CC) features (eliminated 125 zero importance features).

RVM	Training data			Testing data			
	AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
ALL full 368	0.514	0.050	0.529	0.566	1.000	0.000	0.361
ALL full 368 with min-max	0.543	0.065	0.704	0.665	0.721	0.591	0.657
ALL full 368 with L_2	0.455	0.034	0.481	0.566	1.000	0.000	0.361
ALL full 368 with Z -score	0.554	0.080	0.723	0.684	0.779	0.561	0.671
ALL RZF 243	0.491	0.073	0.519	0.566	1.000	0.000	0.361
ALL RZF 243 with min-max	0.607	0.041	0.675	0.638	0.674	0.591	0.632
ALL RZF 243 with L_2	0.508	0.029	0.504	0.566	1.000	0.000	0.361
ALL RZF 243 with Z -score	0.597	0.097	0.592	0.540	0.547	0.530	0.537
ALL RFE 38	0.488	0.052	0.474	0.566	1.000	0.000	0.361
ALL RFE 38 with min-max	0.698	0.010	0.710	0.658	0.686	0.621	0.653
ALL RFE 38 with L_2	0.517	0.084	0.516	0.566	1.000	0.000	0.361
ALL RFE 38 with Z -score	0.671	0.042	0.655	0.691	0.802	0.546	0.676

increasing specificity from 0 to 0.742. Model F46 thus emerged as the best-performing model in our study. Figure 2 shows the comparison of the AUCs between the best-performing model F46 and other competing high-performing models, F6, F14, F18, F38, F42, and ALL 38 (with min-max).

Tables 10 and 11 show the paired-sample t -tests assessing the significance levels of differences in sensitivity and specificity between the various competitive high-performance classifiers and the best-performing RVM classifier we developed through

the automatic optimisation of four different feature sets and feature refinement. We applied the Benjamini-Hochberg correction procedure to reduce false discovery rates in multiple comparisons. The results show that sensitivity of our best-performing RVM (F46) was significantly higher than those of all the other competitive models with P values equal to or smaller than 0.0059; the specificity of our best-performing RVM was statistically higher than those of most of the other high-performing models, except for F18 ($p = 1$).

TABLE 7: Performance of RVM classifiers with paired feature sets.

Feature set	RVM	Training data			Testing data			
		AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
F1	SCA11 + NPC3	0.597	0.045	0.631	0.579	0.628	0.515	0.572
F2	SCA11 + NPC3 with min-max	0.627	0.02	0.657	0.618	0.64	0.591	0.614
F3	SCA11 + NPC3 with L_2	0.624	0.025	0.624	0.572	0.616	0.515	0.566
F4	SCA11 + NPC3 with Z-score	0.625	0.034	0.64	0.599	0.616	0.576	0.595
F5	SCA11 + CC12	0.449	0.049	0.568	0.526	0.279	0.848	0.504
F6	SCA11 + CC12 with min-max	0.563	0.057	0.665	0.638	0.616	0.667	0.637
F7	SCA11 + CC12 with L_2	0.521	0.116	0.631	0.625	0.686	0.545	0.616
F8	SCA11 + CC12 with Z-score	0.604	0.062	0.631	0.612	0.616	0.606	0.609
F9	SCA11 + SS5	0.523	0.045	0.536	0.566	1	0	0.361
F10	SCA11 + SS5 with min-max	0.544	0.094	0.631	0.586	0.605	0.561	0.581
F11	SCA11 + SS5 with L_2	0.5	0.037	0.503	0.566	1	0	0.361
F12	SCA11 + SS5 with Z-score	0.523	0.097	0.608	0.572	0.57	0.576	0.57
F13	NPC3 + CC12	0.662	0.04	0.628	0.566	0.535	0.606	0.565
F14	NPC3 + CC12 with min-max	0.658	0.029	0.635	0.671	0.674	0.667	0.668
F15	NPC3 + CC12 with L_2	0.676	0.034	0.647	0.605	0.581	0.636	0.604
F16	NPC3 + CC12 with Z-score	0.637	0.036	0.621	0.645	0.651	0.636	0.642
F17	NPC3 + SS5	0.524	0.049	0.536	0.566	1	0	0.361
F18	NPC3 + SS5 with min-max	0.665	0.039	0.687	0.678	0.628	0.742	0.677
F19	NPC3 + SS5 with L_2	0.489	0.058	0.572	0.566	1	0	0.361
F20	NPC3 + SS5 with Z-score	0.637	0.031	0.645	0.625	0.605	0.652	0.624
F21	CC12 + SS5	0.523	0.045	0.536	0.566	1	0	0.361
F22	CC12 + SS5 with min-max	0.594	0.028	0.654	0.625	0.628	0.621	0.622
F23	CC12 + SS5 with L_2	0.478	0.051	0.572	0.566	1	0	0.361
F24	CC12 + SS5 with Z-score	0.561	0.028	0.594	0.579	0.616	0.53	0.573

TABLE 8: Performance of RVM classifiers with multiple feature sets.

No.	RVM	Training data			Testing data			
		AUC mean	SD	AUC	Accuracy	Sensitivity	Specificity	Macro-F1
F25	SCA11 + NPC3 + CC12	0.602	0.065	0.634	0.572	0.616	0.515	0.566
F26	SCA11 + NPC3 + CC12 with min-max	0.653	0.030	0.651	0.638	0.663	0.606	0.634
F27	SCA11 + NPC3 + CC12 with L_2	0.614	0.032	0.609	0.546	0.605	0.470	0.537
F28	SCA11 + NPC3 + CC12 with Z-score	0.646	0.015	0.660	0.625	0.686	0.545	0.616
F29	SCA11 + NPC3 + SS5	0.523	0.045	0.536	0.566	1.000	0.000	0.361
F30	SCA11 + NPC3 + SS5 with min-max	0.662	0.034	0.664	0.625	0.616	0.636	0.623
F31	SCA11 + NPC3 + SS5 with L_2	0.509	0.041	0.504	0.566	1.000	0.000	0.361
F32	SCA11 + NPC3 + SS5 with Z-score	0.634	0.037	0.674	0.625	0.628	0.621	0.622
F33	SCA11 + CC12 + SS5	0.523	0.046	0.536	0.566	1.000	0.000	0.361
F34	SCA11 + CC12 + SS5 with min-max	0.551	0.050	0.572	0.599	0.616	0.576	0.595
F35	SCA11 + CC12 + SS5 with L_2	0.487	0.027	0.491	0.566	1.000	0.000	0.361
F36	SCA11 + CC12 + SS5 with Z-score	0.587	0.059	0.659	0.651	0.663	0.636	0.648
F37	NPC3 + CC12 + SS5	0.523	0.046	0.536	0.566	1.000	0.000	0.361
F38	NPC3 + CC12 + SS5 with min-max	0.685	0.038	0.709	0.704	0.721	0.682	0.700
F39	NPC3 + CC12 + SS5 with L_2	0.477	0.049	0.572	0.566	1.000	0.000	0.361
F40	NPC3 + CC12 + SS5 with Z-score	0.643	0.025	0.682	0.671	0.709	0.621	0.665
F41	SCA11 + NPC3 + CC12 + SS5	0.523	0.046	0.536	0.566	1.000	0.000	0.361
F42	SCA11 + NPC3 + CC12 + SS5 with min-max	0.672	0.039	0.740	0.724	0.756	0.682	0.719
F43	SCA11 + NPC3 + CC12 + SS5 with L_2	0.470	0.046	0.473	0.566	1.000	0.000	0.361
F44	SCA11 + NPC3 + CC12 + SS5 with Z-score	0.651	0.054	0.725	0.711	0.756	0.652	0.704
F45	SCA10 + NPC3 + CC12 + SS4	0.525	0.043	0.530	0.566	1.000	0.000	0.361
F46	SCA10 + NPC3 + CC12 + SS4 with min-max	0.668	0.023	0.762	0.763	0.779	0.742	0.760
F47	SCA10 + NPC3 + CC12 + SS4 with L_2	0.517	0.035	0.510	0.566	1.000	0.000	0.361
F48	SCA10 + NPC3 + CC12 + SS4 with Z-score	0.665	0.034	0.727	0.717	0.733	0.697	0.714

4. Discussion

4.1. Features of Patient Online Posts Associated with Psychiatric Medication Nonadherence. To explore the association

between features in the best-performing model and medication adherence outcome, we performed multiple logistic regression. Predictor variables were the standardized frequencies for each of the 29 features included in the best-

TABLE 9: Features included in the best-performing F45 model.

Feature	Name	Description
Syntactic complexity analyzer (SCA10)	MLC	Mean length of clause
	C_S	Clauses per sentence
	VP_T	Verb phrases per T-unit
	C_T	Clauses per T-unit
	T_S	T-units per sentence
	CT_T	Complex T-unit ratio
	CP_T	Coordinate phrases per T-unit
	CP_C	Coordinate phrases per clause
	CN_T	Complex nominals per T-unit
	CN_C	Complex nominals per clause
Syntactic sophistication (SS4)	all_av_construction_freq_stdev	Average construction frequency-all (standard deviation)
	all_av_lemma_freq_stdev	Average lemma frequency-all (standard deviation)
	acad_av_approx_collexeme_stdev	Average approximate collostructional strength- academic (std.)
	all_av_lemma_freq_type	Average lemma frequency (types only)-all
Noun phrase complexity (NPC 3)	dobj_stdev	Dependents per direct object (standard deviation)
	advmod_pobj_deps_NN_struct	Adverbial modifiers per object of the preposition (no pronouns)
	nsubj_NN_stdev	Dependents per nominal subject (no pronouns, standard deviation)
Clause complexity (CC12)	aux_per_cl	Auxiliary verbs per clause
	ccomp_per_cl	Clausal complements per clause
	nsubjpass_per_cl	Passive nominal subjects per clause
	prep_per_cl	Prepositions per clause
	nsubj_per_cl	Nominal subjects per clause
	mark_per_cl	Subordinating conjunctions per clause
	ncomp_per_cl	Nominal complements per clause
	cl_av_deps	Dependents per clause
	cc_per_cl	Clausal coordinating conjunctions per clause
	csubj_per_cl	Clausal subjects per clause
	dep_per_cl	Undefined dependents per clause
prepc_per_cl	The number of prepositional complements per clause	

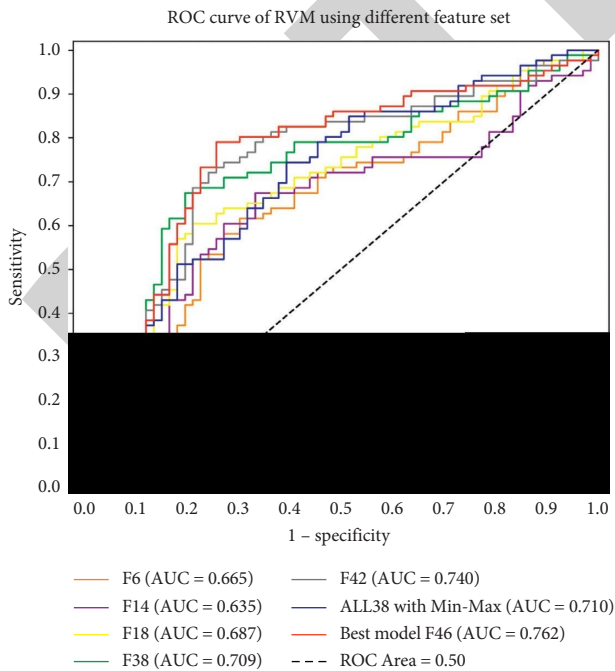


FIGURE 2: AUCs of RVMs on testing data using different feature sets.

performing Bayesian model F46. All analyses were performed in SPSS (26). Continuous predictor variables were standardized using Z-score. We defined statistical significance at 0.001, 0.01, and 0.05 and used a logarithmic scale to display odds ratios and their 95% confidence intervals.

Figure 3 is the forest plot of the multiple logistic regression. Standardized odds ratios and 95% confidence intervals are shown (listed in the right column). The standardized odds ratios (ORs) for each structural feature of patient posts included in the multiple logistic regression model are shown. Standardized odds ratios indicate the effect on an increase of 1 standard deviation (SD) of a feature on the odds of medication nonadherence. In the logistic regression model, medication adherence was the reference class. An odds ratio smaller than 1 indicates that a certain health forum post text feature is more likely to be used by people following psychiatric disorder medication; an odds ratio larger than 1 indicates that a forum post text feature is more likely to be used by people not following medication, and odds ratio of 1 indicates that change in the feature quantity does not affect the medication adherence outcome. The statistical significance of odds ratio is the risk of falsely concluding an association between a feature and medication adherence outcome. We set the statistical significance (P) at a (0.001), b (0.01), and c (0.05). The smaller the P value, the

TABLE 10: Paired-sample *t*-test of the difference in sensitivity between the best-performing model and other models.

No.	Pairs of RVMs	Mean difference	SD	95% confidence interval of difference		<i>P</i> value	Rank	(i/m) Q	Sig.
				Lower	Upper				
1	F46 versus ALL 38 with min-max	0.0930	0.0104	0.0726	0.1134	0.0041	1	0.0083	**
2	F46 versus F6	0.1628	0.0151	0.1332	0.1924	0.0029	2	0.0167	**
3	F46 versus F18	0.1512	0.0145	0.1228	0.1795	0.0030	3	0.0250	**
4	F46 versus F14	0.1047	0.0114	0.0824	0.1269	0.0039	4	0.0333	**
5	F46 versus F38	0.0581	0.0071	0.0442	0.0721	0.0050	5	0.0417	**
6	F46 versus F42	0.0233	0.0031	0.0172	0.0294	0.0059	6	0.0500	**

TABLE 11: Paired-sample *t*-test of the difference in specificity between the best-performing model and other models.

No.	Pairs of RVMs	Mean difference	SD	95% confidence interval of difference		<i>P</i> value	Rank	(i/m) Q	Sig.
				Lower	Upper				
1	F46 versus ALL 38 with min-max	0.1212	0.0115	0.0986	0.1438	0.003	1	0.0083	**
2	F46 versus F6	0.0758	0.0082	0.0596	0.0919	0.0039	2	0.0167	**
3	F46 versus F14	0.0758	0.0082	0.0596	0.0919	0.0039	3	0.0250	**
4	F46 versus F38	0.0606	0.0069	0.0471	0.0741	0.0043	4	0.0333	**
5	F46 versus F42	0.0606	0.0069	0.0471	0.0741	0.0043	5	0.0417	**
6	F46 versus F18	0	0	0	0	1	6	0.0500	

higher the certainty to confirm the feature-outcome association. The results revealed that increased quantities of post structural features like “CT_T” (complex T-unit ratio) (OR, 0.710, 95% CI, 0.541–0.932, *P*, 0.014), “nsubj_per_cl” (nominal subjects per clause) (OR, 0.743, 95% CI, 0.573–0.965, *P*, 0.026), and nsubjpass_per_cl (passive nominal subjects per clause) (OR, 0.763, 95% CI, 0.618–0.943, *P*, 0.012) were associated with greater odds of adherence to psychiatric medication. By contrast, increases in post structural features like “dobj_stdev” (standard deviation of dependents per direct object of nonpronouns) (OR, 1.486, 95% CI, 1.202–1.838, *P* < 0.001), “cl_av_deps” (dependents per clause) (OR, 1.597, 95% CI, 1.202–2.122, *P*, 0.001), and “VP_T” (verb phrases per T-unit) (OR, 2.23, 95% CI, 1.211–4.104, *P*, 0.010) were negatively associated with medication adherence.

4.2. Machine Learning and Statistics Have Different Approaches to Medication Nonadherence Prediction. In many existing studies, the exploration of external explanatory factors and medication adherence outcomes was largely based on the identification of variables which were statistically different between adhering and nonadhering patients. These may include health literacy levels, education, age, culture, and other demographic factors. However, research has shown that statistical significance does not necessarily translate into feature predictivity in machine learning; in other words, variables with high statistical significance do not consequently increase performance of machine learning algorithms. Research shows that addition of statistically significant feature does not improve the performance of machine learning models in health studies.

Our study illustrated a ML-based approach as distinct from existing studies on psychiatric medication adherence prediction.

Our best-performing classifier (F46) included a total of 29 features: 10 syntactic complexity features (SCA10), 4 syntactic sophistication (SS4) features, 3 noun phrase complexity (NPC3) features, and 12 clause complexity (CC12) features. As these features were retained in the model after both zero importance feature elimination (RZF) and recursive feature elimination (SVM_RFE), they were important contributors to the model performance. Among the 29 features, we observed 6 features with statistically significant difference in posts written by adhering versus nonadhering patients (Table 12, Mann–Whitney *U* test).

Three features had statistically higher means in posts from nonadhering patients (NAP) than from adhering patients (AP): “MLC” (mean length of clause) (mean AP, 8.630, mean NAP, 9.369, *P*, 0.003), “acad_av_approx_collexeme_stdev” (standard deviation of average approximate collocation strength in academic English) (mean of AP, 16682.656, mean of NAP, 30199.338, *P*, 0.027), and “dobj_stdev” (standard deviation of dependents per direct object) (mean of AP, 0.722, mean of NAP, 0.888, *P* < 0.001). Logistic regression forest plot shows that there were 15 features positively associated with medication nonadherence: “prepc_per_cl” (the number of prepositional complements per clause), “nsubj_NN_stdev” (dependents per nominal subject (no pronouns, standard deviation)), “prep_per_cl” (prepositions per clause), “ccomp_per_cl” (clausal complements per clause), “C_T” (Clauses per T-unit), “MLC” (mean length of clause), “all_av_construction_freq_stdev” (average construction frequency-all (standard deviation)), “CP_C” (coordinate phrases per

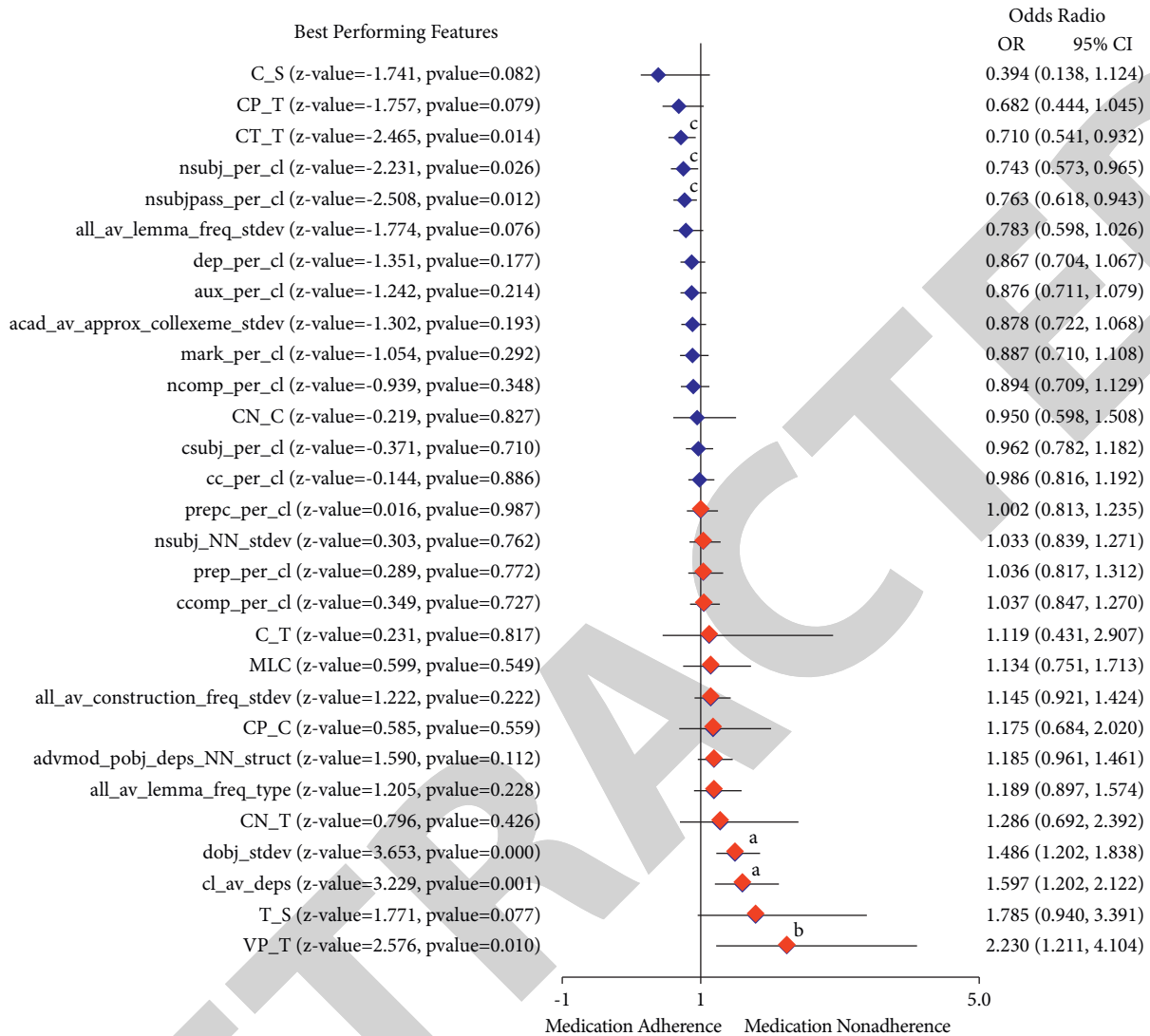


FIGURE 3: Forest plot of logistic regression. ^a $P < 0.001$. ^b $P < 0.01$. ^c $P < 0.05$. Multiple logistic regression is to predict medication nonadherence.

clause), “advmod_pobj_deps_NN_struct” (adverbial modifiers per object of the preposition (no pronouns)), “all_av_lemma_freq_type” (average lemma frequency-all (standard deviation)), “CN_T” (complex nominals per T-unit), “dobj_stddev” (dependents per direct object (standard deviation)), “cl_av_deps” (dependents per clause), “T_S” (T-units per sentence), and “VP_T” (verb phrases per T-unit). Among the 15 features that are highly predictive of medication nonadherence, only 2 had statistically higher means in posts written by nonadhering patients, whereas the remaining 13 features were statistically insignificant.

Three features had statistically higher means in posts from adhering patients than from nonadhering patients: “nsubjpass_per_cl” (nominal subjects per clause) (mean of AP, 0.043, mean of NAP, 0.025, P , 0.025) and “CT_T” (clauses per T-unit) (mean of AP, 0.471, mean of NAP, 0.426, P , 0.046), “nsubj_per_cl” (nominal subjects per clause) (mean of AP, 0.711, mean of NAP, 0.679, P , 0.031). Logistic regression forest plot shows there were 14 features positively

associated with medication adherence: “C_S” (clauses per sentence), “CP_T” (coordinate phrases per T-unit), “CT_T” (complex T-unit ratio), “nsubj_per_cl” (nominal subjects per clause), “nsubjpass_per_cl” (passive nominal subjects per clause), “all_av_lemma_freq_stddev” (average lemma frequency (standard deviation)), “dep_per_cl” (undefined dependents per clause), “aux_per_cl” (auxiliary verbs per clause), “acad_av_approx_collexeme_stddev” (average approximate collostructional strength of academic English (std.)), “mark_per_cl” (subordinating conjunctions per clause), “ncomp_per_cl” (nominal complements per clause), “CN_C” (Complex nominals per clause), “csubj_per_cl” (clausal subjects per clause), and “cc_per_cl” (clausal coordinating conjunctions per clause). Among the 14 features that are highly predictive of psychiatric medication nonadherence, only 3 had statistically higher means in posts written by adhering patients than by nonadhering patients, whereas the remaining 11 features were statistically insignificant.

TABLE 12: Mann–Whitney U test.

Features (29 in total)	Name	Nonadherence mean (std.)	Adherence mean (std.)	P
Syntactic complexity analyzer (SCA10)	MLC	9.369 (684)	8.630 (2.530)	0.003**
	C_S	2.260 (1.223)	2.368 (1.460)	0.5
	VP_T	2.645 (1.362)	2.528 (1.389)	0.378
	C_T	1.900 (0.943)	1.950 (1.009)	0.599
	T_S	1.227 (0.494)	1.209 (0.419)	0.818
	CT_T	0.426 (0.276)	0.471 (0.279)	0.046**
	CP_T	0.334 (0.331)	0.346 (0.378)	0.908
	CP_C	0.202 (0.305)	0.186 (0.184)	0.631
	CN_T	1.585 (1.229)	1.461 (0.979)	0.463
	CN_C	0.825 (0.433)	0.752 (0.381)	0.125
Syntactic sophistication (SS4)	all_av_construction_freq_stdev	624630.667 (259280.32)	628026.977 (280808.55)	0.865
	all_av_lemma_freq_stdev	2237002.019 (759008.785)	2200521.353 (840754.36)	0.773
	acad_av_approx_collexeme_stdev	30199.338 (69193.092)	16682.656 (42391.101)	0.027**
	all_av_lemma_freq_type	1469180.468 (793748.056)	1561373.037 (945196.99)	0.417
Noun phrase complexity (NPC3)	dobj_stdev	0.888 (0.432)	0.722 (0.476)	0**
	advmod_pobj_deps_NN_struct	0.030 (0.054)	0.045 (0.086)	0.291
	nsubj_NN_stdev	0.510 (0.498)	0.584 (0.532)	0.096
Clause complexity (CC12)	aux_per_cl	0.295 (0.154)	0.271 (0.157)	0.155
	ccomp_per_cl	0.124 (0.103)	0.145 (0.128)	0.121
	nsubjpass_per_cl	0.025 (0.045)	0.043 (0.073)	0.025**
	prep_per_cl	0.288 (0.159)	0.325 (0.233)	0.288
	nsubj_per_cl	0.679 (0.167)	0.711 (0.181)	0.031**
	mark_per_cl	0.098 (0.084)	0.108 (0.109)	0.55
	ncomp_per_cl	0.054 (0.085)	0.054 (0.091)	0.407
	cl_av_deps	2.724 (0.399)	2.732 (0.449)	0.492
	cc_per_cl	0.012 (0.035)	0.009 (0.028)	0.128
	csubj_per_cl	0.007 (0.028)	0.011 (0.036)	0.121
	dep_per_cl	0.100 (0.099)	0.109 (0.125)	0.968
	prepc_per_cl	0.023 (0.048)	0.023 (0.048)	0.643

4.3. *Diagnostic Utility of the Bayesian Machine Learning Classifier.* A major advantage of Bayesian machine learning classifiers is that they produce the posterior probabilities of a certain binary outcome dependent on the prior odds and the asymmetrical classification errors of the classifiers. In clinical research, Bayes' nomogram offers a graphical representation of the Bayesian probabilistic predictions [44–50].

In Figure 4, the axis on the left shows the baseline probability of the event of interest, which in our study was the prevalence of medication nonadherence among patients participating in the online mental health forum discussions on psychiatric medications. It was currently as high as 57%, which was calculated based on the total data we collected from the online forum. The middle axis represents likelihood or odds ratio. Likelihood ratio can be positive or negative. A positive likelihood ratio (LR+) is the ratio between

sensitivity and false positivity. In our study, the best-performing classifier (RVM_F46 with min-max normalisation) had a positive likelihood ratio of 3.02 (95% CI: 1.98, 4.63). If we draw a straight line on the nomogram and line up the prior (0.57) on the left axis, with the LR+ (3.02) on the middle axis, we can find the posterior probability on the right axis which was 80% (95% CI: 72%, 86%). The odds ratio of the posterior probability of positive cases was 3.9, which means that around 10 in every 13 psychiatric patients with a positive result as predicted by our model were following their medication regime. The middle axis can also be negative odds ratio which is the ratio between false negative cases and true negative cases. In our study, the negative likelihood ratio was 0.3 (95% CI: 0.2, 0.45). If repeating the same procedure of reading the Bayes' nomograph, we can find the posterior probability on the right axis which was 28% (95% CI: 21%, 37%). The odds ratio of the posterior

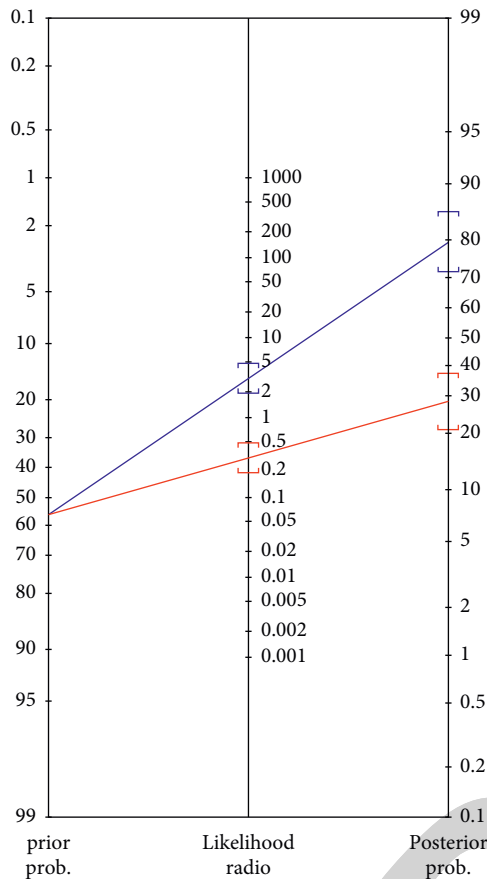


FIGURE 4: Interpreting the diagnostic utility of the best-performing classifier using Bayes' nomogram.

probability of true negative cases was 0.4, meaning that around 10 in every 14 psychiatric patients with a negative test result after screening by our classifier were not adhering to their medications.

5. Conclusion

Medication nonadherence represents a major burden on national health systems. According to the World Health Organization, increasing medication adherence may have a greater impact on public health than any improvement in specific medical treatments. More research is needed to better predict populations at risk of medication non-adherence. We developed clinically informative, easy-to-interpret machine learning classifiers to predict people with psychiatric disorders at risk of medication nonadherence based on the syntactic and structural features of written posts on health forums. Psychiatric medication non-adherence is a large and increasing burden on national health systems. Using Bayesian machine learning techniques and publicly accessible online health forum data, our study illustrates the viability of developing cost-effective, informative decision aids to support the monitoring and prediction of patients at risk of medication nonadherence. Our study has a limitation that the best-performing model comprised high-level, abstract syntactic and grammatical

features which were easier to extract from long written texts. This approach may not be suitable for the short text analysis and automatic classification. The best-performing model we developed requires advanced linguistic expertise to interpret the prediction results. In our future work, we will explore more explainable, intuitive natural language features to improve the interpretability of the machine learning models.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The work was supported by grants from the National Natural Science Foundation of China (no. 61772146) and Natural Science Foundation of Guangdong Province (no. 2021A1515011339).

References

- [1] S. Eduard, *Adherence to Long-Term Therapies: Evidence for Action*, World Health Organization, 2003.
- [2] F. J. Acosta, "Medication adherence in schizophrenia," *World Journal of Psychiatry*, vol. 2, no. 5, p. 74, 2012.
- [3] H. B. Bosworth, C. I. Voils, G. G. Potter, and D. C. Steffens, "The effects of antidepressant medication adherence as well as psychosocial and clinical factors on depression outcome among older adults," *International Journal of Geriatric Psychiatry*, vol. 23, no. 2, pp. 129–134, 2008.
- [4] R. Lingam and J. Scott, "Treatment non-adherence in affective disorders," *Acta Psychiatrica Scandinavica*, vol. 105, no. 3, pp. 164–172, 2002.
- [5] R. Maidment, G. Livingston, and C. Katona, "Just keep taking the tablets?: adherence to antidepressant treatment in older people in primary care," *International Journal of Geriatric Psychiatry*, vol. 17, no. 8, pp. 752–757, 2002.
- [6] R. J. Julius, M. A. Novitsky, and W. R. Dubin, "Medication adherence: a review of the literature and implications for clinical practice," *Journal of Psychiatric Practice*, vol. 15, no. 1, pp. 34–44, 2009.
- [7] Z. A. Marcum, M. A. Sevick, and S. M. Handler, "Medication nonadherence," *JAMA*, vol. 309, no. 20, p. 2105, 2013.
- [8] K. Anderson, S. G. Jue, and K. J. Madaras-Kelly, "Identifying patients at risk for medication mismanagement: using cognitive screens to predict a patient's accuracy in filling a pillbox," *The Consultant Pharmacist*, vol. 23, no. 6, pp. 459–472, 2008.
- [9] R. Horne, J. Weinman, and M. Hankins, "The beliefs about medicines questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication," *Psychology and Health*, vol. 14, no. 1, pp. 1–24, 1999.
- [10] R. Nakajima, F. Watanabe, and M. Kamei, "Factors associated with medication non-adherence among patients with lifestyle-related non-communicable diseases," *Pharmacy*, vol. 9, no. 2, p. 90, 2021.

- [11] M. Kachooei and F. Afsahi, "Relationship between hypertension with irrational health beliefs and health locus of control," *Journal of Education and Health Promotion*, vol. 9, no. 1, p. 110, 2020.
- [12] L. Osterberg and T. Blaschke, "Adherence to medication," *New England Journal of Medicine*, vol. 353, no. 5, pp. 487–497, 2005.
- [13] C. A. McHorney, "The Adherence Estimator: a brief, proximal screener for patient propensity to adhere to prescription medications for chronic disease," *Current Medical Research and Opinion*, vol. 25, no. 1, pp. 215–238, 2008.
- [14] L. M. West, R. Borg Theuma, and M. Cordina, "Health locus of control: its relationship with medication adherence and medication wastage," *Research in Social and Administrative Pharmacy*, vol. 14, no. 11, pp. 1015–1019, 2018.
- [15] R. E. F. Marsden, J. Francis, and I. Garner, "Use of GFCF dcAn investigation into parents' beliefs using the theory of planned behaviour," *Journal of Autism and Developmental Disorders*, vol. 49, no. 9, pp. 3716–3731, 2019.
- [16] T. A. Miller, "Health literacy and adherence to medical treatment in chronic and acute illness: a meta-analysis," *Patient Education and Counseling*, vol. 99, no. 7, pp. 1079–1086, 2016.
- [17] E. Vermeire, H. Hearnshaw, P. Van Royen, and J. Denekens, "Patient adherence to treatment: three decades of research. A comprehensive review," *Journal of Clinical Pharmacy and Therapeutics*, vol. 26, no. 5, pp. 331–342, 2001.
- [18] M. R. DiMatteo, K. B. Haskard-Zolnierok, and L. R. Martin, "Improving patient adherence: a three-factor model to guide practice," *Health Psychology Review*, vol. 6, no. 1, pp. 74–91, 2012.
- [19] M. R. DiMatteo, "Social support and patient Adherence to medical treatment: a meta-analysis," *Health Psychology*, vol. 23, no. 2, pp. 207–218, 2004.
- [20] N. person, "Literacy self-efficacy, and HIV medication adherence," *Patient Education and Counseling*, vol. 65, pp. 253–260, 2007.
- [21] D. H. Howard, J. Gazmararian, and R. M. Parker, "The impact of low health literacy on the medical costs of Medicare managed care enrollees," *The American Journal of Medicine*, vol. 118, no. 4, pp. 371–377, 2005.
- [22] M. R. DiMatteo, "Variations in patients' adherence to medical recommendations," *Medical Care*, vol. 42, no. 3, pp. 200–209, 2004.
- [23] A. Lo, H. Chernoff, T. Zheng, and S.-H. Lo, "Why significant variables aren't automatically good predictors," *Proceedings of the National Academy of Sciences*, vol. 112, no. 45, pp. 13892–13897, 2015.
- [24] A. Lo, H. Chernoff, T. Zheng, and S.-H. Lo, "Framework for making better predictions by directly estimating variables' predictivity," *Proceedings of the National Academy of Sciences*, vol. 113, no. 50, pp. 14277–14282, 2016.
- [25] D. Bzdok, D. Engemann, and B. Thirion, "Inference and prediction diverge in biomedicine," *Patterns*, vol. 1, no. 8, Article ID 100119, 2020.
- [26] D. Bzdok, G. Varoquaux, and B. Thirion, "Neuroimaging research: from null-hypothesis falsification to out-of-sample generalization," *Educational and Psychological Measurement*, vol. 77, no. 5, pp. 868–880, 2016.
- [27] D. N. T. Doan, B. Ku, J. Choi et al., "Predicting dwith prefrontal electroencephalography and event-related potential," *Frontiers in Aging Neuroscience*, vol. 13, 2021.
- [28] A. Onan, "A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer," *Expert Systems with Applications*, vol. 42, no. 20, pp. 6844–6852, 2015.
- [29] S. Lodh, S. Mahajan, C. Vanipriya, A. Bhardwaj, and A. K. Pandit, "Detection of broken and good medical tablets using various machine learning models," in *Advances in Intelligent Computing and Communication*, pp. 43–49, Springer, Singapore, 2021.
- [30] A. Onan, S. Korukoğlu, and H. Bulut, "Ensemble of keyword extraction methods and classifiers in text classification," *Expert Systems with Applications*, vol. 57, pp. 232–247, 2016.
- [31] A. Onan and S. Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification," *Journal of Information Science*, vol. 43, no. 1, pp. 25–38, 2017.
- [32] A. Onan, "Classifier and feature set ensembles for web page classification," *Journal of Information Science*, vol. 42, no. 2, pp. 150–165, 2016.
- [33] S. Mahajan, A. Raina, X.-Z. Gao, and A. Kant Pandit, "Plant recognition using morphological feature extraction and transfer learning over SVM and AdaBoost," *Symmetry*, vol. 13, no. 2, p. 356, 2021.
- [34] A. Onan, "Two-stage topic extraction model for bibliometric data analysis based on word embeddings and clustering," *IEEE Access*, vol. 7, pp. 145614–145633, 2019.
- [35] A. Onan, "Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 23, Article ID e5909, 2021.
- [36] A. Onan, "An ensemble scheme based on language function analysis and feature engineering for text genre classification," *Journal of Information Science*, vol. 44, no. 1, pp. 28–47, 2018.
- [37] K. Kristopher and C. Scotta, "Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices," *The Modern Language Journal*, vol. 102, pp. 333–349, 2018.
- [38] K. Kyle and S. Crossley, "Assessing syntactic sophistication in L2 writing: a usage-based approach," *Language Testing*, vol. 34, no. 4, pp. 513–535, 2017.
- [39] K. Kyle, *Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-Based Indices of, 2016 of Syntactic Sophistication*, Sematic Scholar, 2016.
- [40] X. Huang, L. Zhang, B. Wang, F. Li, and Z. Zhang, "Feature clustering based support vector machine recursive feature elimination for gene selection," *Applied Intelligence*, vol. 48, no. 3, pp. 594–607, 2017.
- [41] Z. Yin and J. Zhang, "Operator functional state classification using least-square support vector machine based recursive feature elimination technique," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 101–115, 2014.
- [42] C. Bowd, I. Lee, M. H. Goldbaum et al., "Predicting glaucomatous progression in glaucoma suspect eyes using relevance vector machine classifiers for combined structural and functional measurements," *Investigative Ophthalmology & Visual Science*, vol. 53, no. 4, p. 2382, 2012.
- [43] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, pp. 211–244, 2001.
- [44] J. Page, "Using Bayes' nomogram to help interpret odds ratios," *Evidence-Based Medicine*, vol. 8, no. 5, pp. 132–134, 2003.
- [45] A. Laupacis, D. L. Sackett, and R. S. Roberts, "An assessment of clinically useful measures of the consequences of treatment," *New England Journal of Medicine*, vol. 318, no. 26, pp. 1728–1733, 1988.

Retraction

Retracted: Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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. Chakraborty, H. F. Mahdi, M. H. A. Al-Abyadh, K. Pant, A. Sharma, and F. Ahmadi, "Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5731532, 10 pages, 2022.

Research Article

Large-Scale Textual Datasets and Deep Learning for the Prediction of Depressed Symptoms

Sudeshna Chakraborty ¹, **Hussain Falih Mahdi** ², **Mohammed Hasan Ali Al-Abyadh** ³,
Kumud Pant ⁴, **Aditi Sharma** ⁵ and **Fardin Ahmadi** ⁶

¹Computer Science & Engineering, Lloyd Institute of Engineering and Technology, Greater Noida, India

²Computer Engineering Department, University of Diyala, Iraq

³Prince Sattam Bin Abdulaziz University Alkharj Saudi Arabia, College of Education- Thamar University, Thamar, Yemen

⁴Department of Biotechnology, Graphic Era Deemed to Be University, Dehradun, Uttarakhand, India

⁵Computer Science Engineering & Information Technology, Institute of Engineering & Technology, Lucknow, Uttar Pradesh, India

⁶Rana University, Kabul, Afghanistan

Correspondence should be addressed to Fardin Ahmadi; fardin.ahmadi@bcs.ru.edu.af

Received 19 January 2022; Accepted 12 March 2022; Published 12 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Sudeshna Chakraborty 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.

Millions of people worldwide suffer from depression. Assessing, treating, and preventing recurrence requires early detection of depressive symptoms as depression-related datasets expand and machine learning improves, intelligent approaches to detect depression in written material may emerge. This study provides an effective method for identifying texts describing self-perceived depressive symptoms by using long short-term memory (LSTM) based recurrent neural networks (RNN). On a huge dataset of a suicide and depression detection dataset taken from Kaggle with 233337 datasets, this information channel featured text-based ten questions. Then, using a one-hot technique, medical and psychiatric practitioners extract strong features from probably depressed symptoms. The characteristics outperform the usual techniques, which rely on word frequencies rather than symptoms to explain the underlying events in text messages. Depression symptoms can be distinguished from nondepression signals by using a deep learning system (nondepression posts). Eventually, depression is predicted by the RNN. In the suggested technique, the frequency of depressive symptoms outweighs their specificity. With correct annotations and symptom-based feature extraction, the method may be applied to different depression datasets. Because of this, chatbots and depression prediction can work together.

1. Introduction

Depression is a regular occurrence in the workplace, school, and home stress can all contribute to depression [1]. Adults are affected by adolescent depression; approximately 0.8 million individuals commit suicide each year [2, 3]. Mental illnesses account for five of the top 10 debilitating conditions, with depression being the most frequent [4]. As a result, depression is a serious illness. More than half of all people have mild depression [5]. Adults in their forties and fifties are particularly vulnerable. When depression is recognized early, it is easier to treat [6–10]. However,

identifying depression symptoms requires time and effort. To predict mental illness, physician interviews and hospital or agency questionnaire surveys [11] are now employed. One-on-one surveys are used in this method.

Instead of interviews or questionnaires, spontaneous writings submitted by users can be used to forecast depression. Clinical psychology has looked into the link between a language user (speaker or writer) and their text [12]. Havigerova et al. found that trip-related informal language might predict depression in recent research [12]. As a result, electronic records and data are becoming more vital in health care. The application of recent breakthroughs in

natural language processing and artificial intelligence to detect depressive symptoms in informal writing is promising for artificial intelligence (AI). Linguistics and computing are used to help computers interpret text. The goal in this scenario is to assign negative or positive polarity to opinions, ideas, and concepts. Automated text analysis in conversations or blog postings can detect depressive symptoms [13–18]. However, there is still a lot to learn about reading letters for melancholy. It is difficult to write about serious depression. Depression symptoms are difficult to diagnose with a single statement. Our automated detection method, we feel, can make a significant scientific contribution. As a result, the present study uses artificial intelligence to detect depressive signs in the text.

Linear discriminant analysis (LDA) is an excellent method for visualizing discriminant data [19–22]. It operates by grouping comparable samples together. Its goal is to improve between-class scatter while lowering within-class dispersion. Facial expression recognition and human activity recognition are examples of real-world LDA applications. The dimensionality of class data is reduced using LDA.

Deep neural networks have lately aided in pattern recognition and AI research [23–34]. It does, however, have two big flaws. The first fault is that it is very tight. Data modeling takes a long time. Restricted Boltzmann machines (RBMs) were used to speed up training in the early days of deep learning. A better instrument for discriminating than others. Convolution neural network (CNN) extracts and trains its data. An abstract feature hierarchy may be created using convolution [24]. Instead of analyzing time-series data, CNNs are employed for image and video analysis. In the examination of sequential data and patterns, RNNs outperform CNNs [30]. For high-dimensional and time-correlated input, RNNs employ LSTM to overcome the problem of vanishing gradients. An LSTM-based RNN is therefore employed in this work to mimic emotional content in text data.

Human physical and mental functions have been extensively studied using machine learning [35–41]. Industry stakeholders are requesting more openness when machine learning algorithms are used to provide crucial forecasts [42]. The major danger is creating and implementing bad AI judgments. The list goes on. Precision medicine practitioners, for example, require more than mere machine learning predictions to support their diagnosis. Other professions, such as medicine, may have similar requirements. In rare cases, this may result in system rejection. Recent research emphasizes the necessity for explainable AI to build trust in machine learning results. Local interpretable model-agnostic explanations (lime), Shapley additive explanation (SHAP), and layerwise relevance propagation are only a few of the modern explanation algorithms that may be used nowadays. Layerwise relevance propagation (LRP). As a result, lime is small and focused on offering quick, posthoc explanations. As a result, when the model is completed, this study will make use of lime to determine why (importance of the attributes). The goal of this project is to identify depressive symptoms in text for a smart chatbot application. Text queries are processed by the server using feature

extraction and deep learning. The findings may lead to additional suggestions from the server. RNN features are developed from all user text input throughout the training phase.

Based on the test results, the trained model determines if the user is sad. To compare proposed features to existing features, LDA is utilized. Finally, we use a widely used method to produce posthoc, local, and understandable machine learning explanations. Here is how the paper makes a difference:

Medical and psychiatric professionals point out certain characteristics that might indicate depression. To imitate emotions, it employs LSTM, attention, and thick layers. Section 2 shows information gathering and analysis. Section 3 depicts methodology. Sections 4 and 5 explain results and conclusion, respectively.

2. Information Gathering and Analysis

Recognizing mental health disorders necessitates the gathering of data. Social media data, such as Facebook status updates, is insufficient. [43]. Use of the massive text-based dataset on the ung.no public information website. On ung.no, young people can anonymously ask questions in Norwegian. Answers and counseling are provided by professionals (doctors, psychologists, nurses, and so on). These are made available to the public via the Internet. Teenagers define and categorize their postings on ung.no. The topic for this week was “emotions and mental health.” They are usually short, but they describe the mental state, symptoms, and behavior. To begin with, some of the writings depict depression that has been medically diagnosed. Many texts examine the history and symptoms of depression, either rejecting or confirming the diagnosis. They appear to be an expression of self-perceived sadness. Clinical diagnoses are mirrored in self-perceived mental states [44–46]. There are a few words that tell tales and portray emotions without using the word “sad.”

It is thought to be depressive symptoms. One of the data categories is depression. The signs of depression were then validated by a competent general practitioner. Melancholy is determined by analyzing a set of phrases and words. The accusations were corroborated by a doctor. In the appendix, you will see possible remarks and/or terms that unhappy kids could use in their searches. To get features for each message, use phrases and words. Look at Table 1 to learn about depression in English. There were 277,552 posts in all, including depressing messages. From that dataset, we used 11,807 and 21,470 postings in our two investigations. Text features are used as binary patterns in a depression prediction machine learning model. The following list of stemmed terms demonstrates the breadth of terminology related to depression [47]. Table 2 displays the snapshot of the dataset taken for the analysis purpose.

3. Methodology

The proposed methodology is discussed in the section, here preprocessing is the first part of the method and then modeling and the proposed model are given.

TABLE 1: Words and phrases associated with depression are commonly used in English.

S.N.	Words
1	Nothing to eat
2	Ending my life is the only option I have left.
3	Suicide
4	Suicidal thoughts and tears
5	Take my own life away from me!
6	Take my life away from me.
7	a void of any kind
8	Sadness
9	Always drained in energy and lacking in inspiration
10	Not a thing
11	Nothing to do

LDA is used for a variety of purposes. To maximize interclass scatterings, LDA seeks to reduce scatterings inside a class.

TABLE 2: : Suicide and depression detection dataset Kaggle 233337 datasets.

Message	Case
Ex-wife threatening suicide recently I left my wife for good because she has cheated on me twice and lied to me so much that I have decided to refuse to go back to her. As of a few days ago, she began threatening suicide. I Have tirelessly spent these past few days talking her out of it and she keeps hesitating because she wants to believe I'll come back. I Know a lot of people will threaten this to get their way, but what happens if she does? What do I do and how am I supposed to handle her death on my hands? I Still love my wife but I cannot deal with getting cheated on again and constantly feeling insecure. I'm worried today may be the day she does it and I hope so much it does not happen.	Suicide
I Need help just help me I'm crying so hard	Suicide
It ends tonight. I Cannot do it anymore.	Suicide
Do you think getting hit by a train would be painful? Guns are hard to come by in my country but trains are not. I Just do not want to suffer through, do you think this would be a painless method of suicide?	Suicide
Iâ€™m scared. Everything just seems to be getting worse and worse. Iâ€™m young and I think Iâ€™m transgender but Iâ€™m not even sure about that. I canâ€™t tell if Iâ€™m just lying to myself or if Iâ€™m actually trans, I feel so overwhelmed with thoughts and emotions and I canâ€™t just take it anymore.	
I just wish I could at least know for sure if I was trans, and even then I have to worry about if my (religious) family will be accepting and if I can do anything to alleviate my pain a bit.	
I Cut myself for the first time yesterday, I barely even drew blood so I canâ€™t even fucking hurt myself correctly. I donâ€™t think Iâ€™ll ever be able to do anything correctly, I want to pursue music but I know thereâ€™s no money to be found in that field unless I become famous but thatâ€™s not happening.	Suicide
Currently, Iâ€™m not seriously debating suicide but the thoughts keep coming back and they just keep getting worse. Iâ€™m is not sure if I can take this much longer, I just wish I was born a girl. I Want to cry.	
Am I weird I do not get affected by compliments if it's coming from someone I know IRL but I feel really good when Internet strangers do it	Nonsuicide
Everyone wants to be "edgy" and it's making me self-conscious I feel like I do not stand out. I Can draw yes and play the guitar but I honestly feel like being stuck in the past, my taste in music is all rock and alt-metal from 2000s to the '90s and it does not make me feel unique it's just my style but seeing as my friends and classmates get more into rap and EDM it's hard for me to feel like I fit in.	
Then I do not feel like I stand out is because of all the others copying a style and if I do that I'd be just another "Quirky kid" who's in a cringe phase.	
Many of my friends say that I look good in grunge style and I kinda agree but it's hard for me to continue that if I cannot even stand out from all the "edgy	Nonsuicide
People who wore crosses and wallet chains and do tiktoks"	
Feels like I do not fit in in all categories, am scared that people might confuse me with a CLOUT CHASER or a fucking TikTok e boy goddamn	
I Hate my life	
Hey, I'm gonna sleep with socks whatcha gonna do? Put them off?! good luck ima gonna sleep with warm feet	Nonsuicide

3.1. *Preprocessing.* The survey questions are put in rows in the dataset, and the survey participants are grouped into columns, resulting in distinct health domain tables. Because the tables are not all organized in the same way, pre-processing is required to categorize the data. For our research, we will only use one-third of the dataset: the survey questions. To eliminate duplicates and make it more

computer-readable, the data was cleaned and modified. The data formats were chosen to allow for comparisons and contrasts between the datasets. To establish a uniform scale across all of the questions, normalization was also necessary. When data is prepared to utilize psychological domain information from functional diagnostic criteria, the data structure is reconstructed. All tables should be reconstructed

using just six functional categories of depression diagnostic criteria. It makes no difference whether there are more or fewer questions because the participants are all the same. The six tables may be consolidated into one because they all have the same row index. When each table is instantly seen, it generates a new dataset with participants as instances and questions as features.

3.2. Classification by Modeling. An ensemble classification approach is used to build the model. Many classification algorithms are used simultaneously using Independent Ensemble Methodology (IEM). The model employs the support vector machine, artificial neural network, K-nearest neighbor (KNN), and decision tree algorithms. In a single training run, each composite classifier is trained on the same piece of training data. A k-fold cross-validation approach is utilized as a part of the assessment process. The ensemble classifier is built by merging the results of all the composite classifiers into a single prediction. An ensemble classification technique employs many independent classifiers to improve prediction accuracy.

An ensemble method, on average, outperforms a single algorithm in terms of prediction performance. The advantages of performance:

- (i) By averaging numerous alternative hypotheses, an incorrect hypothesis is avoided from being chosen.
- (ii) Combining several learning ensemble approaches reduces the possibility of reaching a local minimum, which saves time and money.
- (iii) Using numerous models and diverse representations, we were able to improve the data fit and extend the search area.

The ensemble approach simulates human behavior by looking at a variety of choices. When we compare our preprocessed data to other baseline models, we may conclude that the ensemble strategy for this experiment is a superior technology.

An ensemble model is exemplified by this. The accuracy of predictions is anticipated to increase if all four techniques are used together. Training each of the ensemble's various submodels is required to broaden the scope of the ensemble classifier. To combine the outputs from all of the initial classifiers in our model, we employ a weighted ensemble technique. A weighted ensemble strategy is incredibly broad due to the same outputs of each base classifier. The weights of classifiers are determined by their accuracy on a validation set.

It is fantastic to use a machine learning model to decode time-series data. Therefore, RNNs are employed. [22] RNNs are commonly employed to represent time-sequenced data. In RNNs, previous and present states are linked through recurrent connections. Neural networks rely heavily on memory. A vanishing gradient problem or a processing limit is a common problem for RNN algorithms. The text feature extraction and the suggested model are listed as follows: Figure 1 depicts the sample post with words belonging to depression and nondepression category.

4. Experimental Results

We used data from Kaggle.com. There are depression-related texts included in the collection. Some of the communications were annotated by medical and psychiatric professionals. Testing was conducted on a 32 GB RAM, Windows 10, and the TensorFlow 2.4.1 deep learning tool with an Intel (R) Core (TM) 7700HQ CPU operating at 2.8 GHz and 2.81 GHz.

4.1. Dataset and Experiments. For the first dataset and trials, there were 11,807 messages in total, with 1820 of those identified as depression texts (detailed descriptions of depression symptoms) and 9987 of those classified as non-depression texts (not describing symptoms of depression). These tables show the tenfold classification reports used in most of the training and testing datasets.

During tenfold training, the accuracy and loss are shown in Tables 3–5. Fold training looks to be going well, except for a slight tweak. This approach's confusion matrix is depicted in Figures 2 and 3 for folds 1 and 2. The suggested features outperform one-hot and LSTM with mean recall rates of 0.98 and 0.99 for depression and nondepression, respectively. When comparing precision levels, the precision-recall curve illustrates the trade-off between accuracy and recall. A large area under the curve indicates that the person has strong recall and accuracy. Because high accuracy implies low false positives, and strong recall implies low false negatives, high accuracy implies low false positives.

Accuracy at the 0.99 level indicates that the method is long-lasting.

Figure 4 shows the machine learning model's overall probability. In most ways, a three-dimensional scatter plot is comparable to a two-dimensional scatter plot. Scatter plots are often used to illustrate the relationship between two numbers. Positive or negative, strong or weak, linear or nonlinear relationships between two variables may be depicted in a number of ways. Additionally, scatter plots may aid you in detecting other patterns in the data.

Emotional states' one-hot characteristics, TF-IDF characteristics, and LDA's projected strong characteristics are depicted in three-dimensional renderings in Figures 5–8 in this section.

The mean accuracy (percentage) and forecast accuracy (percentage) for different approaches to all participants are also presented in Table 6.

One of the study's possible benefits is assisting users who show indicators of depression but have not yet been officially diagnosed. In general, the earlier patients get help for depression, the better their outcomes and costs. An intrusive marketing tactic used by mental health organizations to target potential customers based on their web behavior may be deemed intrusive. People are skeptical of this strategy based on preliminary findings. Explainability and interpretability are important factors in overcoming the barrier of using social media data for mental health prediction models.

	Depression	Non-depression
I	████████████████████	
Need	██████████████	
Help	████████████████████	
Just		██████████████
Help	████████████████████	
Me	████████████████████	
Im	████████████████████	
Crying	██████████████	
So		██████████████
hard		██████████████

FIGURE 1: Sample post shows words belonging to the depression or nondepression category.

```

Start with A = 1.
The number of feature sentences is denoted by the letter I.
Between b = 1 and I
Obtaining a feature sentence is a feat.
Using tokenized Feat, separate Words from Feat =
For each word in Tokenized Feat, write a new sentence.
If the word appears in the text,
[A] L = 1
Else
[A] L = 0
End If
Return L
The end of For
    
```

ALGORITHM 1: Text feature extraction

```

Start
Training and Testing of Text
Assign T = the number of text training samples.
Do this for i = 1 to M.
Text = Text (i) to get it.
Obtain features and assign the label of the text to Yi.
End of
Get all the training features L and labels Y.
Train the model!
Calculate the cosine similarity of two vectors: one from the dataset and one from a previously used typed vector.
If the similarity is there, then it is depression; else, there is no depression.
End
    
```

ALGORITHM 2: A proposed model

TABLE 3: For 100 epochs of training, the accuracy and loss are shown in a graph When a patient is suffering from depression.

State	Fold	Precision	Recall	F1-score	Support
Depression	1	0.95	0.97	0.96	189
	2	0.97	0.96	0.96	184
	3	0.96	0.94	0.95	187
	4	0.96	0.96	0.96	162
	5	0.99	0.94	0.97	190
	6	0.97	0.92	0.95	188
	7	0.95	0.99	0.97	172
	8	0.97	0.98	0.98	179
	9	0.97	0.98	0.98	179
	10	0.99	0.96	0.97	187

TABLE 4: 100 epochs of training in a nondepressive condition is shown to illustrate the accuracy and loss.

State	Fold	Precision	Recall	F1-score	Support
Nondepression	1	0.99	0.98	0.98	992
	2	0.98	0.99	0.98	997
	3	0.98	0.98	0.98	994
	4	0.99	0.99	0.99	1019
	5	0.98	0.99	0.99	991
	6	0.98	0.99	0.98	993
	7	0.99	0.98	0.99	1009
	8	0.99	0.99	0.99	1001
	9	0.99	0.99	0.99	1001
	10	0.98	0.99	0.99	993

TABLE 5: 100 epochs of training using the mean value as the state demonstrates the accuracy and loss during tenfold training.

State	Fold	Precision	Recall	F1-score	Support
Mean/Total	1	0.97	0.98	0.97	1181
	2	0.97	0.97	0.97	1181
	3	0.97	0.96	0.96	1181
	4	0.97	0.97	0.97	1181
	5	0.98	0.96	0.98	1181
	6	0.97	0.95	0.96	1181
	7	0.97	0.98	0.98	1181
	8	0.98	0.98	0.98	1181
	9	0.98	0.98	0.98	1181
	10	0.98	0.97	0.98	1181

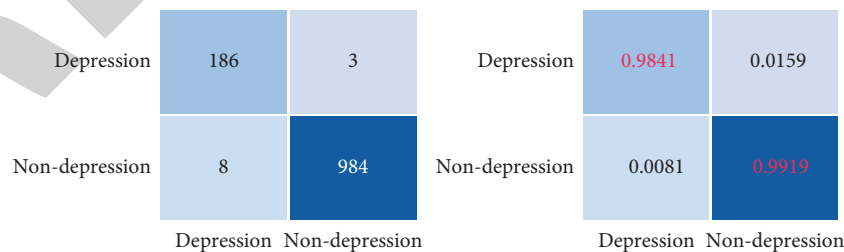


FIGURE 2: The proposed approach produces a fold-1 confusion matrix.

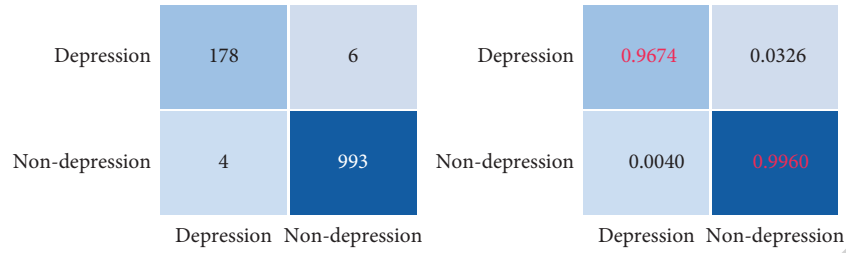


FIGURE 3: The proposed approach produces a fold-2 confusion matrix.

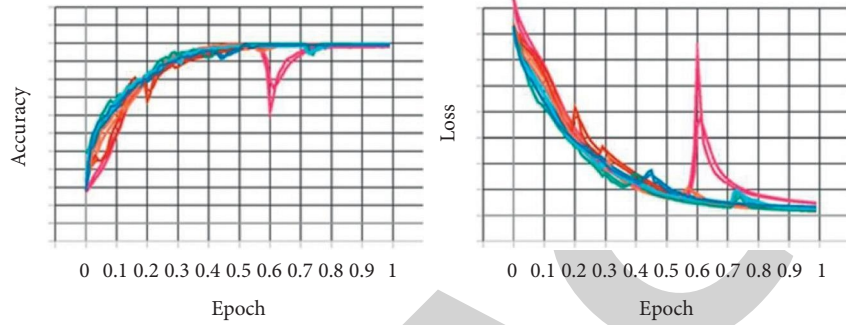


FIGURE 4: During trials on the dataset using the suggested technique, accuracy and loss of 10-folds were observed.

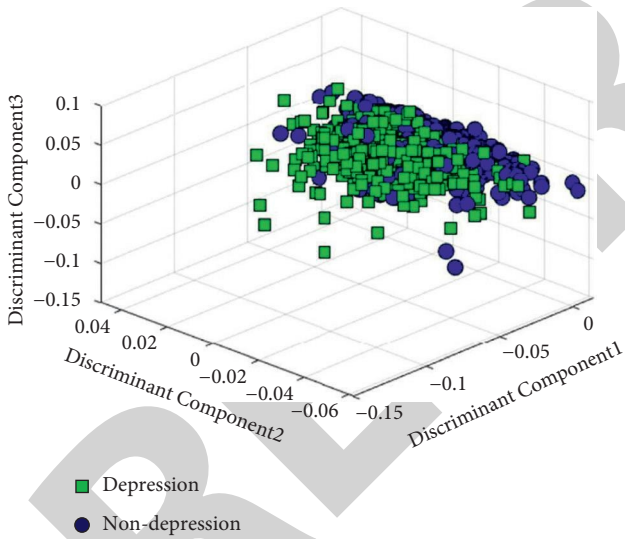


FIGURE 5: Traditional one-hot properties of two emotional states have been mapped out in a three-dimensional fashion by following LDA.

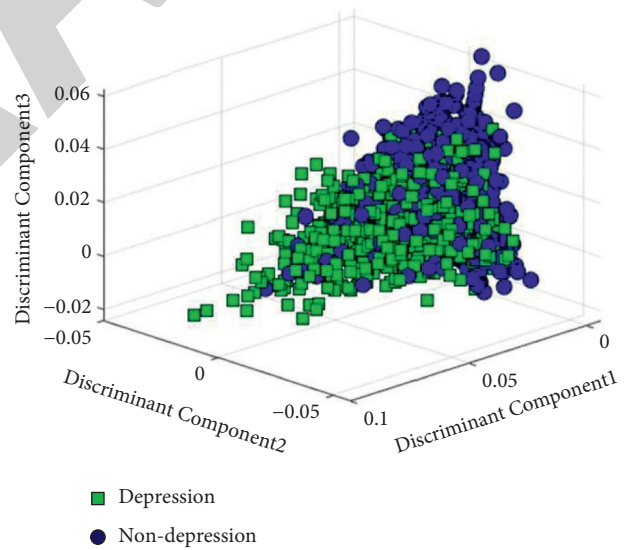


FIGURE 6: A 3-D representation of the typical TF-IDF properties of two emotional states is depicted following LDA.

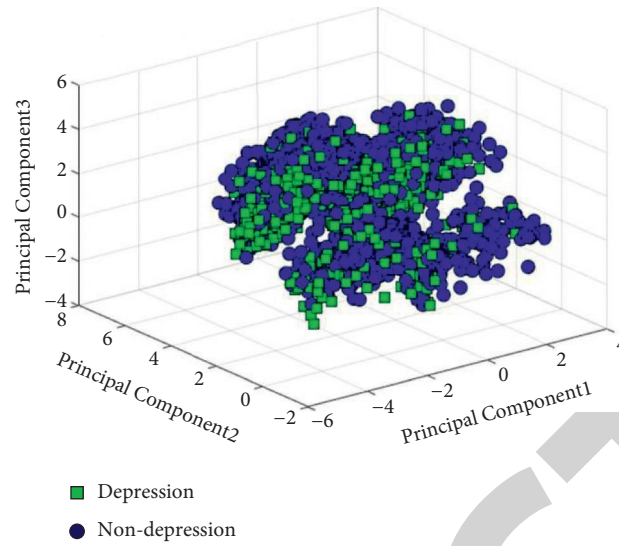


FIGURE 7: Following PCA, shows a three-dimensional representation of the strong attributes predicted of two emotional states in three dimensions.

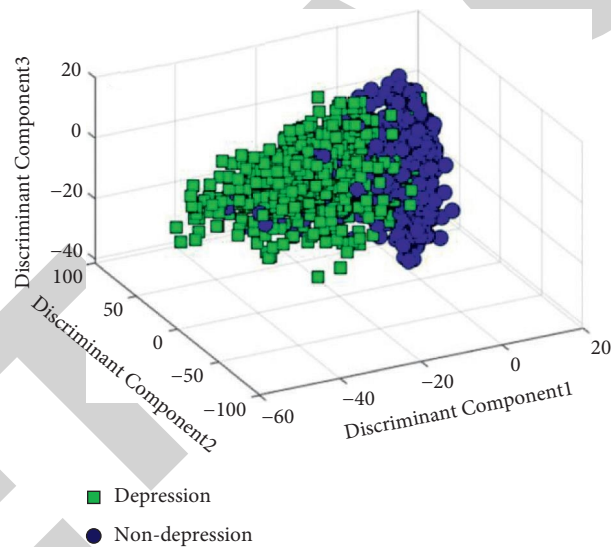


FIGURE 8: Emotional states are shown in a three-dimensional picture as a result of LDA.

TABLE 6: The accuracy of predictions made by various approaches for all people.

S.N.	Approches	Mean accuracy (%)
1.	One-hot logistic regression	83.98
2.	Support vector machine	84.87
3.	Artificial neural network	83.56
4.	TF-IDF decision trees	82.25
5.	K-nearest neighbour	80.58
6.	Decision tree	84.25
7.	Ensemble model	87.69
8.	Usr2Vec	87.02
9.	MIL-SocNet	88.68
10.	TF-IDF deel model	88.26
11.	Proposed method	92.02

5. Conclusion

This study's goal was to develop a multimodal human depression prediction strategy using RNN deep learning and robust depression symptom features. First, text data from suicide datasets for young users is first used. An on-hot approach is then used after extracting words from phrases that describe depressive symptoms. The one-hot features were also used to train an LSTM-based deep RNN to represent and forecast unknown sensor text emotional states. Using the suggested method, the first and second datasets contain 11,807 and 21,807 texts, respectively. However, while mental characteristics appear to be the most important contributors to depression prediction, future analyses of these subsets in isolation and utilizing relevant data will enhance the classification performance and comprehension of the association between characteristics and depression. In the future, our method might be used to extract characteristics from social media, which is a current trend in ML methods. Classifying textual data in this way improves the ensemble system's reliability and sensitivity. Deep learning techniques like DNN might expand the ensemble classification range. As a result, this will be the subject of our next round of research to further refine this approach. Traditional techniques could only reach 91 percent mean recognition performance, suggesting the new approach's robustness. To create effective user interfaces for improved emotional care, the characteristics employed in this study can be leveraged to assist machine learning judgments. Deep learning with a large dataset may be an efficient system to be studied. Using cutting-edge technology, mental health services can assess and predict normal and severe mood problems in real-time.

Data Availability

The dataset has been downloaded from the website ung.no, which is a public Norwegian information website.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] WHO, <https://www-who-int.ezproxy.uio.no/news-room/fact-sheets/detail/suicide>, 2021.
- [2] H. Häfner, K. Maurer, G. Trendler, W. der Heiden, and M. Schmidt, "The early course of schizophrenia and depression," *European Archives of Psychiatry and Clinical Neuroscience*, vol. 255, no. 3, pp. 167–173, 2005.
- [3] P. Fusar-Poli, B. Nelson, L. Valmaggia, A. R. Yung, and P. K. McGuire, "Comorbid depressive and anxiety disorders in 509 individuals with an at-risk mental state: impact on psychopathology and transition to psychosis," *Schizophrenia Bulletin*, vol. 40, no. 1, pp. 120–131, 2014.
- [4] World Health Organization, *Depression and Other Common Mental Disorders: Global Health Estimates*, WHO, Geneva, Switzerland, 2017.
- [5] R. C. Kessler and E. J. Bromet, "The epidemiology of depression across cultures," *Annual Review of Public Health*, vol. 34, no. 1, pp. 119–138, 2013.
- [6] W. H. O Depression, *Other Common Mental Disorders: Global Health Estimates*, World Health Organization, Geneva, Switzerland, 2017.
- [7] A. H. Weinberger, M. Gbedemah, A. M. Martinez, D. Nash, S. Galea, and R. D. Goodwin, "Trends in depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups," *Psychological Medicine*, vol. 48, no. 8, pp. 1308–1315, 2018.
- [8] E. Cambria and B. White, "Jumping nlp curves: a review of natural language processing research [review article]," *IEEE Computational Intelligence Magazine*, vol. 9, no. 2, pp. 48–57, 2014.
- [9] R. W. Picard, "Affective computing: challenges," *International Journal of Human-Computer Studies*, vol. 59, no. 1–2, pp. 55–64, 2003.
- [10] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102–107, 2016.
- [11] S. Wang, G. Peng, Z. Zheng, and Z. Xu, "Capturing emotion distribution for multimedia emotion tagging," *IEEE Transactions on Affective Computing*, vol. 12, no. 4, pp. 821–831, 2021.
- [12] J. M. Havigerová, J. Haviger, D. Kučera, and P. Hoffmannová, "Text-based detection of the risk of depression," *Frontiers in Psychology*, vol. 10, p. 513, 2019.
- [13] J. Tao and T. Tan, "Affective computing: a review," in *Proceedings of the International Conference On Affective Computing And Intelligent Interaction*, pp. 981–995, Beijing, China, October 2005.
- [14] H. Yang, A. Willis, A. De Roeck, and B. Nuseibeh, "A hybrid model for automatic emotion recognition in suicide notes," *Biomedical Informatics Insights*, vol. 5, no. 1, pp. 17–30, 2012.
- [15] B. Desmet and V. Hoste, "Emotion detection in suicide notes," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6351–6358, 2013.
- [16] M. Allouch, A. Azaria, R. Azoulay, E. Ben-Izchak, M. Zwilling, and D. A. Zachor, "Automatic detection of insulting sentences in conversation," in *Proceedings of the 2018 IEEE international conference on the science of electrical engineering in Israel (ICSEE)*, pp. 1–4, IEEE, Eilat, Israel, December 2018.
- [17] S. John, W. Ederyn, and C. Bruce, *The Social Psychology of Telecommunication*, Wiley, London, UK, 1976.
- [18] C. Strapparava and R. Mihalcea, "Learning to identify emotions in text," in *Proceedings of the 2008 ACM symposium On Applied Computing*, pp. 1556–1560, ACM, Sierre, Switzerland, March 2008.
- [19] Y. Ding, X. Chen, Q. Fu, and S. Zhong, "A Depression Recognition Method for College Students Using Deep Integrated Support Vector Algorithm," *IEEE Access*, vol. 8, pp. 75616–75629, 2020.
- [20] E. Yildirim and E. Avci, "Classification of profitability of sands by linear discriminant analysis, 2018," in *Proceedings Of The Zbornik Radova Geo- Expo 2018*, Colombo, Sri Lanka, October 2018.
- [21] R. A. Vannatta and K. N. LaVenía, "Linear discriminant analysis," in *SAGE Research Methods Foundations*, P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, and R. A. Williams, Eds., , 2020.
- [22] S. Gardner-Lubbe, "Linear discriminant analysis for multiple functional data analysis," *Journal of Applied Statistics*, vol. 48, no. 11, pp. 1–17, 2020.
- [23] M. Z. Uddin and J. Torresen, "Activity recognition using smartphone sensors, robust features, and recurrent neural network," in *Proceedings of the 2019 13th international symposium on medical information and communication technology (ISMICT)*, Oslo, Norway, May 2019.

Retraction

Retracted: A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] Z. A. Khan, S. Naz, R. Khan, J. Teo, A. Ghani, and M. A. Almaiah, "A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5112375, 14 pages, 2022.

Research Article

A Neighborhood and Machine Learning-Enabled Information Fusion Approach for the WSNs and Internet of Medical Things

Zard Ali Khan,¹ Sheneela Naz,¹ Rahim Khan ,^{2,3} Jason Teo ,² Abdullah Ghani,² and Mohammed Amin Almaiah ⁴

¹Department of Computer Science, Comsats University, Islamabad, 45550, Pakistan

²Faculty of Computing and Informatics, University of Malaysia Sabah, Malaysia

³Department of Computer Science, Abdul Wali Khan University Mardan, Pakistan

⁴Department of Computer Networks and Communications, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, Saudi Arabia

Correspondence should be addressed to Rahim Khan; rahimkhan@ums.edu.my

Received 30 January 2022; Revised 3 March 2022; Accepted 8 March 2022; Published 11 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Zard Ali Khan 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.

Data redundancy or fusion is one of the common issues associated with the resource-constrained networks such as Wireless Sensor Networks (WSNs) and Internet of Things (IoTs). To resolve this issue, numerous data aggregation or fusion schemes have been presented in the literature. Generally, it is used to decrease the size of the collected data and, thus, improve the performance of the underlined IoTs in terms of congestion control, data accuracy, and lifetime. However, these approaches do not consider neighborhood information of the devices (cluster head in this case) in the data refinement phase. In this paper, a smart and intelligent neighborhood-enabled data aggregation scheme is presented where every device (cluster head) is bounded to refine the collected data before sending it to the concerned server module. For this purpose, the proposed data aggregation scheme is divided into two phases: (i) identification of neighboring nodes, which is based on the MAC address and location, and (ii) data aggregation using k -mean clustering algorithm and Support Vector Machine (SVM). Furthermore, every CH is smart enough to compare data sets of neighboring nodes only; that is, data of nonneighbor is not compared at all. These algorithms were implemented in Network Simulator 2 (NS-2) and were evaluated in terms of various performance metrics, such as the ratio of data redundancy, lifetime, and energy efficiency. Simulation results have verified that the proposed scheme performance is better than the existing approaches.

1. Introduction

Internet of Things (IoT) consist of numerous sensor nodes, devices, and server(s), which are either deployed randomly or in deterministic fashion to probe the physical phenomena after a defined time interval [1]. However, due to their dense deployments, these networks often generate a huge volume of data, mostly redundant or duplicate, which is needed to be refined before the transmission activity is initiated either by a node or cluster head (CH) or server module. As a result, the ratio of the transmitted packets is increased, which is directly proportional to the lifetime of operational IoT networks. Data aggregation is

one of the common approaches that is used to minimize the ratio of redundancy in data captured by sensor nodes residing in closed proximity [2, 3]. Furthermore, the transmission of these redundant data set without proper refinement leads to numerous challenges, that is, wastage of resources such as bandwidth, on-board battery, and congestion throughout the IoT networks [4]. Machine learning-based approaches are assumed among the promising techniques, which are utilized to address these issues preferably with the available resources and infrastructure.

Generally, in Internet of things and resource-limited networks, communication activity is assumed as more

energy starving than processing. Therefore, data captured by various sensor nodes should be refined, that is, the minimum possible ratio of the duplicate data or packets, before transmitting it to the intended destination module (i.e., base station or sink in this case). The data refinement process is carried out by an ordinary node (if WSNs are homogeneous) or CH or both if the underlined WSN is heterogeneous [5]. In both cases, the ultimate goal is to refine raw data captured by sensor nodes, which are called data aggregation or fusion. Thus, the size of the captured data set is reduced with a minimum possible ratio of information loss. Additionally, the number of packets needed to be transmitted by various source nodes is minimized by avoiding duplicate transmissions and, thus, enhances the overall lifespan of both individual node and the entire WSNs [6, 7]. In the literature, various data aggregation and fusion mechanisms are presented to address to aforementioned issue; however, the majority of these approaches are focused on how to control the communication cost as it is assumed as the main consumer of the on-board battery in the WSNs infrastructures [8, 9]. Apart from it, these approaches rely on the duplicate-insensitive functions, which are directly proportional to the duplicate value elimination time [10]. Additionally, the majority of the existing approaches are not sensitive to the duplicate or redundant data values and outliers, which are generated due to malfunctioning of an ordinary node, in the operational WSNs. A periodic clustering approach was proposed where the main objective was to improve transmission efficiency and eliminate data redundancy. This approach is based on two tiers: one is to eliminate redundancy in every period from sensing data if member nodes and the other is for one-way ANOVA model using the k -mean algorithm with three statistical tests to eliminate data redundancy from the participant that make redundant data sets [11]. These approaches are designed according to the resource-constrained nature of the sensor nodes and CH as a slightly higher computation and space overheads are energy starving. Therefore, data aggregation approaches designed for the WSNs should be energy and performance efficient [12]. In addition to it, the accuracy of the captured data is one of the common issues associated with the WSNs, which is usually due to the resource-constrained nature of member devices or sensor nodes [13]. Therefore, the design and development of the energy and performance efficient data aggregation approaches is still an open challenge both for the researcher and organizations. In this paper, energy- and performance-efficient data aggregation approach is presented to resolve the aforementioned issue, that is, duplication data value, with the available resources in the operational WSNs. For this purpose, a simplified distance measure is used to enable a dedicated CH to refine the collected data by eliminating duplicate data values. Furthermore, the proposed data aggregation scheme bound every CH to match data sets of a neighboring node only as it is quite likely that data captured by node reside in closed proximity are highly correlated than other nodes. For this purpose, the proposed data aggregation scheme is divided into two phases, which are given below.

- (1) Identification of neighboring nodes, which is totally based on their MAC address and location information.
- (2) Apply data aggregation using k -mean clustering and Support Vector Machine (SVM) algorithms.

The proposed data aggregation scheme has restricted the concerned CH to refine data sets of those neighboring node that reside in vicinity (i.e., able to communicate directly with each other). The main contributions of this research work are as given below.

- (1) A MAC-enabled mechanism for the identification of neighboring nodes
- (2) A smart data aggregation approach for the resource's constraint networks, such as IoTs
- (3) Neighborhoods-enabled algorithm to avoid unnecessary matching of data values in the data aggregation process

The remaining of this paper is organized as follows. In Section 2, a brief but comprehensive literature review of data aggregation approaches is presented whereas Section 3 depicts a detailed explanation of the proposed scheme (i.e., neighborhood-enabled data) aggregation, which is based on cluster-based networking infrastructure. In Section 4, the proposed data aggregation algorithm with proper mathematical background is presented. In subsequent section, simulation results of proposed and existing approaches in terms of various performance metrics are presented. Finally, concluding remarks and future direction are provided.

2. Related Work

Various approaches have been proposed in the literature for cluster-based data aggregation, such as k -mean clustering-based, Euclidean distance, cosine distance, one-way ANOVA model, Jaccard function, analysis of variance, and so on. The authors [14] presented the EK-means clustering approach for classification of data set to reduce the volume of sending data over the network. The approach is based on two steps: (i) elimination of data redundancy at sensor level using Euclidean distance measure and (ii) group similar data sets or values and, thus, minimizing the number of packets required to be sent using the EK-means approach. The author [15] has proposed a modified k -means approach to eliminate data redundancy and enhance the sensor network lifetime and forward minimum possible packets to the intended destination module that is the base station in this case. The authors [16] have proposed a k -mean clustering data aggregation approach to eliminate irrelevant data from the sensor network.

This approach works in three steps: (1) check similarity in sensor node level; (2) the sensor node data convert into groups using the k -mean algorithm; and (3) finally, check the human activity with SA score using Euclidean distance between cluster and centroid of data to decide whether to send it or not to the concerned sink.

The author [17] presents a cluster-based periodic sensor network (CPSN) data aggregation approach with aims to eliminate data redundancy and analyze the performance of data latency, accuracy, and energy consumption. This approach consists of two phases: local aggregation and cluster-based aggregation. In CH level, use three methods of distance functions, one-way ANOVA model, and set similarity function. The authors [18] have proposed a cluster-based data aggregation approach using the Comb-Needle model, which aims to minimize communication cost and energy consumption. The authors [19] have proposed a clustering approach in Underwater Sensor Network (UWSN) based on aggregation with Euclidean distance, aiming at data redundancy and analyzing network throughput and energy consumption. Author [11] proposed a periodic clustering approach for underwater in WSNs with the objective of efficient transmission and eliminate data redundancy. This approach is based on two tiers: one to eliminate redundancy in every period from sensing data if member nodes and the other for one-way ANOVA model using the k -mean algorithm with three statistical tests to eliminate data redundancy from the participant that make redundant data sets. The authors of [20] focuses on increasing network lifetime and reduce data transmission. For that, the two-phase protocol was proposed; in the first phase eliminate similarities between sensor nodes data, and in the second phase use distance function to find similarities between sets. This paper focused on decreasing data transmission that sensed by sensor nodes. For that, the authors were proposed a two-layer framework; the first layer divides nodes into clusters and the second layer is full sampling layer where data aggregation was fully performed that minimize the energy consumption of the network [21]. The authors proposed an efficient in-network approach for overcoming the whole transmitting data from member sensor nodes via CH to the base station; for that two-phase scheme was proposed, where in the first phase captured reading goes to an appropriate stratum and the range of stratum is decided; in the second phase two condition are checked, if the reading is minimum, then its compared with previous minimum reading where the smaller value will be next minimum value of stratum, and it should be minimum from new reading, the second condition is for maximum 3 value, the new value of reading compared with old maximum value and the greater value will be next maximum value for stratum [22].

In [23], the authors have proposed a two-level approach for reducing temporal and spatial data redundancy from WSNs to enhance network lifetime and data reliability. The first level of the approach works on the end-node using the Kalman filter and the second level works on the base station using sink level algorithms. However, the energy consumption is high from existing schemes when compared and also no comparative analysis was done among all approaches. The authors have proposed a spatial correlation approach for the elimination of data redundancy from WSNs. The approach works on two levels: source and aggregator. The source level works based on similarity function and the aggregator works based on correlation technique. However, the loss of data is higher than other

existing approaches when compared, and also no comparative analysis was done by the authors [24]. This study has proposed a data aggregation approach that uses less number of limited resources of sensor nodes to reduce data redundancy from WSNs. The approach works based on two levels: sensor node level and cluster head. The first level works based on Exponential Moving Average and threshold-based mechanism. The second level works based on an extended version of Euclidean distance. However, the proposed scheme shows high energy consumption when compared with other existing schemes [25]. The author has proposed a two-level data aggregation mechanism for WSNs to eliminate data redundancy to enhance network lifetime and save energy consumption. In the first level, cluster member nodes minimize data redundancy and also send error-free data to concerned CH. In the second level, k -mean algorithm is used for data aggregation. However, no comparative analysis was done [26].

The authors work on reducing the redundant data and maintaining the integrity of data to send minimum data to the base station or final user and to achieve the above goals; the authors has used two phases method first local aggregation and the second is an aggregators level. Local aggregation used a link function to measure frequency and delete similarities. The aggregators level used the Jaccard similarity function for further aggregation [27]. The authors of the paper focused on increasing network lifetime and minimizing data redundancy for that proposed an Enhanced Clustering Hierarchy (ECP) approach. The objective of ECP is to overlap member nodes and use neighbors sleeping walking approach for minimizing data redundancy [28]. Redundant data use the network resources and decline the network performance by the congestion increasing. Due to the speedy growth of Internet data, various data redundancy methods have been proposed in recent years [29].

Many existing methods provided an appropriate solution to enhance the performance of the network by eliminating data redundancy from the network. It has been generally agreed that data redundancy elimination offers a huge advantage in practice. Usually, the advantage of eliminating the data redundancy is the enhanced network performance in terms of increasing throughput of the network and the decreasing end-to-end delay. However, the proposed techniques are not effective for eliminating redundant and irrelevant data from the network. For the above problem, we proposed a neighborhood-enabled scheme that solves related issues with resource-constrained network.

3. Proposed Neighborhood-Enabled Data Aggregation Approach for the IoTs

In this section, a detailed description of the proposed Neighborhood-Enabled and Cluster-Based Data Aggregation Scheme (NCDAS) is presented, which is designed specifically for the wireless sensor networks. The proposed approach bounds every CH module to refine data values, preferably through a neighborhood based aggregation mechanism, which are captured by the member devices. Actually, the proposed aggregation approach is based on the

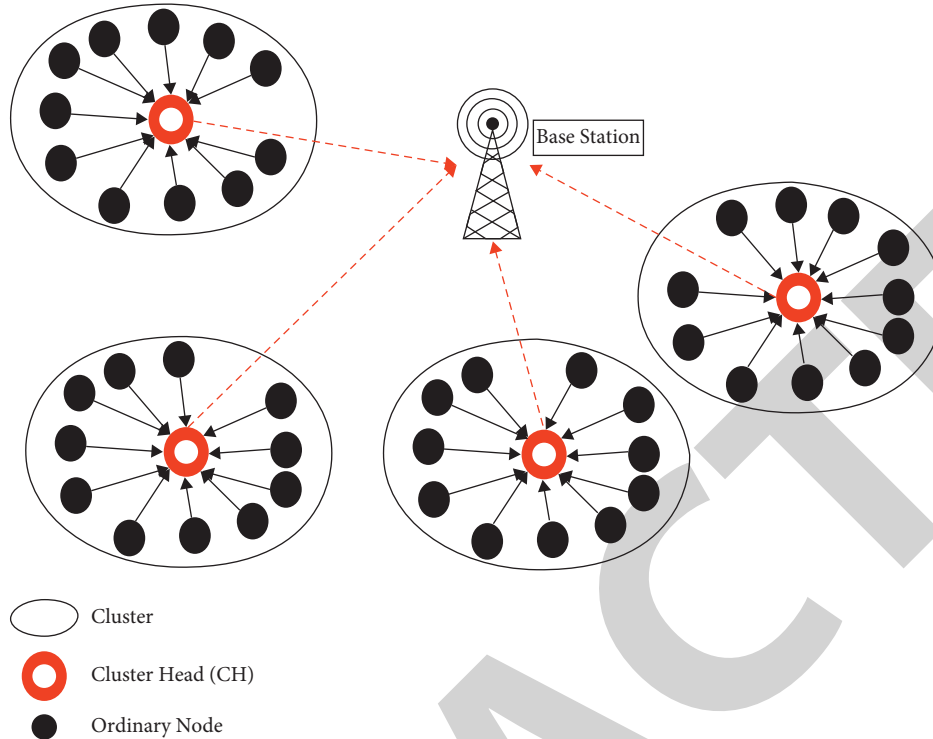


FIGURE 1: Proposed network topology for data aggregation.

usual perception about member devices in IoTs that similarity index in data values captured by neighboring nodes is very high than other nodes. Therefore, the proposed scheme forces every CH module to compare data values of the nodes that are deployed in closed proximity in the IoTs. A detailed description of this mechanism is provided below.

3.1. Network Model. Hierarchical wireless sensor networks are used where member nodes C_i are bounded to communicate via the respective CH module preferably those which are deployed in the coverage area with an approximate ratio of 1 : 20, as shown in Figure 1. Additionally, as the deployment process of WSNs is random, therefore, an imbalanced clustering approach is adopted where it is not necessary that every CH has equal number of member devices or nodes in the operational WSNs. Apart from it, Ad-Hoc On-Demand Distance Vector (AODV) protocol is used for the transmission of packets even if multiple nodes are eager to communicate simultaneously with the intended CH module. Likewise, User Data-Gram Protocol (UDP) is used as data communication protocol. Omnidirectional antenna model is used, whereas power consumption of packets transmission and receiving is 0.860 W and 0.5 W, respectively. Furthermore, packet size is 512 bytes and every member node is assumed to rely on it on-board battery.

3.2. Overview of the Proposed Scheme. Initially, every node has to capture data value after a defined time interval (i.e., 30 seconds in this case). As soon as data value is captured, it is transmitted to the intended destination devices, that is,

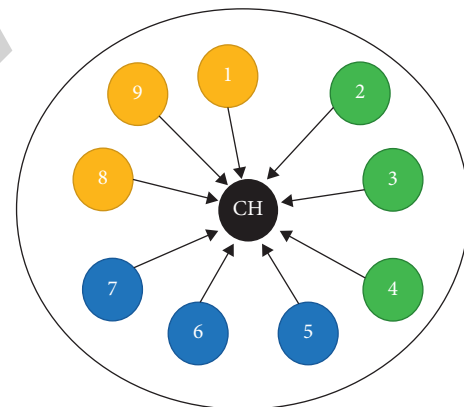


FIGURE 2: Aggregation at CH.

nearest CH module in the proposed data aggregation and communication approach. Since every member node is assumed as a source device, then it is highly likely that CH module receives data values from multiple member nodes after the defined interval of time. Furthermore, CH module is bounded to transmit these data values to a centralized unit (i.e., server module or base station). However, before activation of the transmission activity, CH module needs to refine these values and checks it for possible noise and redundant data values. Generally, in IoTs, sensor nodes that are deployed in close proximity capture similar data values, and transmitting these redundant values is not only a wastage of resources but creates other problems, such as congestion and collision. Therefore, CH module is responsible for eliminating possible redundant data values and

sends a refined version to the base station module. The proposed neighborhood-enabled data aggregation technique not only minimizes the ratio of redundant data values but at the same time avoids excessive matching of data values. For example, assume that CH module has received data values from nine member nodes as shown in Figure 2. In this figure, sensor node 1 has two closed neighbors, that is, sensor nodes 2 and 9. Generally, it is highly likely that data values captured by sensor node 1 are similar (approximately) to the data values captured by sensor nodes 2 and 9, which is due to the fact that these nodes are deployed in close proximity. However, data values captured by sensor node 6 are likely to be different from those of sensor node 1 or we can say similarity index between capture data values of sensor nodes 1 and 6 is very low or at minimum possible level in the IoTs. Therefore, matching captured data values of these two nodes, i.e., sensor nodes 1 and 6, is not only a wastage of time, as similarity index is very low, but seems not feasible as well as far as resource-limited nature of the IoTs is concerned. Thus, the proposed neighborhood-enabled data aggregation scheme not only refines data values captured by various member devices but at the same time avoids excessive matching or processing of irrelevant data in the IoTs.

3.3. Neighborhood Discovery Phase. In this phase, a detailed description of the mechanism that is adopted in the proposed data aggregation approach to find neighboring devices or nodes in the operational IoTs is provided. For this purpose, every CH module is needed to broadcast a message with a hop count value equal to one in the payload. This message is received by those devices or nodes, which are deployed in close proximity or neighborhood to the concerned CH module in the IoTs. These nodes update or modify the received message parameters and rebroadcast them as soon as their backoff timer is expired. Backoff time is used to ensure a collision-free transmission of data values, preferably in scenarios where multiple devices or nodes are interested to communicate simultaneously. Backoff time is a random variable and is computed using the following equation.

$$\text{Backoff timer} = \text{rand}(20 - 1000). \quad (1)$$

For example, if sensor nodes 4 and 5 are interested in sending an updated version of the received message, then it is highly likely that their packets will collide and retransmission will be required. However, if backoff timer is used, then it is highly likely that waiting time interval of these devices is different, which results in successful transmission of the updated packet in the IoTs. This mechanism is repeatedly applied by every CH module and sensor nodes until each and every device has collected information about neighboring devices in the IoTs. In addition to it, the proposed neighborhood-enabled scheme forces every sensor node to share its neighborhood information with the concerned CH module as well. This information is used by the respective CH module to decide which nodes data is needed to be matched to possible ratio of the redundant data values as shown in Figure 2. Furthermore, in order to understand

the working mechanism (that is refinement of the captured data values) of the proposed neighborhood-enabled data aggregation scheme, a simplified workflow diagram is presented in Figure 3, where every step is depicted clearly. Every sensor node is assumed as a source device and is bounded to capture data values by interacting directly with the underlined phenomenon. Once the data is captured, then it is transmitted to the concerned CH module, preferably CH deployed in the coverage area of the transceiver module. As soon as the captured data values are received by the concerned CH module probably from different sources, then refinement or aggregation activity is initiated as shown in Figure 3. Data values of neighboring nodes are matched with each other to find redundant data values and eliminate them (if any exist). This mechanism is repeatedly applied to CH module after a defined time interval, and CH module sends a refined version, preferably with minimum possible redundant data values, to the concerned base station or server module. It is to be noted that the concerned CH module either sends the refined data directly to the respective base station or through multihop communication (if direct communication is not feasible). Apart from it, two data values particularly from different sources are assumed as similar if their difference or distance is less than a defined threshold value, that is, 0.1 in the proposed model, and it is computed using Euclidean distance measure as shown in the following equation, where p and q are two points.

$$d(p, q) = \sqrt{(q - p)^2}. \quad (2)$$

In the proposed neighborhood-enabled data aggregation approach, k -mean clustering is used to divide the deployed IoTs into clusters where it is not necessary that every cluster has similar member nodes (i.e., imbalanced clustering mechanism). In addition to the Euclidean distance measure, other distance measures are used to thoroughly examine their efficiency particularly from execution time, accuracy, and data refinement perspective. We have observed that support vector machine (SVM) is the best possible solution as far as efficiency of the proposed scheme is concerned in the IoTs.

4. Proposed Neighborhood-Enabled Data Aggregation Algorithm

The proposed data aggregation algorithm has two phases, that is, (i) neighborhood discovery and (ii) refinement of data using k -mean clustering algorithm and SVM. It is to be noted that this algorithm is designed to be executed on the concerned cluster head module and is not feasible for the ordinary nodes due to their limited processing power capacities in the IoTs. After the deployment process of the IoTs, every node is bounded to become member of the nearest possible CH module in the IoTs, where ordinary nodes and CH modules are represented as $N_n = \{0, 1, 2, 3, N_n\}$ and $k = \{k_1, k_2, \dots, k_n\}$, respectively. Every CH module is needed to broadcast a message with a hop count value equal to one in the payload. This message is received by those devices or nodes, which are deployed in closed proximity or

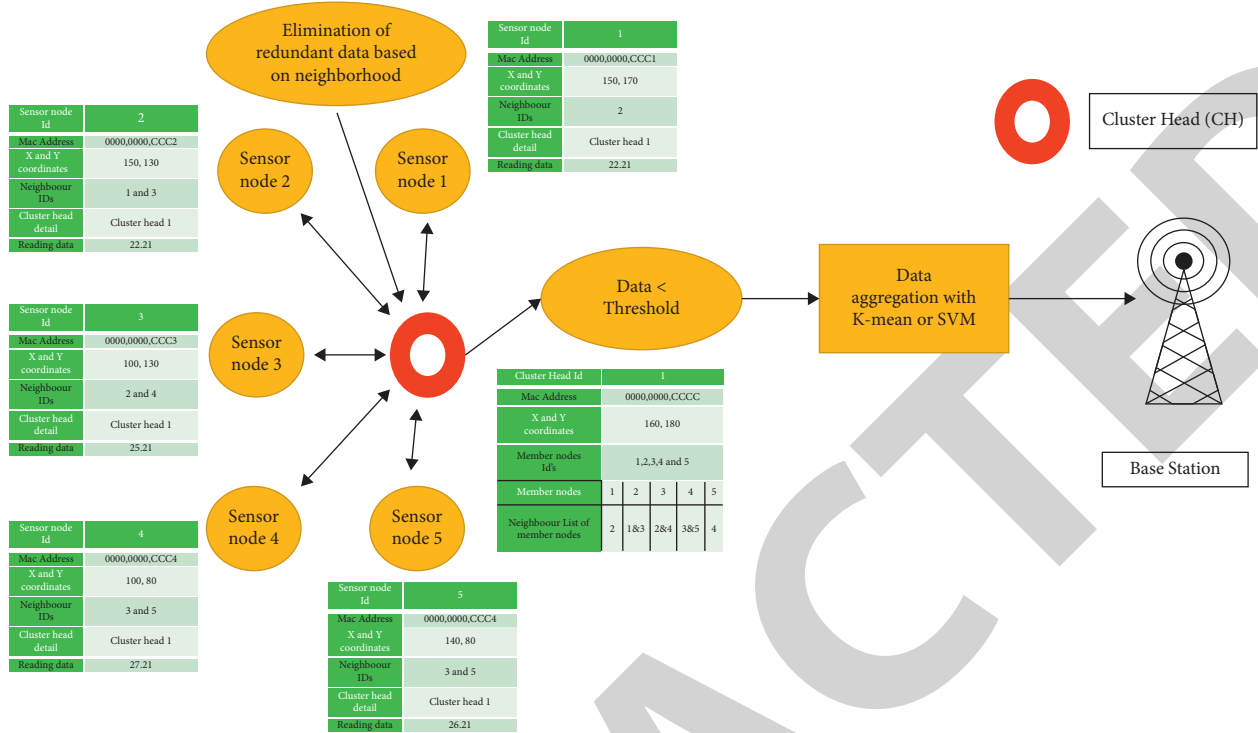


FIGURE 3: Proposed scheme for data redundancy elimination.

neighborhood of the concerned CH module in the IoTs. These nodes update or modify the received message parameters and rebroadcast them as soon as their backoff timer is expired. Backoff time is used to ensure a collision-free transmission of data values preferably in scenarios where multiple devices or nodes are interested to communicate simultaneously. Backoff time is a random variable.

4.1. Neighborhood Identification Mechanism for Ordinary Nodes in the IoTs. In this section, a sophisticated mechanism to find neighboring devices or nodes of a particular member node is presented. As described in CH neighbor discovery section, every member node is forced to broadcast an updated version of message that is received from the nearest CH module in the IoTs. This message is not only received by the respective CH module, but it is also received by those nodes that are deployed in direct coverage area of the concerned node. For example, in Figure 2, message broadcasted by sensor node 8 is received by sensor nodes 7 and 9 as well in addition to the respective CH module. Therefore, these nodes assumed that sensor node 8 resides in close proximity and, thus, it is added to the neighboring node class. This mechanism is repeatedly applied by each and every sensor node in the operational IoTs. In order to further clarify, the proposed neighborhood-enabled aggregation approach allows CH module, as shown in Figure 2, to match data values of node 8 with the captured values of green-color nodes only. Similarly, for sensor node, data values of yellow nodes are matched against each other, whereas existing algorithms or techniques compare data values of one node against all possible

member nodes in a given cluster. For cluster presented in Figure 2, CH is bounded to compare data values of sensor node one with captured data values of every other member node in this cluster (i.e., 2 to 9), which is time-consuming and costly.

4.2. Refinement of the Captured Data Using Euclidean Distance Measure. As described above, k -mean clustering is used to impose a hierarchical structure on the deployed nodes, preferably an imbalanced clustering approach. Additionally, k -mean clustering is used to refine the captured values that is received from multiple sources (i.e., sensor nodes in this case). To minimize the possible ratio of data redundancy, various distance measures (e.g., Euclidean, SVM, and k -mean clustering) have been utilized in the proposed neighborhood-enabled data aggregation mechanism. These methods are not only used to minimize duplicate data values but are equally utilized to improve lifetime of the underlined IoTs particularly with available resources. Sensor nodes are bounded to send data values to the concerned CH after a defined time interval, that is, sampling rate, in the IoTs. However, a tedious and time-consuming task for CH module is the identification of neighboring nodes, particularly whose data values are needed to be matched, that is necessary to avoid unnecessary comparison in the operational IoTs. For this purpose, Euclidean distance measure with a feasible threshold value is used to identify sensor nodes that reside in close proximity [30]. Thus, if distance between two member nodes is greater than the defined threshold value, then these nodes are assumed as nonneighbors and CH is forced to neglect these nodes while

performing refinement of data values of the concerned member node or device in the IoTs (see Algorithm 1).

4.3. Data Aggregation Using k -Mean and SVM Algorithms. As soon as the neighborhood discovery phase is completed, then every CH module has to eliminate (if possible) or reduce the possible ratio of the duplicate data values in the IoTs. The collected data is refined by the respective CH module through the neighborhood-enable aggregation approach as described above. However, it is highly likely that certain ratio of duplicate data values still exists even in the refined data set. Therefore, the proposed scheme has adopted two well-known algorithms, that is, k mean clustering and SVM classification, to further refine the underlined data sets by elimination duplicate data values. As both of these algorithms are computationally expensive, therefore only CH modules are bounded to apply these algorithms to the refined data sets to further improve accuracy and precision of the underlined decision support system. k -mean is a well-known clustering algorithm that has been proved as an effective way to find the redundant values in a given data set particularly those generated by sensor nodes in the IoTs. The key objective of the k -mean clustering algorithm is to create groups of clusters or data sets in such a way that data sets in the same cluster are very similar and data sets in different clusters are quite different. The key idea of k -means clustering is to organize k centroids in each cluster. The initial step is to take every point belonging to a specified data set and associate it with the closest centroid. The initial step is completed when no point is incomplete, and the initial grouping is finalized. After the first step, new k centroids are needed to be calculated for every cluster sequentially. When we get these k new centroids, a new binding has to be completed between the similar data set points and the closest new centroid. A loop has been made, and as a result of a loop, the k centroids change step by step their location until no further changes are done. After having recognized the final clusters that have redundant data sets, the CH removes redundancy from that cluster to decrease the amount of data transferred to BS. SVM is a machine learning algorithm

that is used for regression or classification problems. However, SVM is typically used in classification challenges. In the context of our paper, we used SVM for classification purposes that divide nodes into two classes: redundant and nonredundant. The SVM classifications are primarily based on the following equation.

$$F(X) = T * C + N, \quad (3)$$

where N is ID of sensor node, C is a constant value, and T is the threshold value.

$$\begin{aligned} H1 &= \{F(X) \leq 0\} \\ H2 &= \{F(X) > 0\}. \end{aligned} \quad (4)$$

$H1$ and $H2$ are two hyperplanes, where one is greater than zero and another is less than or equal to zero. Actually, we have utilized equation (3) to separate the captured data into two classes, where $1 > 0$ and $1 \leq 0$. Let us assume three support vectors (i.e., S_1 , S_2 , and S_3). By integrating coordinates of the sensor nodes as described above, we get

$$S_1 = \begin{pmatrix} X1 \\ Y1 \end{pmatrix}, S_2 = \begin{pmatrix} X2 \\ Y2 \end{pmatrix} \text{ and } S_3 = \begin{pmatrix} X3 \\ Y3 \end{pmatrix}. \quad (5)$$

We assume bias value is 1.

$$S_1 = \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix}, S_2 = \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} \text{ and } S_3 = \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix}. \quad (6)$$

Now, we take three parameters for the above vectors.

$$\begin{aligned} \alpha_1 \vec{S}_1 \vec{S}_1 + \alpha_2 \vec{S}_2 \vec{S}_1 + \alpha_3 \vec{S}_3 \vec{S}_1 &= 1 \\ \alpha_1 \vec{S}_1 \vec{S}_2 + \alpha_2 \vec{S}_2 \vec{S}_2 + \alpha_3 \vec{S}_3 \vec{S}_2 &= 1 \\ \alpha_1 \vec{S}_1 \vec{S}_3 + \alpha_2 \vec{S}_2 \vec{S}_3 + \alpha_3 \vec{S}_3 \vec{S}_3 &= 1. \end{aligned} \quad (7)$$

So, we get the value of \vec{S}_1 , \vec{S}_2 , and \vec{S}_3 .

$$\begin{aligned} \vec{S}_1 &= \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix}, \vec{S}_2 = \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix}, \vec{S}_3 = \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \\ \alpha_1 \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix} \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix} &= 1 \\ \alpha_1 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} &= 1 \\ \alpha_1 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} &= 1. \end{aligned} \quad (8)$$

```

(1) Input: Analysis of neighbors nodes in IoTs ( $N_v \in \text{IoTs}$ )
(2) Output: Elimination of redundant data based on neighbors nodes in IoTs ( $N_v \in \text{IoTs}$ )
(3) begin
(4) Ordinary nodes  $N_n = \{0, 1, 2, 3, N_n\}$  with MAC address value and location
(5) Clusters  $K_n = \{K_1, K_2, \dots, K_n - 1\}$ 
(6)   While every  $k_i \in \text{IoTs}$  do
(7)     Generate a message
(8)     Set a join request value 1
(9)     Broadcast message
(10)  end while
(11) While every Node(i)  $\in \text{IoTs}$  do
(12) If  $\text{RSSI}(k_i \geq K_i + 1 \dots n)$  then
(13)   Update the message
(14)   Set destination  $K_i$ 
(15)   Backoff Timer = rand(20–1000 milliseconds)
(16)   Re-broadcast message
(17)   end if
(18) end while
(19) While Nodes  $\in k_i$  do
(20) Calculate Euclidean distance( $E_d$ ) among all nodes
(21)   If ( $E_d$  of Node(i) and (j) is  $\leq t_d$ ) then
(22)     Both nodes I and j are neighbors
(23)     Check data redundancy among
(24)     Eliminate redundant data captured by node
(25)   else
(26)     Both nodes I and j are not neighbors
(27)   end if
(28) end while
(29) While Node(i)  $\in K_i$  do
(30)   If Node(i).data  $\leq$  threshold then
(31)     Aggregate data using k-mean or SVM
(32)     Send aggregate data to the Base station
(33)   else
(34)     Discard data
(35)   end if
(36) end while

```

ALGORITHM 1: NCDAS Algorithm.

After simplifying the above equation, the values of α_1 , α_2 , and α_3 are computed for classes discriminates; the following equation is used.

$$\vec{F} = \sum_{i=1}^n \alpha_i \vec{S}. \quad (9)$$

Put values in equation (4)

$$\alpha_1 \begin{pmatrix} X1 \\ Y1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} X2 \\ Y2 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} X3 \\ Y3 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 X1 + \alpha_2 X2 + \alpha_3 X3 \\ \alpha_1 Y1 + \alpha_2 Y2 + \alpha_3 Y3 \\ \alpha_1 + \alpha_2 + \alpha_3 \end{pmatrix}. \quad (10)$$

$$\text{in } \begin{pmatrix} \alpha_1 X1 + \alpha_2 X2 + \alpha_3 X3 \\ \alpha_1 Y1 + \alpha_2 Y2 + \alpha_3 Y3 \\ \alpha_1 + \alpha_2 + \alpha_3 \end{pmatrix}.$$

$\alpha_1 + \alpha_2 + \alpha_3$ is the threshold value that is used for separation of classes in equation (3).

5. Experimental Results

In order to evaluate and verify effectiveness of the proposed NCDAS, extensive simulations in terms of various performance metrics have been performed, such as packet delivery ratio, network lifetime, energy consumption, throughput,

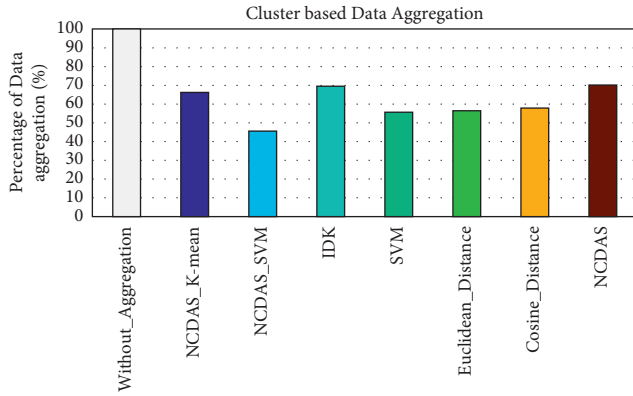


FIGURE 4: Data aggregation at CH of the proposed versus existing state-of-the-art schemes.

and end-to-end delay. For comparison, we have compared the performance of the proposed data aggregation approached against the latest field proven approaches, such as [16, 17, 31, 32]. These algorithms were implemented in network simulator 2 (NS-2) using similar topological infrastructures. Apart from this, these approaches were thoroughly checked for varying sensor nodes, threshold values, and readings. Various parameter used in the simulation setup are presented in Table 1.

5.1. Data Aggregation at CH. Cluster head is assumed as one of the core components of the heterogeneous WSNs as majority of the processing is carried out at this level. Therefore, it is mandatory to evaluate the performance of the proposed data aggregation scheme at this level. In the proposed data aggregation approach, CH is not only responsible for refining the capture data values of its member devices but at the same acting as destination devices for member devices as well. Simulation results show that the performance of the proposed data aggregation scheme is far more better than the existing state-of-the-art approaches, as depicted in Figure 4. From Figure 4, we have observed that k -mean and SVM with embedded neighborhood-enabled (proposed) approach aggregates or refine the captured data values up to 66.23% and 45.28%, respectively, whereas the aggregation ratio of IDK, SVM, Euclidean distance, cosine distance, and NCDAS approaches are 69.56%, 55.56%, 56.24%, 57.56%, and 70.23%, respectively. However, 100% data is sent to the BS if no aggregation activity is performed. In terms of data redundancy or duplication, the proposed neighborhood-enabled approach eliminates 98.4% of the duplicate data values.

5.2. Energy Efficiency at CH. Energy efficiency is assumed as an essential component to evaluate the performance of the newly developed algorithm, specifically those that are designed for the wireless sensor networks. It is due to the fact that member devices in these networks solely rely on the on-board batteries; thus, prolonged lifetime of both individual device and whole network is highly appreciated by the research community. For this purpose, we have evaluated energy consumption of the proposed and existing approaches at the CH

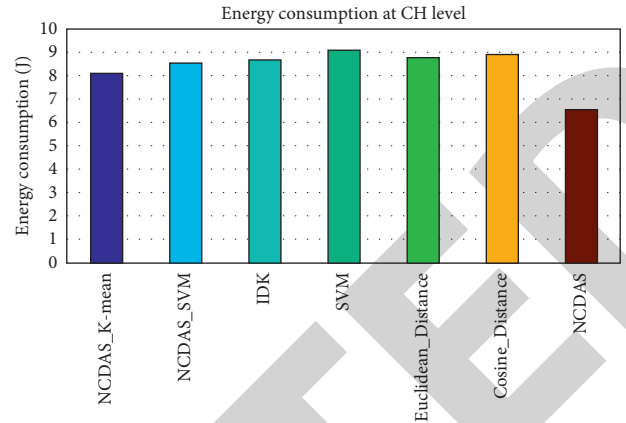


FIGURE 5: Average energy consumption at CH of the proposed versus existing state-of-the-art schemes.

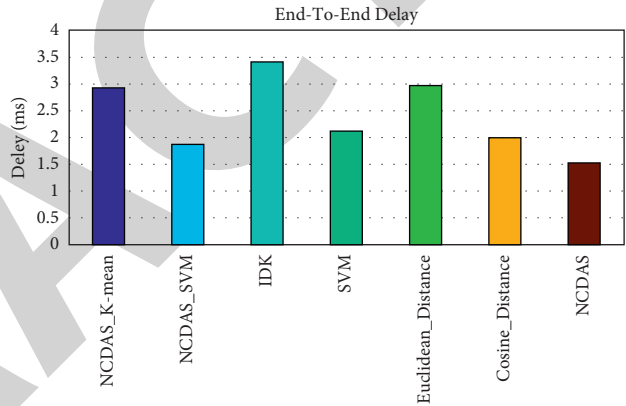


FIGURE 6: End-to-end delay of the proposed versus existing state-of-the-art schemes.

level, which is bounded by the proposed scheme to refine the captured data values of its member devices or sensor before sending it to the gateway module. Simulation results clearly depict that the average energy consumption of the proposed scheme is far better than the existing state-of-the-art schemes particularly at CH level in the IoTs. Furthermore, we have observed that the NCDAS, specifically when not embedded with the k -means and SVM algorithms, consumes 6.55 J, with k -mean and SVM 8.1 8.54 J of the on-board battery, respectively, at CH, whereas IDK, SVM, cosine distance, and Euclidean distance are 8.6789 J, 9.1 J, 8.92 J, and 8.76 J, respectively, as shown in Figure 5. Likewise, when the proposed scheme is integrated with the k -mean to form a hybrid refinement mechanism, then the simulation results show that its performance is according to the expectations, that is, better than existing state-of-the-art approaches. The proposed approach when merged with the k -mean has less energy consumption from a hybrid of proposed scheme and SVM (i.e., 5.15%), from IDK is 6.77%, from SVM is 10.98%, from Euclidean distance is 7.53%, and from cosine distance is 9.19%.

5.3. End-to-End Delay. The end-to-end delay of the NCDAS is 1.532 ms, which is less than from existing schemes and proposed hybrid schemes. The proposed scheme with k -

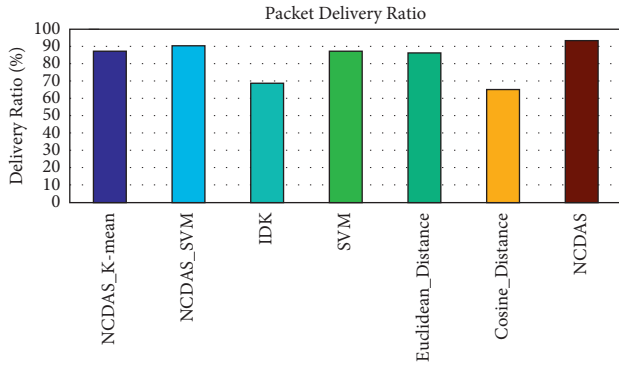


FIGURE 7: Packet delivery ratio of the proposed versus existing state-of-the-art schemes.

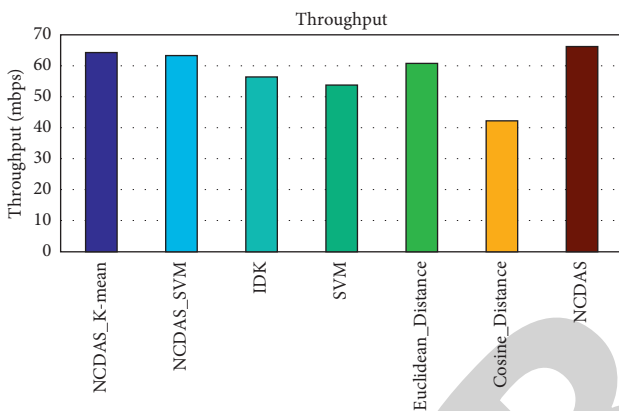


FIGURE 8: Throughput of the proposed versus existing state-of-the-art schemes.

mean has an end-to-end delay of 2.92 ms and the proposed with SVM is 1.87 ms, whereas end-to-end delay of IDK, SVM, Euclidean distance, and cosine distance are 3.41 ms, 2.11 ms, 2.96 ms, and 1.96 ms, respectively. The proposed scheme when integrated with the SVM approach is 35% than proposed with k -mean, 45% from IDK, 37% from Euclidean, 6% from SVM, and 12% from cosine are lesser when compared with these approaches (Figure 6).

5.4. Average Packet Delivery Ratio (APDR). APDR is considered as an important evaluation metric to judge the performance of the newly developed approach using well-known topological infrastructures. To evaluate the performance of the proposed data aggregation approach, specifically in terms of APDR, the proposed scheme along with existing approaches is implemented in realistic environment of IoTs. During the simulations, we have observed that the APDR of the NCDAS is better than the existing state-of-the-art schemes as shown in Figure 7. APDR of the proposed scheme is integrated with k -mean is 87.21% and with SVM is 90.21%, where IDK is 68.58%, SVM is 87.11%, Euclidean distance is 86.24%, and cosine distance is 65.15%. The ratio of the NCDAS is 93.22%, which is good from other schemes. The PDR of proposed scheme with SVM is 3.44% from proposed with k -mean, 31.53% from IDK, 3.5% from SVM,

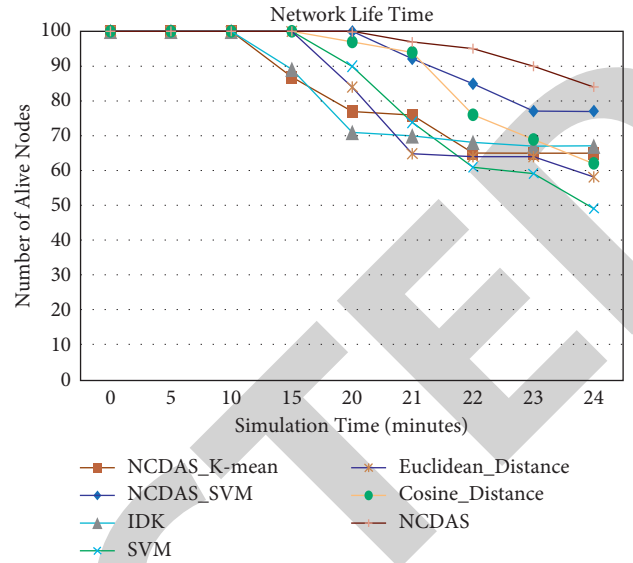


FIGURE 9: Number of alive nodes in proposed scheme versus existing state-of-the-art schemes.

4.60% from Euclidean distance, and 38.46% from cosine distance are higher when compared with these schemes.

5.5. Throughput. The throughput of the sensor network is defined as the ratio of packets that are delivered successfully to the intended destination device (i.e., the base station in this case). Additionally, it is to be noted that we have assumed that IoTs become inactive as soon as the very first node consumes its on-board battery completely. Throughput analysis of the NCDAS and existing schemes is shown in Figure 8, which depicts that the proposed scheme has a relatively higher throughput than the existing state-of-the-art schemes. Likewise, the proposed scheme performance, specifically in terms of throughput, is improved further if it is integrated with the k -mean than the hybrid of the proposed SVM, proposed IDK, proposed SVM, proposed Euclidean, and proposed cosine, respectively. Apart from it, we have observed that a hybrid of the proposed k -mean performs better in terms of packet loss ratio, which is 1.54%. Similarly, when the proposed scheme is integrated with SVM, then packet loss ratio is 12.84%, which is better than that IDK, SVM, and Euclidean i.e., 19.12%, 5.54%, and 52.54%, respectively.

5.6. Network Lifetime. Lifetime is one of the critical and vital evaluation metrics that is used by the research community to examine the performance of the newly developed scheme for the IoTs and other resource-constrained networks. Network's lifetime of the proposed neighborhood-enabled data aggregation and existing state-of-the-art schemes is presented in Figure 9. During the simulation setup of the proposed data aggregation approach, we have observed that the NCDAS approach and the proposed integrated with SVM leads to the minimum possible set of dead nodes (i.e., 16 and 23 in this case). Similarly, the proposed scheme is

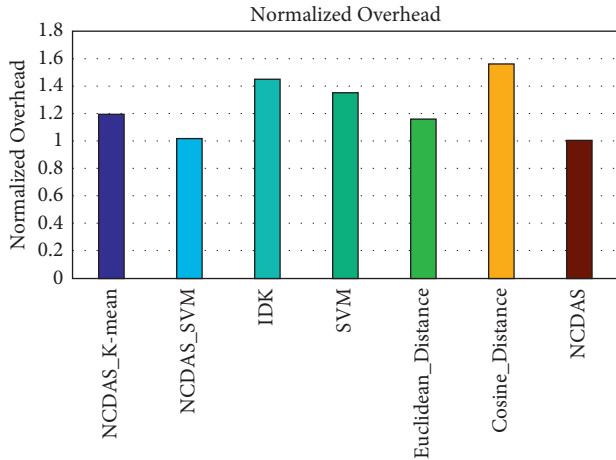


FIGURE 10: Normalized overhead of the proposed versus existing state-of-the-art schemes.

when integrated with the k -mean; then death ratio of member devices is 35. Likewise, for IDK, SVM, and cosine, the ratio of dead nodes is 33, 51, and 42, respectively. Additionally, the proposed with SVM is an ideal solution, which is depicted in Figure 9. Moreover, the proposed scheme when combined with the k -mean, its dead-to-alive node ratio is 34.28%, from IDK is 30.30%, from SVM is 53.06%, from Euclidean distance is 45.38%, and from cosine is 39.47% when these approaches are compared.

5.7. Normalized Routing Overhead. In Figure 10, the normalized overhead of proposed neighborhood-enabled data aggregation and existing schemes is presented. It shows that the NCDAS has less normalized overhead than the existing schemes. The normalized overhead of proposed when integrated with k -mean is 1.19 and with SVM is 1.015, where with IDK is 1.45, SVM 1.35, Euclidean distance 1.16, and cosine distance is 1.56. The NCDAS overhead is 1.001, which is less from other schemes when compared with these schemes.

6. Comprehensive Comparison of the Proposed and Existing State-of-the-Art Schemes

The aim of this comprehensive study is to compare the proposed scheme with existing ones, to find which scheme is best among all. Here, we compare the NCDAS and proposed with SVM and k -mean with four existing schemes. The performance metrics values of all schemes are presented in Table 2. Based on the results of Figure 4, the proposed with SVM is giving the best result in data aggregation, where all evaluation parameters suggest that the proposed with SVM aggregate data accurately among all. In Figures 5 and 8, it is presented that the NCDAS and the proposed with k -mean in terms of energy consumption is the lowest, and in terms of throughput of the network, it is higher among all schemes. Based on the result of Figures 10 and 6, it shows that the NCDAS and the proposed with SVM in terms of normalized routing overhead and end-to-end delay is the lowest among

TABLE 1: Simulation parameters.

Routing protocol	AODV
Number of nodes	100
Network size	1000 × 1000 m
MAC protocol	IEEE 802.11
Antenna model	Omnidirectional antenna
Initial energy	10.0 J
Transmission power	0.860 W
Receiving power	0.5 W
Packet size	512 bytes

all schemes. In Figure 7, the packet delivery ratio is shown in all schemes such that the NCDAS and the proposed with SVM has a good packet delivery ratio among all. The proposed with SVM in terms of network life is good based on results shown in Figure 9; the number of dead nodes during simulation time is smaller among all when compared. The performance metric results of all schemes when compared to that of the proposed scheme have the best results among all aspects.

7. Complexity Analysis of the Proposed and Existing State-of-the-Art Schemes

In this section, we discuss all scheme time complexity, which is shown in Table 3. We consider the time complexity of all schemes, which are used in the data aggregation phase and the reason is that the previous papers only consider the data aggregation algorithm time complexity. The Euclidean and cosine distance has $\Theta(n^2)$ time complexity and IDK have $\Theta(n^2) * \Theta(n)$ [16, 17, 32], where the SVM [31] time complexity is $\Theta(n) * \Theta(3n)$; the proposed with k -mean and proposed with SVM have $\Theta(n) * \Theta(n)$ and $\Theta(n) * \Theta(n) * \Theta(3n)$ time complexity; here in proposed with SVM we did not consider the training timing complexity, only considering the classification time complexity of the work. The NCDAS has $\Theta(n)$; based on the above results, the proposed has the best time complexity among all when compared.

8. Further Discussion

In this section, we give further explanations of our proposed schemes. In the experimental section, we saw the comparison of existing and proposed schemes while applying. The performance of all schemes is evaluated based on performance metrics. From data aggregation and redundancy point of view, the proposed scheme with SVM gives the best result among the other schemes and actual data reduce to 45.28%. The proposed with SVM more reduces data from proposed with k -mean 31.63%, from IDK is 34.90%, from SVM is 18.50%, from a Euclidean distance is 19.48%, from cosine distance is 21.33%, and from NCDAS is 35.536% when these schemes are compared.

In terms of energy consumption, the NCDAS and the proposed with k -mean consume less energy at CH among other schemes; the proposed scheme saves more energy compared to other schemes. The NCDAS saves energy, that

TABLE 2: A detailed comparative analysis of the proposed and existing state-of-the-art schemes.

Schemes	Data aggregation	Energy consumption (J)	End-to-end delay (ms)	Normalized overhead	PDR (%)	Throughput (Mbps)	Dead nodes
NCDAS	70.23%	6.55	1.532	1.001	93.22	66.26	16
Proposed with SVM	45.28%	8.54	1.87	1.015	90.21	63.22	23
Proposed with k -mean	66.23%	8.1	2.92	1.19	87.21	64.22	35
IDK	69.56%	8.678	3.41	1.45	68.58	63.84	33
SVM	55.56%	9.1	2.11	1.35	87.11	53.91	51
Euclidean distance	56.24%	8.92	2.96	1.16	86.24	60.84	42
Cosine distance	57.6%	8.76	1.96	1.56	65.15	42.01	38

TABLE 3: Complexity analysis of the proposed and existing schemes.

Schemes	Time complexity
NCDAS	$\Theta(n)$
Proposed with SVM	$\Theta(n) * \Theta(n) * \Theta(3n)$
Proposed with k -mean	$\Theta(n) * \Theta(n)$
IDK	$\Theta(n^2) * \Theta(n)$
SVM	$\Theta(n) * \Theta(3n)$
Cosine distance	$\Theta(n^2)$
Euclidean distance	$\Theta(n^2)$

is, 30.38% from proposed with SVM, from the proposed with k -mean 23.66%, from IDK 32.48%, from SVM 61.06%, from Euclidean distance and cosine distance are 36.18% and 33.74% when compared with these schemes in terms of energy consumption at CH. From the end-to-end delay point of view, we see that the NCDAS and the proposed with SVM scheme give a good result in terms of the end-to-end delay among others when compared with Euclidean distance, cosine distance, SVM, IDK, and proposed with k -mean schemes. The Euclidean distance and IDK give the highest end-to-end delay from other schemes; the Euclidean from proposed with SVM is 37% where IDK is 45%. We see that SVM and cosine distance also give minimum end-to-end delay from IDK, Euclidean distance, and proposed with k -mean. From the packet delivery ratio point of view, the proposed with SVM and the NCDAS gives the highest delivery ratio among other schemes when these schemes are compared. The PDR of proposed with k -mean and SVM are higher than Euclidean distance, cosine distance, and IDK. The lowest delivery ratio is of cosine distance, which is 65%. From the throughput point of view, the proposed with k -mean has the highest throughput among other schemes. We see that IDK also has good throughput from other schemes in terms of lowest throughput is of cosine distance, which is 42.1 Mbps. The result of SVM, particularly when combined with the proposed scheme, are better than Euclidean distance and SVM as shown in the results section. From the network lifetime point of view, the NCDAS and the proposed with SVM have a good network lifetime among other schemes. We see that the cosine distance and SVM have not a good network lifetime when compared with other schemes. The proposed with SVM is considered the best scheme among all in terms of data aggregation, energy consumption, delay, delivery ratio, and network lifetime. Based on the

above results, our proposed scheme is best for a large and small network. The user can use this scheme in his own interest; in this paper, we did not consider the integrity of data; in general, data are secured and without any loss are sent to the base station.

Although the proposed data aggregation scheme has resolved the challenging issue of the redundant data in the data which is captured by the various sensor nodes deployed in closed proximity of the phenomena, however, a common issue with the proposed approach and other data aggregation approaches is that these approaches have compromised on data loss of the smallest possible portion of the underlined data. Although the proposed approach is very effective in reducing the ratio of duplicate data values, a significant portion of data is lost. Finally, in the proposed data aggregation approach, every sensor node along with cluster head module is bounded to compute or find its neighborhood information, which consumes a significant portion of the available power. Moreover, this process is time-consuming specifically in situations if deployment density of the underlined sensor nodes is very densed.

9. Conclusion

In this paper, we have proposed a neighborhood-enabled and machine learning-enabled data aggregation scheme for the resource-constrained network. In our proposed approach first, we find neighbors in the cluster and then use data aggregation and redundancy check at CH level, where data comes from member sensor nodes that sense the physical environment. When neighbor nodes are found, then data redundancy check only with neighbor nodes and then k -mean and SVM data aggregation algorithms are applied on remaining data, which further clean data. After using data aggregation, the refined data are sent to the base station. We see that our approach extends network lifetime, increases throughput and packet delivery ratio, and decreases energy consumption and the end-to-end delay. In the future, this scheme will be used on real data set for more validation and evaluate this scheme with real-time data aggregation schemes.

Data Availability

The data sets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research work was supported by the Faculty of Computing and Informatics, University of Malaysia Sabah, Malaysia.

References

- [1] B. David Deebak and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Networks*, vol. 97, Article ID 102022, 2020.
- [2] N.-Tu Nguyen, "An efficient minimum-latency collision-free scheduling algorithm for data aggregation in wireless sensor networks," *IEEE Systems Journal*, vol. 12, pp. 2214–2225, 2017.
- [3] M. El Fissaoui, A. Beni-Hssane, and M. Saadi, "Multimobile agent itinerary planning-based energy and fault aware data aggregation in wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 1–11, 2018.
- [4] Q. Wang, D. Lin, P. Yang, and Z. Zhang, "An energy-efficient compressive sensing-based clustering routing protocol for WSNs," *IEEE Sensors Journal*, vol. 19, no. 10, pp. 3950–3960, 2019.
- [5] A. Li, W. Liu, L. Zeng, C. Fa, and Y. Tan, "An efficient data aggregation scheme based on differentiated threshold configuring joint optimal relay selection in WSNs," *IEEE Access*, vol. 9, Article ID 19254, 2021.
- [6] S. Vimal, M. Khari, R. G. Crespo, L. Kalavani, N. Dey, and M. Kaliappan, "Energy enhancement using Multiobjective Ant colony optimization with Double Q learning algorithm for IoT based cognitive radio networks," *Computer Communications*, vol. 154, pp. 481–490, 2020.
- [7] A. L. R. Madureira, F. R. C. Araújo, and L. N. Sampaio, "On supporting IoT data aggregation through programmable data planes," *Computer Networks*, vol. 177, Article ID 107330, 2020.
- [8] F. Xiong, H. Zheng, L. Ruan et al., "Energy-saving data aggregation for multi-UAV system," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 8, pp. 9002–9016, 2020.
- [9] S. Farajzadeh, J. Rezazadeh, R. Farahbakhsh, K. Sandrasegaran, and M. Abbasian Dehkordi, "A survey on data aggregation techniques in IoT sensor networks," *Wireless Networks*, vol. 26, no. 2, pp. 1243–1263, 2020.
- [10] B. Pourghableh and N. J. Navimipour, "Data aggregation mechanisms in the Internet of things: a systematic review of the literature and recommendations for future research," *Journal of Network and Computer Applications*, vol. 97, pp. 23–34, 2017.
- [11] M. Rida, A. Makhoul, H. Harb, D. Laiymani, and M. Barhamgi, "EK-means: a new clustering approach for datasets classification in sensor networks," *Ad Hoc Networks*, vol. 84, pp. 158–169, 2019.
- [12] K. Sarangi and I. Bhattacharya, "A study on data aggregation techniques in wireless sensor network in static and dynamic scenarios," *Innovations in Systems and Software Engineering*, vol. 15, no. 1, pp. 3–16, 2019.
- [13] W. Alghamdi, M. Rezvani, H. Wu, and S. S. Kanhere, "Routing-aware and malicious node detection in a concealed data aggregation for WSNs," *ACM Transactions on Sensor Networks*, vol. 15, no. 2, pp. 1–20, 2019.
- [14] H. Hassan, A. Makhoul, and R. Couturier, "An enhanced K-means and ANOVA-based clustering approach for similarity aggregation in underwater wireless sensor networks," *IEEE Sensors Journal*, vol. 15, pp. 5483–5493, 2015.
- [15] A. K. Idrees, "Distributed data aggregation based modified k-means technique for energy conservation in periodic wireless sensor networks," in *Proceedings of the 2018 IEEE Middle East and North Africa Communications Conference (MENACOMM)*, pp. 1–6, IEEE, Jounieh, Lebanon, April, 2018.
- [16] S. Pattamaset and J. S. Choi, "Irrelevant data elimination based on a k-means clustering algorithm for efficient data aggregation and human activity classification in smart home sensor networks," *International Journal of Distributed Sensor Networks*, vol. 16, Article ID 1550147720929828, 2020.
- [17] H. Hassan, "Comparison of different data aggregation techniques in distributed sensor networks," *IEEE Access*, vol. 5, pp. 4250–4263, 2017.
- [18] M. Shanmukhi and O. B. V. Ramanaiah, "Cluster-based comb-needle model for energy-efficient data aggregation in wireless sensor networks," in *Proceedings of the 2015 Applications and Innovations in Mobile Computing (AIMoC)*, pp. 42–47, IEEE, Kolkata, India, February, 2015.
- [19] K. T.-M. Tran and S.-H. Oh, "UWSNs: a round-based clustering scheme for data redundancy resolve," *International Journal of Distributed Sensor Networks*, vol. 10, no. 4, Article ID 383912, 2014.
- [20] H. Hassan, "A distance-based data aggregation technique for periodic sensor networks," *ACM Transactions on Sensor Networks*, vol. 13, pp. 1–40, 2017.
- [21] H. Lin, W. Wei, P. Zhao et al., "Energy-efficient compressed data aggregation in underwater acoustic sensor networks," *Wireless Networks*, vol. 22, no. 6, pp. 1985–1997, 2016.
- [22] S. R. U. Jan, M. A. Jan, R. Khan, H. Ullah, M. Alam, and M. Usman, "An energy-efficient and congestion control data-driven approach for cluster-based sensor network," *Mobile Networks and Applications*, vol. 24, no. 4, pp. 1295–1305, 2019.
- [23] Z. Yemeni, H. Wang, W. M. Ismael, Y. Wang, and Z. Chen, "Reliable spatial and temporal data redundancy reduction approach for wsn," *Computer Networks*, vol. 185, Article ID 107701, 2021.
- [24] R. Maivizhi and P. Yogesh, "Spatial correlation based data redundancy elimination for data aggregation in wireless sensor networks," in *Proceedings of the 2020 International Conference on Innovative Trends in Information Technology (ICITIIT)*, pp. 1–5, IEEE, Kottayam, India, February, 2020.
- [25] S. R. Jan, R. Khan, F. Khan et al., "Marginal and average weight-enabled data aggregation mechanism for the resource-constrained networks," *Computer Communications*, vol. 174, pp. 101–108, 2021.
- [26] X. Xiao, "An energy-efficient clustering routing protocol based on data aggregation for underwater acoustic sensor networks," in *Proceedings of the Global Oceans 2020: Singapore-US Gulf Coast*, pp. 1–6, IEEE, Biloxi, MS, USA, October, 2020.
- [27] J. M. Bahi, A. Makhoul, and M. Medlej, "A two tiers data aggregation scheme for periodic sensor networks," *Ad Hoc & Sensor Wireless Networks*, vol. 21, pp. 77–100, 2014.
- [28] H. El Alami and A. Najid, "ECH: an enhanced clustering hierarchy approach to maximize lifetime of wireless sensor networks," *IEEE Access*, vol. 7, Article ID 107142, 2019.

Retraction

Retracted: Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] A. Hasan, T. Sharma, A. Khan, and M. Hasan Ali Al-Abyadh, "Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8153791, 8 pages, 2022.

Research Article

Analysing Hate Speech against Migrants and Women through Tweets Using Ensembled Deep Learning Model

Asif Hasan ¹, Tripti Sharma ², Azizuddin Khan ³,
and Mohammed Hasan Ali Al-Abyadh ^{4,5}

¹Department of Psychology, Aligarh Muslim University, Aligarh 202001, India

²IT Department, Maharaja Surajmal Institute of Technology, New Delhi 110058, India

³Department of Humanities and Social Sciences, Indian Institute of Technology, Bombay Powai, Mumbai 400076, India

⁴Mental Health-College of Education, Prince Sattam Bin Abdulaziz University, Alkharj, Saudi Arabia

⁵College of Education, Thamar University, Thamar, Yemen

Correspondence should be addressed to Tripti Sharma; tripti_sharma@msit.in

Received 30 December 2021; Accepted 23 February 2022; Published 10 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Asif Hasan 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.

Twitter's popularity has exploded in the previous few years, making it one of the most widely used social media sites. As a result of this development, the strategies described in this study are now more beneficial. Additionally, there has been an increase in the number of people who express their views in demeaning ways to others. As a result, hate speech has piqued interest in the subject of sentiment analysis, which has developed various algorithms for detecting emotions in social networks using intuitive means. This paper proposes the deep learning model to classify the sentiments in two separate analyses. In the first analysis, the tweets are classified based on the hate speech against the migrants and the women. In the second analysis, the detection is performed using a deep learning model to organise whether the hate speech is performed by a single or a group of users. During the text analysis, word embedding is implemented using the combination of deep learning models such as BiLSTM, CNN, and MLP. These models are integrated with word embedding methods such as inverse glove (global vector), document frequency (TF-IDF), and transformer-based embedding.

1. Introduction

Academic literature and political debates continue to focus on the question of free speech. Hate speech is tolerated in some form or another in many countries [1]. While hate speech and hate crimes have traditionally had devastating effects on individuals and communities, hate speech laws are based on substantive equality. Nazi intentions for the annihilation of the Jewish population were followed by public hatred campaigns, which made hate speech an issue in international law following the Second World War [2]. Social media platforms are concerned about inappropriate user-generated content [3]. Hate speech thrives on social media sites like Twitter because of a lack of accountability and lack of control [4]. Even though social media businesses pay people to censor material, the volume of social media posts is too great for humans to keep track of the concerned

people. Our goal in this work is to develop a system for automatically detecting hate speech against migrants and woman through tweets using the ensembled deep learning model. In this work, deep learning models have been trained to tackle the challenges provided by the competition for the automatic identification of hate speech against migrants and woman [5]. It is time to stop arguing about the value of women's contributions to society. According to a recently published UNDP report, women's unpaid domestic and care work in Montenegro achieved a predicted value of 122 million euros in the first three months of COVID-19 in 2020 [6]. When it comes to the workplace, over half of all women have experienced a breach of their rights. This is just the beginning; Montenegro has been experiencing a worrying trend of women and girls being subjected to misogynistic attitudes, sexist language, hate speech, bullying, and other forms of intolerance and discrimination. Despite the

abundance of evidence that women and girls make significant contributions to society, we have observed a worrying trend of women and girls being subjected to misogynistic attitudes, sexist language, hate speech, bullying, and other forms of intolerance and discrimination. This can be seen in the public realm and media, particularly in social media conversation. When it appears that women are being used as a political pawn by those with divergent political views and ideologies, it is a deeply troubling development that could undo years of steady progress in raising awareness of the importance of safeguarding universal values such as equality and dignity for all. An outbreak of violence against women directly results from COVID-19's pandemic and has already worsened longstanding disparities in the United States.

Deep learning has done a significant work in the area of sentiment analysis. So, this paper has presented the use of deep learning for analysing the hate speech against migrants and the woman. The application of this work is to identify the people or the users who are posting the messages against the woman and the migrants.

The organisation of the paper is as follows: Section 1 describes the introduction of the proposed work. In section 2 literature survey will be discussed, and section 3 covers the methodology section. Section 4 explains the implementation details with the results, and section 5 defines the conclusion and future work followed by the references in the last section.

2. Literature Survey

The second Global Summit on Religion, Peace and Security was conducted from April 29 to May 1, 2019, by the United Nations (UN). The summit's goal was to produce a Plan of Action to oppose hate speech and protect religious minorities and migrants (United Nations, 2019) [7]. Adama Dieng the UN's Speech Adviser for the prevention of Genocide state that "Political opportunism is fueling an increase in the hate speech." This statement draws the analogies between 1930s Europe and present political climate on the continent [8]. He warned that "big atrocities often begin with little actions and language." Dieng is worried about the far-right parties' political plans in Europe following the 2015 refugee crisis. Many of these political parties have taken advantage of Europe's migrant policies to set the stage for a broader discussion on migration. As a result, they have successfully entered mainstream politics by appealing to xenophobia and islamophobia. A hostile climate toward migrants has been fostered by their employment of anti-immigrant rhetoric and the dissemination of false information [9]. On the issue of limiting the nationality of migrants' children and their access to public health care, Chilean president Sebastian Piera ran for office. Immigrants have been referred to as "animals" by President Trump [10]. "You see migrants and refugees being insulted and dehumanised daily." As Dieng pointed out during his remarks at the summit, you hear politicians utilising that segment of the public as a scapegoat. United Nations concerns about hate speech are not new, but the specific extent of antimigrant and antirefugee rhetoric is becoming

increasingly crucial. Since the beginning of the year, over one million asylum seekers have arrived in Europe alone, mainly from the war-torn countries of Libya, Iraq, Afghanistan, and Syria. As the number of migrants fleeing conflict has risen, so has the number of attacks, which have grown.

International human rights treaties and soft law instruments are being used to combat hate speech directed towards refugees and migrants. Concerned about this new problem, the UN launched a comprehensive initiative that culminated in July 2018 with UN member states presenting their renewed commitment to human rights and fundamental values of international law in the form of the Global Compact for Migration. This bolstered the UN's agreement to combat xenophobia and other types of prejudice. Even though 152 nations signed the compact, far-right leaders and parties overwhelmingly rejected it, demonstrating the centrality of antimigrant sentiment in their schedule. No international treaty or convention defines hate speech, but it can be found in human rights treaties and soft laws that limit freedom of expression in general. When construed according to international law, domestic laws should give protection against hate speech to refugees, even while hate speech rules are denied domestically. International law provides a clear outline for the concept of prohibited hate speech. To substantiate this claim, this paper will demonstrate that antihate speech legislation complies with international standards on the limits of freedom of expression while also providing sufficient justification for these laws to protect migrants and refugees under international law.

There are many articles based on hate speech detection, but the novelty of this work is to analyse the hate speech against the woman and the migrants using the deep learning model [11].

3. Proposed Methodology

Researchers employ neural networks, which are similar to the human brain in that they connect nodes that process and organise information. As a result of the usage of insults, phrases, and other disparaging statements as training data, an intelligent system can "learn" the patterns and structure of language to forecast incoming tweets and identify objectionable ones. Essential pronouns and other factors can fundamentally alter the meaning of a statement in some sentences. "According to a researcher at the University of Jaén, "with our technology and the support of language resources, we can identify expressions associated with hate speech [12]." A study published in ACM Transactions on Internet Technology, entitled 'Detecting Misogyny and Xenophobia in Spanish Tweets Using Language Technologies', describes how the researchers generated four lists of words in Spanish containing offensive and insulting expressions and words against women and migrants. Artificial intelligence uses this information to identify hate speech focused exclusively on these two demographic groups. Researchers at the University of Jaén are constantly adding new lexical resources to this technology, such as dictionaries and word lists, to improve its accuracy and effectiveness.

Previously hate speech was analyzed through machine learning algorithms. Later on, deep learning algorithms are widely used for performing the analysis. This article is also based on a deep learning approach to perform a better analysis in the hate speech analysis. In the first analysis, hate speech is classified against the woman using deep learning models and in the second analysis is to detect whether the hate speech is performed by the individual or the group of people. To perform the implementation, two different datasets are used, i.e., English dataset based on English text, and the next one is Spanish dataset based on Spanish text [13]. The deep learning model is implemented by the convolutional neural networks with different number of max pooling layer, dropout layers, activation function, and many more. The deep learning models are integrated with the word embedding model such as inverse glove (global vector), document frequency (TF-IDF), and transformer-based embedding, the implemented models are the combination of the convolutional neural network, bilong short-term memory, and multilayer perceptron. The description of the following models are as follows:

3.1. Datasets. The dataset split into two parts, i.e., training dataset and the testing dataset. In the training dataset, there are 31,963 records, and the testing dataset consists of 17,198 records. These are further split separately into English dataset and the Spanish dataset. The training dataset is used for building the model, and the testing dataset is used for the validation of the model.

3.2. Model 1: Global Vector (Glove) with Convolutional Neural Network, Bi-Long Short-Term Memory, and Multilayer Perceptron. In this method, the text dataset based on English and Spanish language is classified using the deep learning model for the analysis of hate speech against woman and migrants. The word embedding is done through the glove model for performing distributed word representation Figure 1 displays the processing of the English and the Spanish text dataset for the hate speech classification with the help of the Glove model and the ensembled deep learning model.

In this model, firstly, the word embedding model is applied, i.e., glove [14]. Glove encodes a corpus into pre-trained weights. Next, deep neural networks use this embedding layer as an input layer. Two convolutional layers, two dropout layers, two max-pooling layers, a flatten layer, and a dense layer were all used in the CNN model [15]. A convolution, dropout, and max-pooling layer follow this embedding layer before the flatten and dense layers are applied. A dropout layer followed by a dense layer in subtask is considered the most effective. We utilized an embedding dimension of 300 for each language subtask to minimize losses and get the most accurate results and applied the “Adam” optimizer. The dataset for subtask B was so imbalanced that we used the “ADASYN” oversampling technique to balance the data [16]. Internal activation is based on “ReLU,” and the final output dense layer is on the sigmoid [17].

Two hidden layers of a multilayer perceptron (MLP) integrate the LSTM with the MLP [18]. The LSTM neural network processes word embeddings one at a time while maintaining the order of the words [19]. Hyperbolic tangent activation is used to handle the output of the LSTM neural network. This is a 600-by-600-pixel vector. The MLP network has three layers:

- (i) An input layer with 600 neurons
- (ii) A hidden layer with 1,600 neurons and 100 neurons activated by ReLU
- (iii) An output layer with the rest of the neurons (a single neuron with sigmoid activation multilayer per function)
- (iv) Between each pair of layers, dropout units have been inserted, each with a probability of 0.8%. To avoid overfitting, the dropout units provide a possibility of deletion to each input neuron during training
- (v) All of our models have been built using the Keras library

3.3. Model 2: TF-IDF (Term Frequency- Inverse Document Frequency) with Multilayer perceptron, Support Vector Machine, and XGBoost (Extreme Gradient Boosting) Classifier [20]. In this model, the word embedding TF-IDF model is integrated with multilayer perceptron, support vector machine, and XGBoost (extreme gradient boosting) classifier.

- (i) TF-IDF (term frequency-inverse document frequency). Documents are distinct from each other based on the terms they use. Classifications are analyzed using the TF-IDF approach to determine the importance of terms. The TF-IDF score of each term or word in the document is used instead of frequency in this method. Similar to the bag of word algorithm, the TF-IDF score of each term replaces the word count.

$$TF * IDF(t, d) = TF_{t,d} * \frac{\log N}{DF_t} \quad (1)$$

In the above equation, TF represents the term frequency and IDF defines the inverse document frequency. t, d is showing the number of terms t appears in a document d . N denotes the number of documents.

If you want to calculate the TF-IDF score of an individual phrase inside an entire document, you can do so by multiplying the total number of documents in the document by the frequency of that term (DF) [21].

- (ii) Multilayer. Multilayer perceptron is a feedforward artificial neural network (ANN) that includes input and hidden layers and output and feedback layers. Backpropagation is used to update the weights of all nodes in an MLP. During training, nonlinear mapping is learned by utilising nonlinear activation functions and numerous layers. Nonlinear activation is used to generate the label in MLP [22].

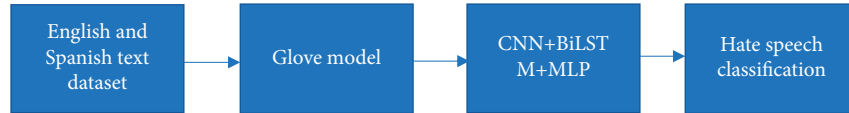


FIGURE 1: Glove model with ensemble CNN, BiLSTM, and MLP.

TABLE 1: Model comparison based on English dataset for the hate speech classification against woman.

Model	Accuracy	Precision	F1 score
Support vector machine	87.58	86.42	85.43
Multilayer perceptron	88.48	87.38	86.31
Random forest	88.88	87.48	86.36
CNN	93.79	91.9	90.14
Glove-BiLSTM + CNN + MLP	94.81	93.2	92.19
TF-IDF + MLP + SVM + XGB	95.63	94.1	93.2
Transformer-CNN and MLP	95.78	94.6	93.4

TABLE 2: Model comparison based on Spanish dataset for the hate speech classification against woman.

Model	Accuracy	Precision	F1 score
Support vector machine	87.51	86.37	85.14
Multilayer perceptron	87.49	86.46	85.43
Random forest	87.83	86.78	85.63
CNN	92.89	91.77	91.01
Glove-BiLSTM + CNN + MLP	93.42	92.74	91.36
TF-IDF + MLP + SVM + XGB	94.65	93.81	92.68
Transformer-CNN and MLP	95.71	94.62	93.52

TABLE 3: Model comparison based on English dataset for the hate speech detection by the individual or group of people.

Model	Accuracy	Precision	F1 score
Support vector machine	86.59	85.23	84.76
Multilayer perceptron	87.49	86.23	85.17
Random forest	88.83	87.15	86.33
CNN	94.79	92.23	91.74
Glove-BiLSTM + CNN + MLP	95.61	94.61	93.28
TF-IDF + MLP + SVM + XGB	96.1	95.15	94.17
Transformer-CNN and MLP	96.23	95.23	94.19

TABLE 4: Model comparison based on Spanish dataset for the hate speech detection by the individual or group of people.

Model	Accuracy	Precision	F1 score
Support vector machine	87.58	86.47	85.14
Multilayer perceptron	88.48	87.13	86.38
Random forest	88.88	87.52	86.46
CNN	93.79	92.16	91.53
Glove-BiLSTM + CNN + MLP	94.23	93.49	92.43
TF-IDF + MLP + SVM + XGB	95.12	94.76	93.71
Transformer-CNN and MLP	95.63	94.39	93.82

(iii) Support vector machine. One of the most common classification and regression algorithms is the Support Vector Machine (SVM), a supervised technique for classification and regression issues. To locate the decision border between two classes, it uses a vector space model that is as far away from the data points as possible, and the support vectors

are data points near the hyperplane that divides the classes [22, 23].

(iv) XGBoost classifier. In addition to tree-based models, another sort of an ensemble model is a “boosting model.” For example, machine learning ensemble meta-algorithms for minimising bias and variance in supervised learning are known as

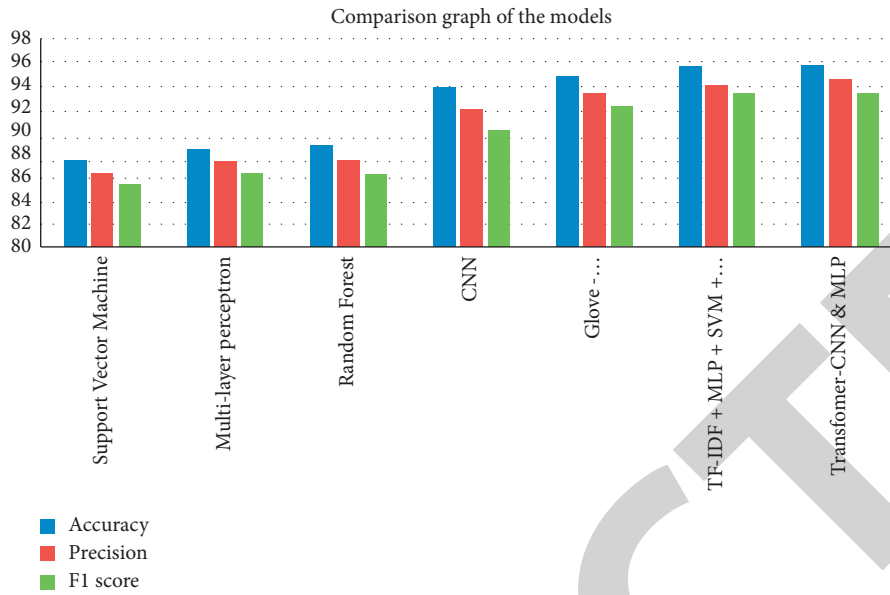


FIGURE 2: Comparison graph of the models based on English dataset for the hate speech classification against a woman.

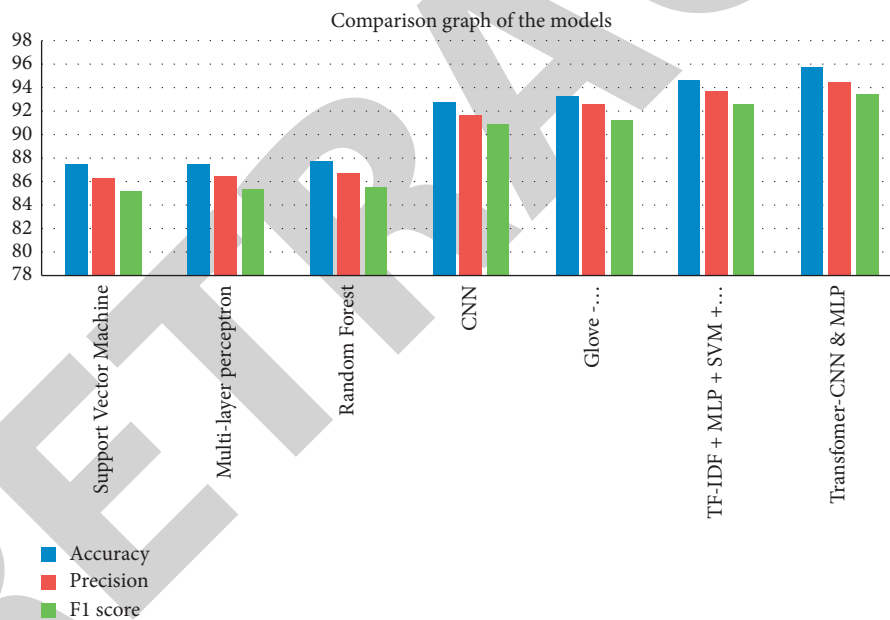


FIGURE 3: Comparison graph of the models based on Spanish dataset for the hate speech classification against a woman.

“boosts.” They are a class of machine learning algorithms that turn weak learners into strong ones. For example, a classifier with a slender correlation with the actual categorization is said to be a weak learner (it can label samples better than random guessing).

3.4. Model 3: Transformer Model with Convolutional Neural Network and the Multilayer Perceptron. In this proposed model, the word embedding is done through transformer model; i.e., electra and the classification are performed by

using the ensemble convolutional neural network and the multilayer perceptron model.

The electra model is possible to corrupt the input by replacing some tokens with (MASK) and then train a model to recreate them using pretraining approaches such as BERT. To be effective, they typically demand a considerable amount of computing power to perform well in downstream NLP tasks [24]. A more sample-efficient pretraining task termed replacement token detection has been offered as an alternative. By replacing tokens with plausible alternatives from a tiny generator network, the technique corrupts input rather than hiding it. Finally, instead of predicting the

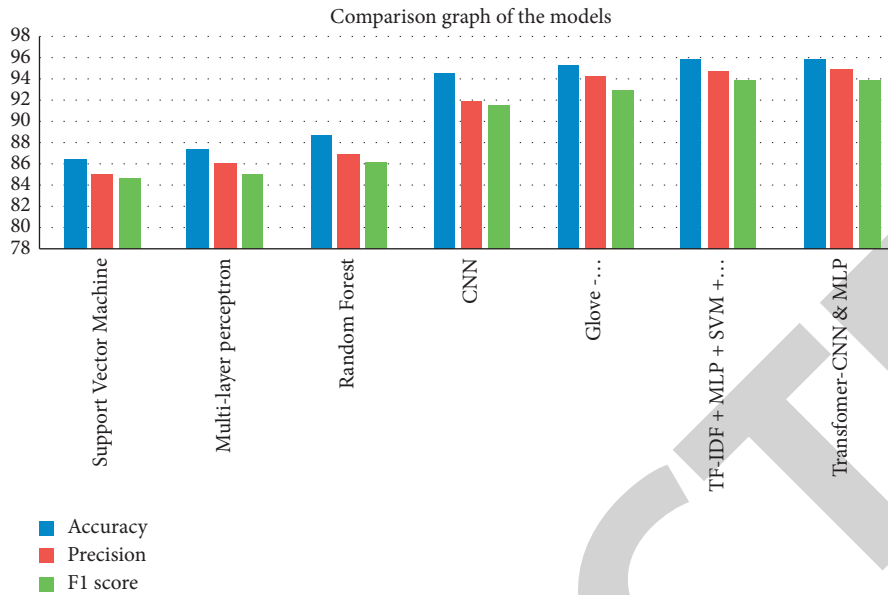


FIGURE 4: Graph comparison based on English dataset for the hate speech detection by the individual or group of people.

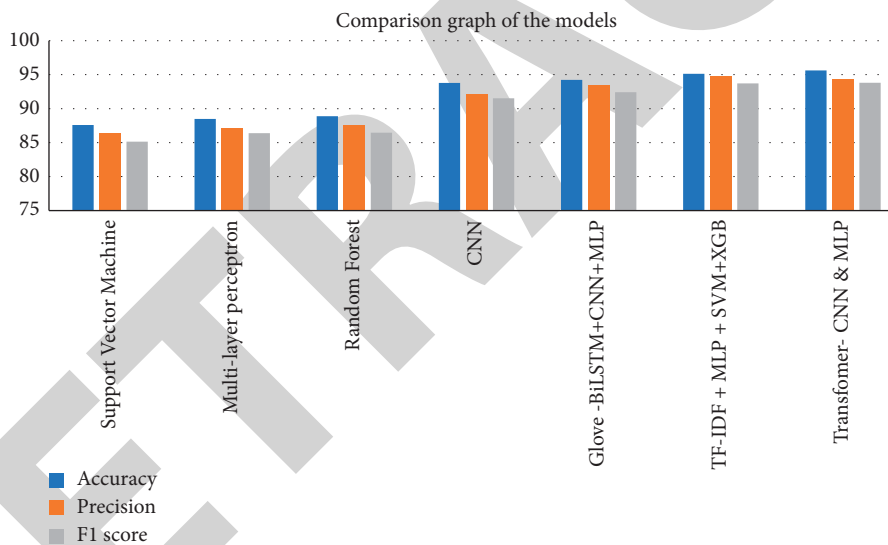


FIGURE 5: Graph comparison based on Spanish dataset for the hate speech detection by the individual or group of people.

original identities of the corrupted tokens, the electra model trains a discriminative model that predicts whether a generator sample replaced each token in the corrupted input. Because this new pretraining job has specified overall input tokens rather than just a limited fraction that was masked away, it is more efficient than MLM, according to a series of rigorous studies.

Given the same model size, data, and computing power, the contextual representations learnt this way beat those trained by BERT. On the GLUE natural language understanding benchmark, an algorithm trained on one GPU for four days outperforms the state-of-the-art GPT (learned using 30 times more CPU) [25]. Regarding performance at scale, the technique is comparable to RoBERT and XLNet while utilising less than a quarter of their computation and

surpassing them when using the same amount of computing [26].

4. Results and Discussion

To perform the implementation, a python programming language is used. The evaluation parameters used for detecting the models are accuracy, precision, and the F1 score. All the models are implemented separately, and the results are compared. The model implemented using glove and the LSTM has performed better than the other states-of-the-art algorithms. In Tables 1 and 2, comparison results of the proposed method are given. The results shown in Table 1 was generated through the English language dataset, and Table 2 displays the result achieved through the Spanish

dataset. All the three above discussed models in section 3 are applied on both datasets, i.e., English and Spanish.

Above Tables 1–4 and Figures 2–5 have shown the comparison of the models based on English and Spanish dataset. The classification is done to perform hate speech classification against the woman and the migrants. The results have shown that the model implemented with the transformer model and the deep learning model has achieved a better accuracy than the other models.

5. Conclusions and Future Work

This paper focuses on text categorization in which two datasets are based on English and Spanish languages. These datasets are used for performing the hate speech classification against women and migrants. In the first analysis, three different types of deep learning models based on artificial neural networks are applied to the English dataset for performing the hate speech analysis, and in the second analysis, the same deep learning models are applied to the Spanish dataset. The parameters used for evaluating the model are accuracy, precision, and the F1 score. The model implemented with the transformer word embedding and the hybrid approach of the convolutional neural network and multilayer perceptron has achieved more than 95% accuracy, which is better than the other state-of-the-art algorithms. In the future work, to enhance the classification performance some more algorithms based on convolutional neural networks, capsule networks will be implemented.

Data Availability

Dataset has been downloaded from the https://www.kaggle.com/dv1453/twitter-sentiment-analysis-analytics-vidya?select=test_tweets_anuFYb8.csv, and some more content is added in the dataset through Extracting Tweets from Twitter using tweepy library.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] J. W. Howard, "Free speech and hate speech," *Annual Review of Political Science*, vol. 22, no. 1, pp. 93–109, 2019.
- [2] Y. Gorny, "The jewish press and the holocaust, 1939-1945," vol. 27, no. 1, pp. 147–149, 2011, <https://www.researchgate.net/journal/Holocaust-and-Genocide-Studies-1476-7937>.
- [3] H. Nasser Alsager, "Towards a stylometric authorship recognition model for the social media texts in Arabic," *Arab World English Journal*, vol. 11, no. 4, pp. 490–507, 2020.
- [4] B. Erdem, "Fighting infodemic becomes must after COVID-19 pandemic's onslaught on truth, knowledge," *European Journal of Natural Sciences and Medicine*, vol. 4, no. 2, p. 110, 2021.
- [5] P. Arya, A. Bhagat, and R. Nair, "Improved performance of machine learning algorithms via ensemble learning methods of sentiment analysis," *International Journal on Emerging Technologies*, vol. 10, no. 2, pp. 110–116, 2019.
- [6] E. Dugarova, "Unpaid care work in times of the COVID-19 crisis: gendered impacts, emerging evidence and promising policy responses," in *Proceedings of the UN Expert Group Meeting "Families in development: Assessing progress, challenges and emerging issues. Focus on modalities for IYF+30*, June 2020.
- [7] United Nations, *United Nations Strategy and Plan of Action on Hate Speech*, United Nations Report, 2019.
- [8] A. Dieng, "Remarks by Adama Dieng," *Proceedings of the ASIL Annual Meeting*, vol. 114, pp. 340–343, 2020.
- [9] C. Williams and M. Graham, "A World on the mmm and social work," *British Journal of Social Work*, vol. 44, pp. i1–i17, 2014.
- [10] L. Finley and L. Esposito, "The immigrant as bogeyman: examining donald Trump and the right's anti-immigrant, anti-PC rhetoric," *Humanity and Society*, vol. 44, no. 2, pp. 178–197, 2020.
- [11] R. Alshalan and H. Al-Khalifa, "A deep learning approach for automatic hate speech detection in the saudi twitter sphere," *Applied Sciences (Switzerland)*, vol. 10, no. 23, 2020.
- [12] F.-M. Plaza-Del-Arco, M. D. Molina-González, L. A. Ureña-López, and M. T. Martín-Valdivia, "Detecting misogyny and xenophobia in spanish tweets using language technologies," *ACM Transactions on Internet Technology*, vol. 20, no. 2, pp. 1–19, 2020.
- [13] T. Davidson, D. Warmley, M. Macy, and I. Weber, "Automated hate speech detection and the problem of offensive language," 2017, <https://arxiv.org/pdf/1703.04009.pdf>.
- [14] J. Pennington, R. Socher, and C. Manning, "Glove: global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543, January 2014.
- [15] H. Krishnan, M. S. Elayidom, and T. Santhanakrishnan, "Optimization assisted convolutional neural network for sentiment analysis with weighted holoentropy-based features," *International Journal of Information Technology and Decision Making*, vol. 20, no. 4, pp. 1261–1297, 2021.
- [16] H. A. Gameng, B. D. Gerardo, and R. P. Medina, "A modified adaptive synthetic smote approach in graduation success rate classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 8, no. 6, pp. 3053–3057, 2019.
- [17] R. Nair and A. Bhagat, "Genes expression classification using improved deep learning method," *International Journal on Emerging Technologies*, vol. 10, no. 3, pp. 64–68, 2019.
- [18] N. Munasatya and S. Novianto, "Natural language processing untuk analisis sentimen presiden jokowi menggunakan multi layer perceptron natural language processing for president jokowi sentiment analysis using multi layer perceptron," *Agustus*, vol. 19, no. 3, 2020.
- [19] Y. Zhang, P. Tiwari, D. Song et al., "Learning interaction dynamics with an interactive LSTM for conversational sentiment analysis," *Neural Networks*, vol. 133, pp. 40–56, 2021.
- [20] Y. Huang, R. Wang, B. Huang, B. Wei, S. L. Zheng, and M. Chen, "Sentiment classification of crowdsourcing participant's reviews text based on LDA topic model," *IEEE Access*, vol. 9, pp. 108131–108143, 2021.
- [21] D. Endalie and G. Haile, "Automated Amharic news categorization using deep learning models," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3774607, 9 pages, 2021.
- [22] H. el Rifai, L. al Qadi, and A. Elnagar, "Arabic text classification: the need for multi-labeling systems," *Neural Computing & Applications*, vol. 34, no. 2, pp. 1135–1159, 2021.

Retraction

Retracted: Psychological Analysis for Depression Detection from Social Networking Sites

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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. Gupta, L. Goel, A. Singh, A. Prasad, and M. A. Ullah, "Psychological Analysis for Depression Detection from Social Networking Sites," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4395358, 14 pages, 2022.

Research Article

Psychological Analysis for Depression Detection from Social Networking Sites

Sonam Gupta ¹, Lipika Goel ², Arjun Singh ³, Ajay Prasad ⁴,
and Mohammad Aman Ullah ⁵

¹Department of Computer Science and Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, India

²Gokaraju Rangaraju Institute of Engineering and Technology, Hyderabad, India

³School of Computing and Information Technology, Manipal University Jaipur, Jaipur, India

⁴University of Petroleum and Energy Studies, Dehradun, India

⁵Department of Computer Science and Engineering, International Islamic University Chittagong, Chittagong, Bangladesh

Correspondence should be addressed to Mohammad Aman Ullah; aman_cse@iiuc.ac.bd

Received 7 December 2021; Revised 28 February 2022; Accepted 24 March 2022; Published 6 April 2022

Academic Editor: Alexander Hošovský

Copyright © 2022 Sonam Gupta 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.

Rapid technological advancements are altering people's communication styles. With the growth of the Internet, social networks (Twitter, Facebook, Telegram, and Instagram) have become popular forums for people to share their thoughts, psychological behavior, and emotions. Psychological analysis analyzes text and extracts facts, features, and important information from the opinions of users. Researchers working on psychological analysis rely on social networks for the detection of depression-related behavior and activity. Social networks provide innumerable data on mindsets of a person's onset of depression, such as low sociology and activities such as undergoing medical treatment, a primary emphasis on oneself, and a high rate of activity during the day and night. In this paper, we used five machine learning classifiers—decision trees, K-nearest neighbor, support vector machines, logistic regression, and LSTM—for depression detection in tweets. The dataset is collected in two forms—balanced and imbalanced—where the oversampling of techniques is studied technically. The results show that the LSTM classification model outperforms the other baseline models in the depression detection healthcare approach for both balanced and imbalanced data.

1. Introduction

Psychological analysis is a process in which psychological data are extracted from text-based data. To eliminate emotions, opinions, and judgement-forming, text data are used. Having opinions or views toward some products or any topic is human psychology, which defines what one thinks about the products or topic. Nowadays, the way of expressing various emotions and giving opinions has changed drastically with the advancement of social media and Internet technology. People use blogs, product recommendation and review websites, and other social media to give opinions about products, movies, and political parties and on current important topics. Famous social media platforms such as Facebook, Twitter, and Reddit have become the most reliable platforms for sharing opinions and reviews among a new generation of Internet users [1].

Business firms and organizations use people-oriented psychological feedback to increase their products' value and quality.

1.1. Role of Psychology. Human beings usually have a greater sense of emotions and feelings; these feelings, when merged with technology, can be converted into useful tools. Another word used for human feelings is psychology. Psychology is used in various research works, such as identification of mental health issues, product reviewing, customer satisfaction identification, and business advertisements. In business advertisements, psychology helps in decision-making, e.g., to buy a product online, when the product is not available to touch and feel, then product reviews help a customer take decisions on whether the product is good to buy or not. Usually, online reviews are a mix of true and false

opinions. Reviews with high polarity toward positive psychology increase the product value [2]. Reviews also hint about defaults and needs in a product that help business firms improve their products and satisfy their customers. Half of the business world depends on customer feedback and reviews. Another application of psychological analysis is market research that includes collecting data through social media and other websites. These data help in understanding the current market trend and advertisements' quality that affect people. The reviews provided by people for a product or a movie follow a collect-and-pass-through recommendation system, which depicts the polarity of psychology that shows whether people like the product or the movie or not. People's opinions and reviews can be analyzed thoroughly, which helps boost business performance and design future services [3].

Second, psychology plays an important role in mental health research, such as identification of anxiety attacks, major depressive disorders, bipolar disorder, and many more. We focus on major depressive disorder in this paper. Depression is also known as major depressive disorder and is commonly found in people with anxiety issues. Every 1 in 5 people have been suffering or have suffered from depression [4]. Of the total world population, 4.5% of the population is suffering from depressive disorder. Depression has some common symptoms such as anxiety attacks, loss of appetite, feeling sad for a long duration (1 month), and losing interest in favorite activities. Usually, during the initial state of depressive disorder, an individual avoids social gatherings, shows lack of energy, makes fewer efforts to communicate with friends and family, and shows a feeling of incompleteness. Due to lack of social interaction and the fear of being judged, depression survivors indulge in social networking to share their thoughts and feelings with people similar to themselves (Munmun et al, 2021). By being hidden and still saying what they had in their mind, this made social media more useful for depression detection research work. Mainly, the Twitter platform comprises many eco-groups where people of the same interest connect with each other through, for example, the "depression group," "mental health club," etc. We use the Twitter platform in this work, which provides sufficient data for depression detection and classifying users into depressed and nondepressed categories.

Words convey different psychologies and tell about the current psychological state of a being. For example, consider words such as "not feeling good" or "why people behave like this to me." Sentences comprise words with both positive and negative polarity, but with the use of "not" and "why," the very meaning and feelings change. The dictionary software LIWC (Linguistic Inquiry and Word Count) is usually used for analyzing words and extracting their meaning in terms of psychology [5–7]. This dictionary software comprises multiple categories according to the types of human psychology, which helps in psychological analysis using textual data only. Multiple factors are considered when we perform psychological analysis to detect mental health illness, such as temporal factors, emotional factors (positive: "excited," "wonderful," and "lovely") (negative: "empty," "lonely," and "forgetful"), and use of

personal pronouns ("I," "me," and "them") (Verma et.al, 2020).

1.2. Psychological Analysis Using Social Media. Social media platforms such as Facebook, Twitter, Instagram, and Reddit consist of rich amounts of data to perform psychological classification tasks. Psychological analysis is all about extracting psychological data from textual data. Twitter is a famous social media platform used by millions of people worldwide. People share their opinions about current hot topics or any political chaos and discuss famous incidents. The opinions or views in Twitter are in the forms of tweets, which have a maximum limit of 250 words approximately. This limit made Twitter data important. As tweets are meant to have a certain word limit, users use specific words to express their views and emotions. Twitter, nowadays known as a hub of political and government people—mainly politicians and government officials, is used for various announcements about international and national projects [8]. Users discuss the views given by various political parties. During elections, this kind of discussion will help in determining results of elections if psychological analysis is carried out using tweets collect from the Twitter platform. Another famous social networking platform is Facebook. As the Internet expanded, Facebook became famous in 2007, and now, in 2021, Facebook currently has 5 billion users around the world. People who use the platform provide useful insights into businesses, movies, politics, and trends. Facebook generates a vast number of data every year, which comprise images, text, and links. Facebook comprises multiple online communities where people of likeable interest connect with each other and share their psychological views and opinions. Online communities, such as political groups that discuss their favorite and hated political personalities, talk about parties that have different opinions. These types of data are useful in understanding what kind of psychology people have for their political parties. Another group is mental health groups where people discuss their mental health—how they are fighting every day to live and undergoing treatment [9]. Data from these types of groups provide sufficient text data to understand the psychology of people in these groups. Identifying psychology through text data is much easier than identifying psychology through image-based data as images require use of deep learning classifiers and high-quality image-based data. However, with text-based data, psychological analysis is easily performed on these giant social networking platforms. A lot of research has been carried out in this field, and it uses artificial intelligence, which in combination with LIWC provides satisfactory results. Facebook contains ads for various companies and data about companies that we search on the Google platform. Usually, these ads sometimes use surveys to understand user psychology as to what kind of image they have about the business organization. As shown in Figure 1, there are some basic human emotions that are usually conveyed through specific words. Facebook provides some reaction emoticons in the "Posts" section that contain emoticons similar to these basic human emotions. Figure 2 shows

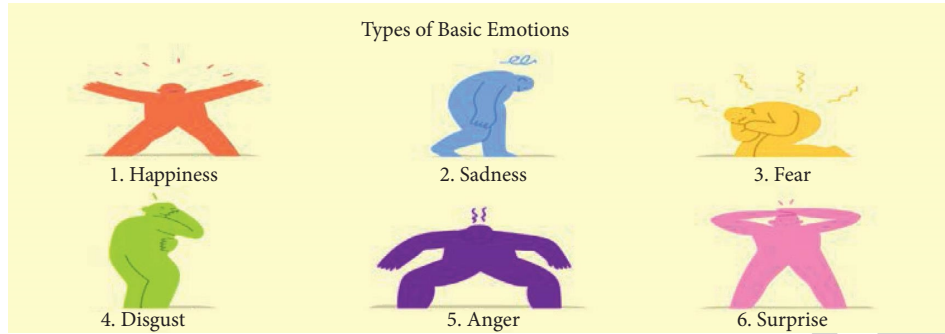


FIGURE 1: Commonly used human emotions.

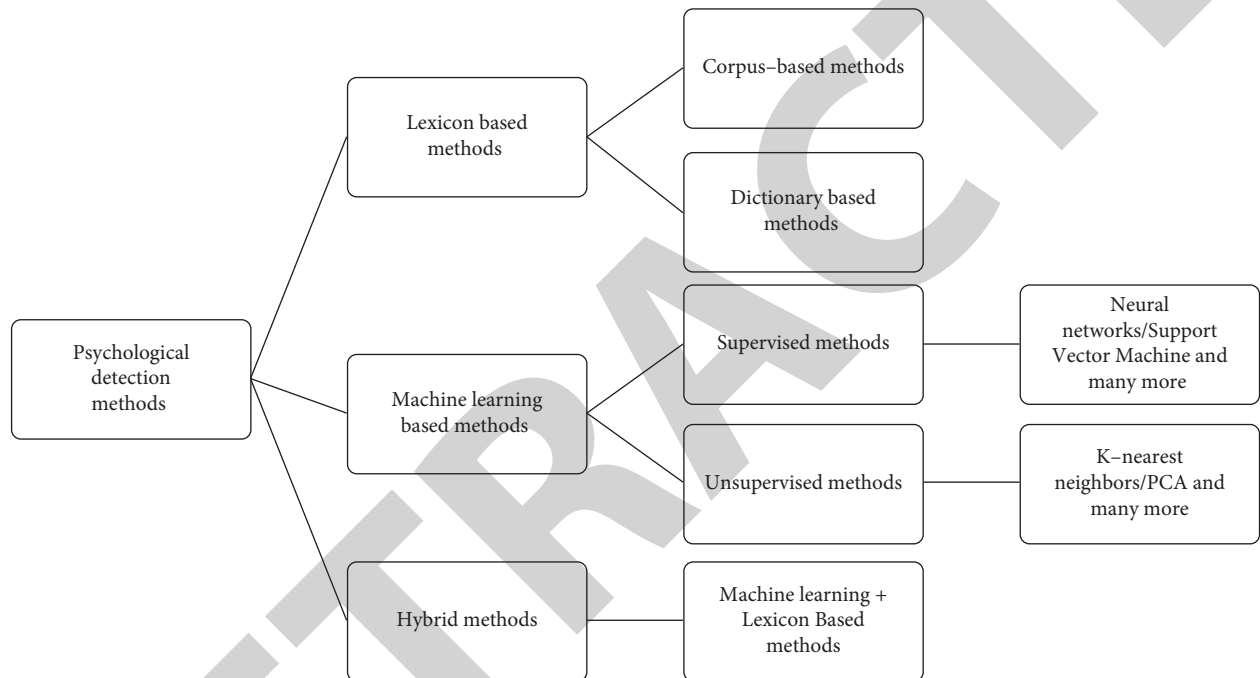


FIGURE 2: Various methods for psychological classification.

various methods for psychological classification, which include dictionary-based and machine learning-based methods. In dictionary-based methods, the sets of words are classified using dictionary values, and machine learning-based methods include supervised and unsupervised learning methods, such as neural networks, that provide some of the best results. There are hybrid-based methods that combine dictionary-based methods with machine learning-based methods.

In this paper, we have selected and trained the tweet data. We have performed data preprocessing on datasets and removal of the raw data from the dataset. After value extraction from the datasets, training the data of tweets and cross-validating the training dataset were carried out. We used 5 machine learning classifiers—support vector machines, decision trees, logistic regression, K-nearest neighbor, and LSTM—for depression detection in tweets. The results show that the LSTM classification model outperforms the other baseline models in the depression detection approach. To handle the imbalanced dataset, the oversampling

and undersampling of class imbalance approaches are implemented and analyzed.

In this paper, the study is carried out on two datasets that involve the imbalance and the balance set, and different techniques, such as SMOTE and RUS, are used for oversampling and undersampling to work with the dataset. From the research gap, we are trying to conclude that the performance of LSTM is better than that of other machine learning classifiers.

The methodology that has been followed in this paper is described in Figure 3.

The paper is organized as follows:

Section 2 comprises a literature review. Section 3 comprises the prerequisites of depression detection. Section 4 describes the proposed method, and Section 5 includes the experiment setting, which includes a subsection of data information, modeling that involves the study of five machine learning classifiers used in this research, data preprocessing, and performance measurement. Section 6 includes discussion and results, as well as various

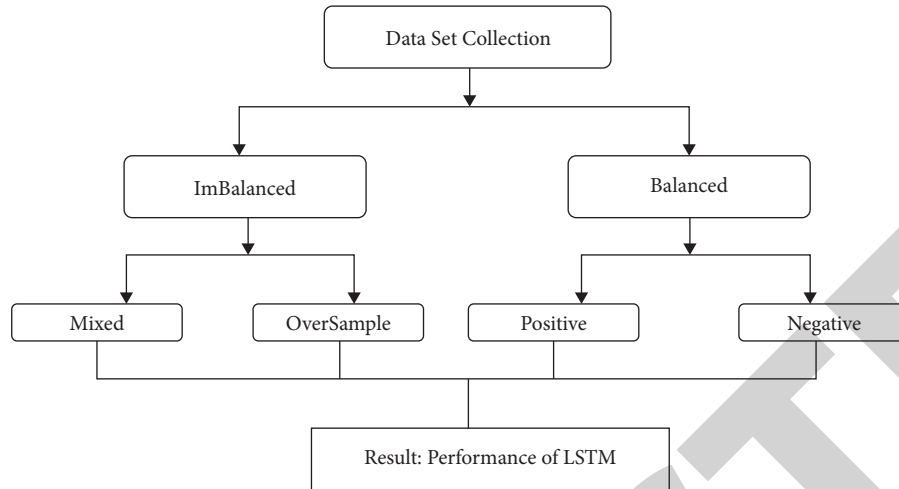


FIGURE 3: Methodology.

measurement comparisons in the form of a graph, and finally, the conclusion and future scope are mentioned in Section 7.

2. Literature Review

Psychological analysis can be carried out using various methods, as explained in Figure 2, with the text-based dataset. In this section, we discuss the previous works performed using various techniques for psychological analysis and the depression detection task.

Using linguistics in depression detection is quite useful as it shows that words used by nondepressed and depressed people may differ. Depressed individuals mainly focus on themselves. In 2014, Nguyen studied two online discussion groups, namely, “control” and “clinical” groups. The “control” group comprised people with a similar interest and fun-loving people, whereas the “clinical” group comprised people suffering from mental illnesses such bipolar disorder, major depressive disorder, SAD, and anxiety attacks. People in the clinical group discussed their issues freely and took advice for medication and intervention. The author finds a difference in online communities involving people of these two groups. People in the “clinical” group usually use first-person pronouns (“I”, “me,” and “my”) in comparison with “control” group people, who use fewer first-person pronouns and discuss various activities such as dancing, singing, and running. This work reveals that use of language plays an important role in depression detection as words describe what someone’s current mental state is.

In 2005, Pennebaker used LIWC (Linguistic Inquiry and Word Count), a piece of dictionary software, for analyzing textual data to obtain meaningful insights. Detecting depression using Facebook comments and using 4 categories from the LIWC software to analyze words in the comments were the author’s objectives. These categories are language-based, time-based, or emotions-based or include all features. These factors have all minor characteristics of human language use and conditions such as related emotions, time periods, and use of nouns to find out the very meaning of

human speech. Using the KNN (K-nearest neighbour) machine learning algorithm for classification, an accuracy of 65% was achieved.

In 2019, Gaikar used SVMs (support vector machines) and Naïve Bayes classifiers to detect depression-related words and sentences and detected the types of depression from those identified words. The authors trained both the classifiers using the bipolar disorder dataset and depression illness-related dataset. The best accuracy achieved was 85% using a machine learning classifier. In [10], the author used audio and text-based data. These data play an important role in natural language processing, and they range from images, texts to videos, and small clips. With text data, the information is limited, but with the audio of depression survivors, it is easy to understand some facts about the affected person’s body language. A total of 142 individuals underwent depression detection tests by being asked some questions. They recorded the answers in audio forms and text forms directly from the 142 people, and feature extraction was carried out using the long short-term memory model. The results show that people with depression use more pauses than nondepressed persons and that using first-person pronouns is common among depression survivors; however, nondepressed persons use fewer pauses while talking and focus less on them.

In [11], the author used the Twitter platform to conduct the task of identification of depression using text-based data as Twitter provides short and useful linguistic phrases that directly shows the current mental state of the user. The data were gathered from the CLPsych 2015 conference in which the latest 3000 public tweets are available. For feature extraction, the author used the bag-of-words approach that is famous for identifying mental health illness using machine learning. The method shows word frequencies and the number associated with them and applies various machine learning classifiers in which Naïve Bayes/1-gram gives a recall value accuracy of 86% with 82%, respectively.

Reference [12] focuses on early detection of depression using neural networks. Depressive disorder has various stages, which include initial, intermediate, major, and severe

disorders. The initial level of depressive disorder includes visibility of symptoms appearing in affected people such as low appetite, feeling of suicide, social inferiority, and comparing themselves with others. The author suggests identifying depression at its initial stage where it is less harmful is easily curable. Using CNNs (convolutional neural networks), behavioral characteristics from text-based data (CLPsych) are extracted, and some improvements in the early detection parameter ERDE are introduced.

In 2018, Lang used speech data to detect depression, proposing a deep convolution neural network to process speech-based data. Authors use secondary datasets such as AVEC2013 and AVEC2014 depression datasets that comprise 340 short videos of 292 selected people. These data were collected through human beings and computer system interactions using webcams and microphones. The model gives an RMSE value of 9.0001 and an MAE of 7.4211. In Ref. [13], the author analyzed Facebook comments data using various machine learning algorithms. Facebook contains plenty of data that comprise videos, images, and text-based data. Using only text-based data from comments of various publicly available pages, the author analyzed them using machine learning classifiers such as decision trees, KNN, support vector machines, and ensemble classifiers, in which decision trees achieved the best accuracy of 73%.

In Ref. [9], the authors proposed a lexicon-based approach for detection of depression. They constructed a lexicon using Word2Vec, a semantic relationship graph, and the label propagation algorithm. The authors based it on 111,052 Weibo microblogs from 1868 users who were depressed or nondepressed. They have used five classification methods and considered six features to predict depression. The results show that the lexicon generated proved to be better for classification.

In Ref. [14], the researchers focused on mood analysis of human beings with the help of machine learning and deep learning tools of artificial intelligence. In this paper, authors also focused on limitations of artificial intelligence to detect depression. In Ref. [15], the author implemented a machine learning model with preprocessed data for automatic depression classification. Kinect captured a skeletal model used for data extraction and preprocessing. The model achieved 96.47% accuracy in old-age group and 53.85% accuracy in young-age group. In Ref. [16], the authors carried out prediction of depression by noticing the human behavior. For prediction, authors used smart phone datasets. The model achieved 96.44%–98.14% accuracy. In Ref. [17], the researchers implemented an automatic depression detection [AUTO DEP] model by using facial expressions of human beings. The authors used a linear binary pattern descriptor model for feature extraction. The performance of the automatic depression detection model is the same as that of usual previous models. The evaluation of the model was performed on MATLAB and the linear binary pattern descriptor FPGA using Xilinx VIVADO 16.4. Table 1 shows the consolidated literature survey performed by us before research.

After the survey, we have identified the following research questions:

RQ1: Will solving the class imbalance problem improve the performance of the psychological analysis model?

RQ2: Which technique of class imbalance (oversampling/undersampling) gives better results for depression detection?

RQ3: Which classification model outperforms the other baseline models in the depression detection approach?

The above-mentioned research questions have been answered in the subsequent sections.

3. Prerequisites for Depression Detection

Mental health conditions are the common problem for people. According to the WHO, 20% of people suffer from this problem. Adults and children with mental illnesses, such as depression, memory loss, hypertension, and anxiety attacks, were the primarily affected ones. Depression is the fastest growing health disorder; it depends on the mood, which comprises components of motivational and emotional conditions. Traditionally, mental health experts have mainly used clinical examination procedures depending on self-reporting of emotional, behavioral, and cognitive dimensions. To diagnose and calculate the predictions on the condition's evolution, machine learning algorithms have been used since 1998. There are various types of approaches used to detect depression as follows:

3.1. Machine Learning Approach. SVMs, decision trees, random forests, and Naïve Bayes are examples of supervised and unsupervised machine learning algorithms. These algorithms conduct longitudinal, temporal, and sequential analysis on the mood of human beings, and they also analyze social media and social networks. These algorithms detect depression by plotting graphs. The longitudinal analysis plots the user activities with time, and the results show less social interaction, more negative things, strongly clustered ego networks, increased interrelation, and medicinal issues.

3.2. Lexicon-Based Approach. The lexicon-based psychological analysis algorithm uses psychological normalization and evidence-based combined functions. In this, they used Rtexttools of machine learning for comparison of lexicon psychology. The lexicon analysis types in a corpus are adverb lexicons, network word lexicons, and negative word lexicons.

In this approach, the psychological calculation using the number of positive and negative words in a document can be carried out by using the following formula [9]:

$$\text{Score}(w) = \text{pos}(w) - \text{neg}(w), \quad (1)$$

where

TABLE 1: Literature survey.

Author	Dataset	Model	Result
Nguyen et al. [18]	Clinical group and control group	Lasso	Accuracy 93
Pennebaker et al. [6]	Facebook	Knn	Accuracy 65
Gaikar et al. [19]	Bipolar disorder and depression illness	SVM	Accuracy 85%
Hanai et al. [10]	Audio and text based	LSTM	F1: 0.44 and precision: 0.59%
Lang and Cao [20]	Twitter	Naive Bayes	Accuracy 86%
Trotzek et al. [12]	Reddit message	CNN	Accuracy 87%
Lang and Cao [20]	Speech data	Deep CNN	RMSE: 9.0001 and MAE: 7.4211
Islam et al. [8]	Facebook comments	KNN	Accuracy 73%
Li et al. [9]	Weibo microblogs	Logistic regression	Accuracy: 77% and precision: 77%
Li et al. [21]	Shandong mental health center	Kinetic captured skeleton	Accuracy 96.47%
Asare Kennedy et al. [16]	Smart phone dataset	SVM	Accuracy 96.44–98.14%
Tadalagi and Joshi [17]	Facial expression	SVM	Accuracy 72.8%

$$\begin{aligned} \text{pos}(w) &= pdf(w)N_{\text{pos}} \times 1 df(w), \\ \text{neg}(w) &= ndf(w)N_{\text{neg}} \times 1 df(w), \end{aligned} \quad (2)$$

$$\begin{aligned} N_{\text{pos}} &= \sum_{w \in \text{vocab}} p df(w), \\ N_{\text{neg}} &= \sum_{w \in \text{vocab}} n df(w). \end{aligned} \quad (3)$$

In the above-mentioned equation, N_{pos} shows the total number of positive words in the current tweet-based data. N_{pos} is the sum of the positive document frequency of each word in the document. N_{neg} shows the total number of negative words in the tweet-based dataset.

The two methods of lexicon are as follows:

3.2.1. Dictionary-Based Method. The dictionary-based approach is a part of a statistical model that is used to encode the symbols from the datasets. This approach does not encode the single symbols as a variable length of bit vectors. It encodes the variable length, starting into a single token.

The token starts with the dictionary index. If tokens are smaller, then they are replaced with bit vectors. The dictionary-based approach is an easily understandable approach.

3.2.2. Corpus-Based Method. The corpus-based approach is based on study of language, which is based on the language of the text corpus. It is the reliable analysis of languages, psychologies, and disorders. The corpus-based approach collects the natural context of language and performs psychological analysis with minimum experimental interference. The text corpus has been used for linguistic search and to compile dictionaries since 1985, and this method is used for psychological analysis. A phrase's semantic orientation is calculated, where the orientation is determined by mutual information.

Semantic Orientation (phrase) = PMI (phrase, "Amazing") PMI (phrase, "Destitute")

The words "Amazing" and "Destitute" are used for calculating the SO of phrases.

3.3. Artificial Neural Network. The most powerful neural network is the ANN, which is a part of machine learning. It is

employed in a variety of fields, including computer vision, digital image processing, psychological analysis, and word sequence prediction, and it produces low-error results. ANN algorithms are modeled as the human brain. ANNs work just like how a human brain uses neurons and nerves and learns from the past data. Similarly, the ANN can learn from the past data, predictions, and classifications.

3.4. Class Imbalance Approach. The two types of the class imbalance approach are

- (1) **SMOTE**—Smote balances the class distribution and handles the imbalance problem in the dataset. It calculates the Euclidean distance between the minority class and other instances to find out the K-nearest neighbors.
- (2) **RUS**—It is an undersampling technique of the size of the class in no higher instances. The majority is reduced from the source dataset. It is a simple and easy method of undersampling. A subset of the majority class is selected and merged with the minority class samples of the dataset.

The aim of this study is to identify the individuals in depression on Twitter using tweets, which are short messages created by the individual users on Twitter. We are using 5 machine learning classifiers—SVMs, KNN, decision trees, logistic regression, and LSTM—for the classification job, and feasibility of this classification is examined through evaluation metrics precision, F1-score, and recall. To handle the imbalanced dataset, the oversampling and undersampling of class imbalance approaches are implemented and analyzed.

4. Proposed Methodology

In this paper, we have selected and trained the tweet data. We have performed data preprocessing on datasets and removal of the raw data from the dataset. After value extraction from the datasets, training the data of tweets and cross-validating the training dataset have been carried out. The sampling process is performed over the dataset so that it can sample the data according to the psychology.

The classifier is implemented, which will classify the actual data and apply various learning methods to it. In the classification technique, the data are categorized in two parts: one is imbalanced data, and the other one is balanced data. In imbalanced data, the data which do not contain any missed values can be oversampled. Balanced data are categorized into two types: (1) positive tweets and (2) negative tweets. The flowchart is shown in Figure 4.

5. Experiment Setting

5.1. Data Acquisition. In this study, we considered two datasets: Sentiment_140 and Sentiment_tweets3; these are publicly available on <https://www.kaggle.com> for research and study purposes.

Sentiment_140: This is one of the most famous datasets for sentiment analysis and for natural language processing. The dataset consists of 1.6 million tweets, which is extracted using Twitter API. The tweets have annotation 0 = negative, 2 = neutral, and 4 = positive, but we use different annotations for our study, which is 0 = negative and 1 = positive. There are 3 fields in the dataset:

Id = id of the tweet.

Text = the tweet data.

Label = the polarity of tweets (0 = negative and 1 = positive).

Sentiment_tweets3: This dataset consist of 14000 tweets, which is publicly available for use, from which only 6000 tweets are taken for use. For this study, we combined these two datasets and collectively used them; if the dataset is imbalanced, then the missing values of the dataset are checked, and the data are oversampled.

In the imbalanced dataset, first the number of instances with positive and negative tweets is checked. This will degrade the overall predictive performance of the model. The imbalanced data modeling is carried out with oversampling data techniques.

5.2. Modeling. Modeling is the mathematical expression that represents the data in the context of a problem, often a business problem. The main aim of modeling is to form data for insights. Machine learning algorithms perform processing of data. To process the data, some models are used. The following are the data processing models used in this paper:

- (1) **Decision tree (DT):** A decision tree is a decision support tool that includes the chances of event outcomes, resource cost, and utility. It is used in predictive modeling approaches, statistics, and machine learning.

One of the important parameters in decision trees is information gain, which minimizes the amount of information required to differentiate between two data points for partition[18].

$$\text{Info}(D) = \sum_m p_i \log z(p_i),$$

$$\text{InfoA}(D) = \sum_{vj=1} \frac{|DJ|}{|D|} * \text{Info}(Dj). \quad (4)$$

Here, p_i shows the probability that a tuple in dataset D may belongs to class c_i .

$\text{Info}(D)$ is the sum of the mean amount of data required to determine a data object class D to which it belongs. The info gain is calculated as

$$\text{Gain}(A) = \text{info}(D) - \text{info A}(d). \quad (5)$$

- (2) **Support vector machine (SVM):** A support vector machine is a model that uses the classification algorithm for two-group classification problems. An SVM is a fast and dependable algorithm. It consists of a line that separates two data objects, known as the decision boundary. This acts as the main separation axis. The equation of the separation axis is

$$Y = mx + c$$

where m stands for the slope.

Now, the hyperplane equation separating the data objects are

H: $wt(x)+b=0$, where b stands for the bias term.

- (3) **K-nearest neighbor (KNN):** This algorithm is used for classification and regression. The KNN algorithm assumes the similarity between the new case data and available cases and put the new cases into the category that is most similar to the available categories. The KNN algorithm utilizes the Euclidean distance formula to determine the minimum distance between two data points in the plane with coordinates (x, y) and (a, b) , which is represented as

$$\text{Dist}((x, y), (a, b)) = \sqrt{(x-a)^2 + (y-b)^2}. \quad (6)$$

- (4) **LSTM:** It is an artificial recurrent neural network architecture used in the field of deep learning. It can only process single data but can also process an entire sequence of data.

LSTM transition function[9]:

$$It = \sigma (Wi \cdot [ht-1, xt] + bi).$$

$$Ft = \sigma (Wf \cdot [ht-1, xt] + bf).$$

$$Qt = \tanh (Wt \cdot [ht-1, xt] + bq).$$

$$Ot = \sigma (Wo \cdot [ht-1, xt] + bo).$$

$$Ct = ft \otimes ct-1 + it \otimes qt.$$

$$Ht = ot \otimes \tanh (ct).$$

The LSTM model contains multiple iterative steps for every time stamp. Here, $ht-1$ shows the hidden states (old), xt shows the input of the current time stamp, and ft is the forget gate that partially removes data, which are redundant or not useful from the old memory cell, for further processing in LSTM with ot as the output gate that selects an appropriate output for the current time stamp. The sigmoid

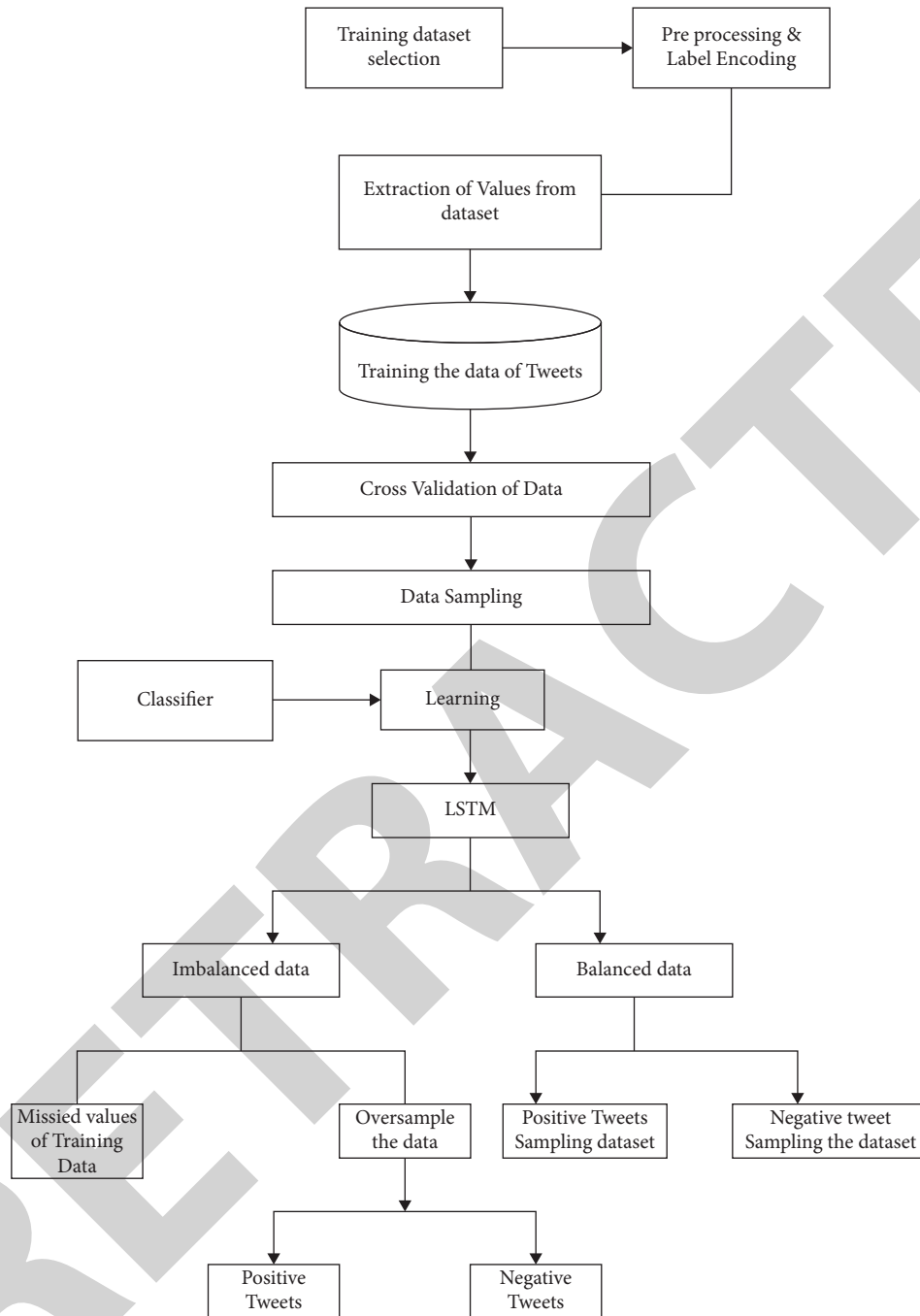


FIGURE 4: Proposed methodology for handling the data of tweets.

function is used here that has a range $[0, 1]$, which is shown using the σ symbol, with \otimes used for multiplication of elements.

5.3. Data Preprocessing. The data collected from the social media platform contain some error or may contain useless text, which causes difficulty in semantic analysis. As the dataset we are using is free from emojis, there is no need for emoji processing. Second, the stop words removal task is performed by using NLTK in Python. We can download the list of stop words, and stemming is used, which ignores

stop words and creates systems by removing suffixes or prefixes that are used with the word. In this study, we use a snowball stemmer, which is different from a porter stemmer, as it allows performing multiple language stemmers. Tfidfvectorizer is used to tokenize the given document.

5.4. Performance Measure. The performance measurement of datasets is carried out with the help of a confusion matrix from where we get the results, which calculates the values of accuracy, precision, and recall with the help of positive and

negative values of datasets. The formulas used for the calculation of values are

$$\begin{aligned} \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \\ \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{F1 - Score} &= \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \end{aligned} \quad (7)$$

6. Result and Discussion

We have performed the experiments in this proposed work in three folds. Experiment 1 performs depression detection using the imbalanced datasets. In Experiment 2, the SMOTE technique of oversampling has been implemented to handle the class imbalance issue of the training dataset. Experiment 3 implements the RUS technique of undersampling for solving the imbalance issue in the dataset. The performance of the classifiers discussed above is analyzed and compared. The classification algorithms are analyzed using scikit-learn library, and we plotted the graphs using Matplotlib library.

Table 2 tabulates the results of the performance measurement with the imbalanced dataset. Table 3 lists the performance of the various classification models using the balanced dataset of SMOTE. Table 4 tabulates the results of the performance measurement of the learning models using the balanced dataset of RUS.

From Table 2 it is noted that the average precision, recall, F1 score, and accuracy with the imbalanced dataset are 0.74, 0.65, 0.68, and 0.53, respectively, whereas the average precision, recall, F1 score, and accuracy with the balanced dataset using SMOTE are 0.77, 0.67, 0.71, and 0.72, respectively. There is a percentage increase of 4.05, 3.07, and 35.84 in the values of precision, recall, and accuracy, respectively. The results show that solving the class imbalance problem improves the performance of the psychological analysis model.

On comparison of Table 3 and Table 4, there is a percentage increase of 4.05, 1.51, and 4.34 in precision, recall, and accuracy, respectively, on using the SMOTE technique of oversampling in contrast with the RUS approach. The results illustrate that the SMOTE approach to handle class imbalance gives better results for depression detection.

From Tables 2–4, it can also be observed that the values of precision, recall, F1 score, and accuracy are high when LSTM is used as a learning model. The highest recall value of 0.75 is achieved using LSTM. Similar results with the accuracy value as high as 0.83 and the highest precision of 0.84 have been recorded with LSTM as the learning model. The observations mentioned above infer that LSTM outperforms the other baseline models in the depression detection approach.

TABLE 2: Results of performance measures with imbalanced dataset.

Classifier	Precision	Recall	F1-score	Accuracy
Decision tree	0.73	0.63	0.67	0.03
SVM	0.76	0.67	0.71	0.52
KNN	0.67	0.56	0.61	0.65
LR	0.75	0.68	0.71	0.71
LSTM	0.79	0.72	0.74	0.78
Average	0.74	0.65	0.68	0.53

TABLE 3: Results of performance measures using SMOTE.

Classifier	Precision	Recall	F1-score	Accuracy
Decision tree	0.76	0.64	0.69	0.68
SVM	0.79	0.69	0.73	0.62
KNN	0.69	0.59	0.63	0.71
LR	0.77	0.72	0.74	0.76
LSTM	0.84	0.75	0.79	0.83
Average	0.77	0.67	0.71	0.72

TABLE 4: Result of performance measure using RUS.

Classifier	Precision	Recall	F1-score	Accuracy
DT	0.72	0.64	1.80	0.67
SVM	0.77	0.67	0.71	0.59
KNN	0.67	0.57	0.61	0.68
LR	0.75	0.69	0.07	0.72
LSTM	0.82	0.73	0.77	0.80
Average	0.74	0.66	0.79	0.69

6.1. *Balanced Data.* Figures 5–8 show the graphical representation of the performance measurement of precision, recall, F1score, and accuracy using the imbalanced dataset on the various leaning models. The graphs indicate that LSTM outperforms other baseline learning models in performance measurement.

Figures 9–12 show the graphical representation of the performance measurement of precision, recall, F1score, and accuracy using the balanced dataset on the various leaning models after the application of the SMOTE oversampling technique. From the graphs, it can be clearly inferred that on resolving the class imbalance issue of the training dataset, the predictive performance of the models increases. The models are less biased and give more accurate results for depression detection.

Figures 13–16 show the graphical representation of the performance measurement of precision, recall, F1 score, and accuracy using the balanced dataset on the various leaning models after the application of the RUS undersampling technique. The plotted graphs indicate that although RUS gave better results than the models trained with imbalanced datasets, it did not outperform the oversampling approach of class imbalance.

From the above-mentioned statistical analysis and graphical interpretations, the answers to the research questions stated above are clearly provided as follows:

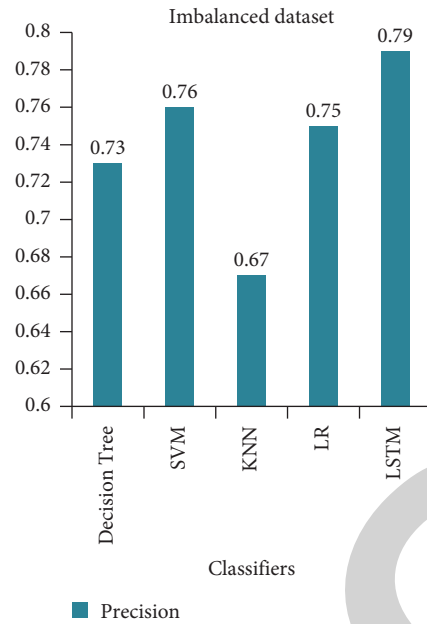


FIGURE 5: The precision value of imbalanced dataset.

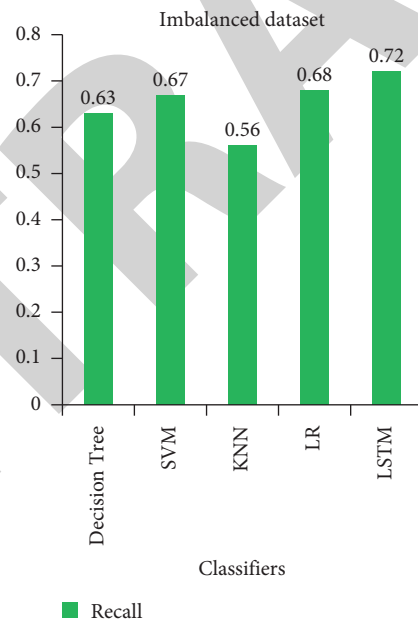


FIGURE 6: The recall value of imbalanced dataset.

For RQ1: The solution of the class imbalance problem improves the performance of the psychological analysis model.

For RQ2: The SMOTE oversampling technique of class imbalance gives better results for depression detection using psychological analysis.

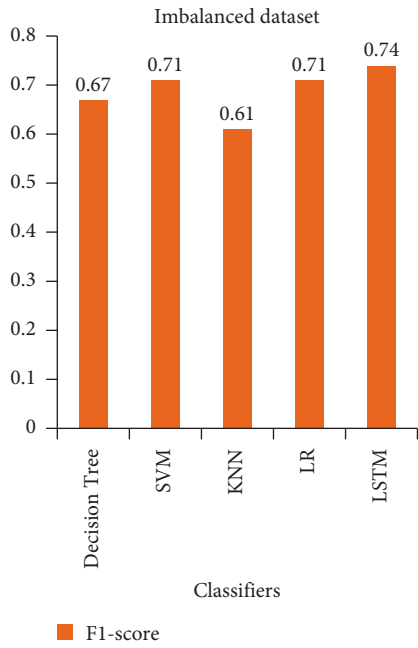


FIGURE 7: The F1-score value of imbalanced dataset.

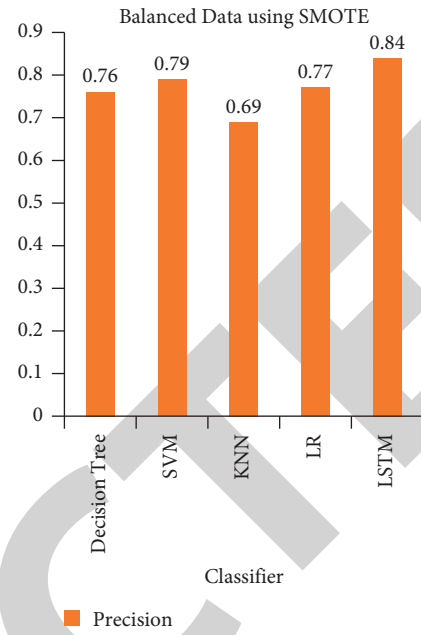


FIGURE 9: The precision value of balanced dataset using SMOTE.

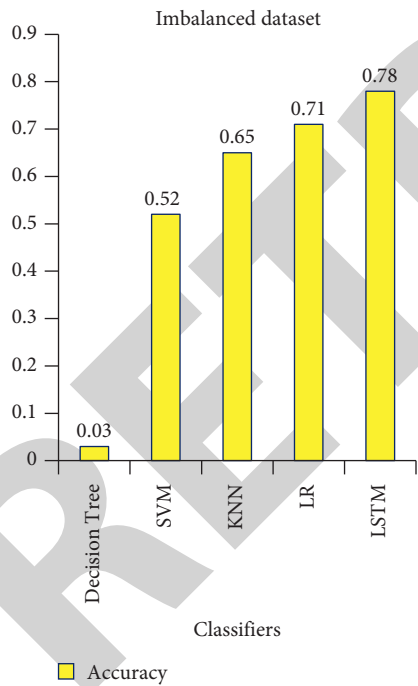


FIGURE 8: The accuracy value of imbalanced dataset.

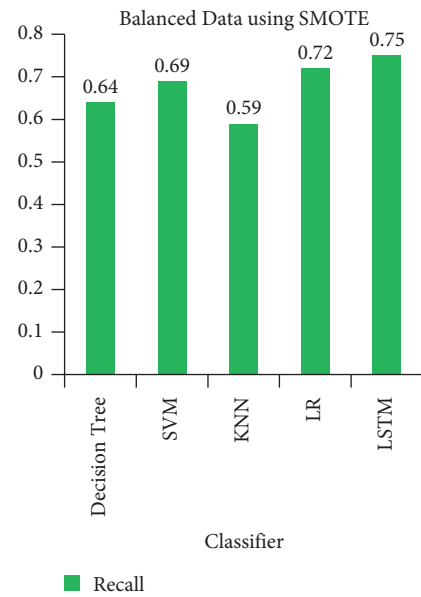


FIGURE 10: The recall value of balanced dataset using SMOTE.

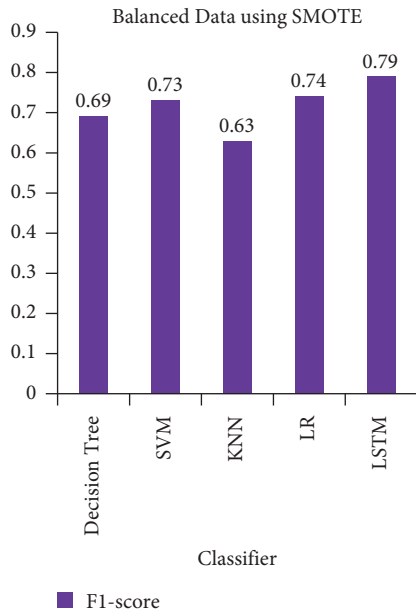


FIGURE 11: The F1-score value of balanced dataset using SMOTE.

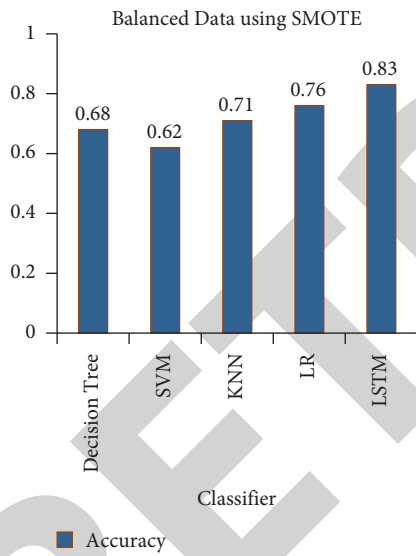


FIGURE 12: The accuracy value of balanced dataset using SMOTE.

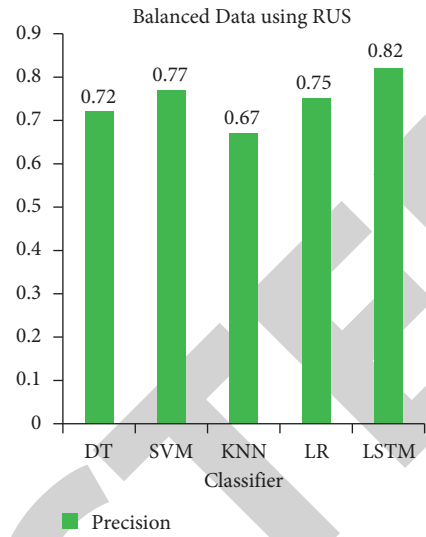


FIGURE 13: The precision value of balanced dataset using RUS.

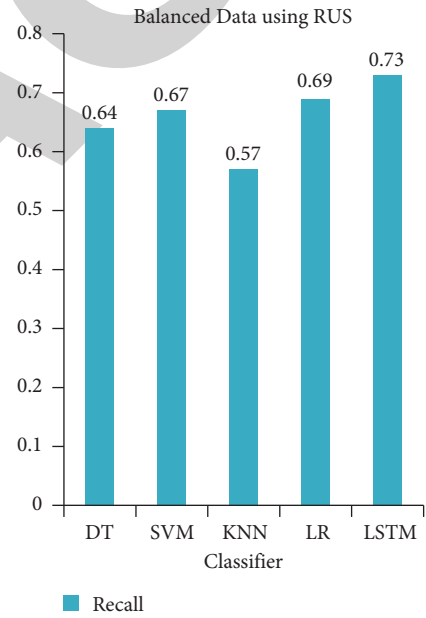


FIGURE 14: The recall value of balanced dataset using RUS.

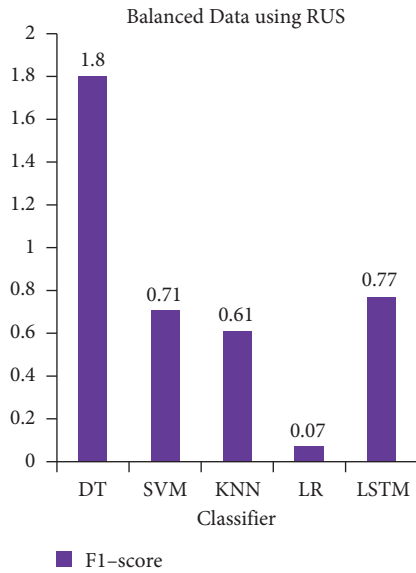


FIGURE 15: The $F1$ -score value of balanced dataset using RUS.

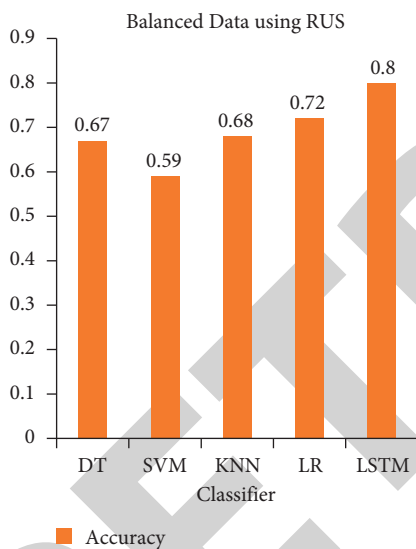


FIGURE 16: The accuracy value of balanced dataset using RUS.

For RQ3: The LSTM classification model outperforms the other baseline models in the depression detection approach.

7. Conclusion and Future Scope

This study describes various techniques for psychological analysis for depression identification, including machine learning and lexicon (dictionary and corpus)-based approaches, with the goal of detecting depression using machine learning classifiers. We also investigated the classification of both balanced and imbalanced techniques. Machine learning algorithms such as decision trees, support vector machines, and LSTM show good accuracy, but a hybrid approach may provide better results in detecting mental health disorders than individual classifiers. Younger

generations are quite active on social media, expressing their thoughts and opinions on a variety of topics and circumstances. Business firms use these platforms for their benefits, such as collecting customer reviews and complaints about services. These opinions help in identifying customer psychology (positive, negative, and neutral) toward products and any activity. The social networking platform has proved to be a great tool for identifying the mental state of social networking users. Because depression is a new problem in the society, examining social network data is helpful in detecting depression. This study discusses various obstacles in psychological analysis, but these issues will be addressed in the future as more psychological analysis research is conducted. The limitation to this research is that it can help in depression detection on social media. A large number of people of the world do not use social media because they find themselves uncomfortable on the social platform, which will remain undiagnosed.

For future work, we may identify the ability to estimate depressive disorder and daily activity, which leads to depression analysis with more accuracy. By merging LSTM and SVMs, we can create a hybrid model that can more accurately identify depression, benefiting people all across the world. As we know, LSTM can handle large datasets, and SVM performance in text classification for psychological analysis has greater accuracy; combining SVMs with LSTM may give better results for the imbalanced dataset as well as the balanced data set. In balanced datasets, SMOTE and RVS are used in the detection of depression. People may suffer from anxiety, depression, or suicide thoughts, and we can predict the same using AI and machine learning techniques. However, what about those who do not use social networking platforms? Developing a system that can identify depressive behavior in persons who do not use social media is a research objective in the future [20–24].

Data Availability

The data collected during the data collection phase are available from the corresponding authors upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. Verma, S. Gupta, and L. Goel, "A neural network based hybrid model for depression detection in twitter," in *Proceedings of the International Conference on Advances in Computing and Data Sciences*, pp. 164–175, Springer, Valletta, Malta, April 2020.
- [2] M. De Choudhary, F. A. Bakar, and N. M. Nawi, "Predicting depression using social media posts," *Journal of Soft Computing and Data Mining*, vol. 2, no. 2, pp. 39–48, 2021.
- [3] M. M. Mariani, P. V. Rodrigo, and W. Jochen, "AI in Marketing, Consumer Research and Psychology: A Systematic Literature Review and Research Agenda," *Psychology & Marketing*, vol. 39, no. 4, 2021.
- [4] J. A. Naslund, T. Deepak, A. Aditya et al., "Digital training for non-specialist health workers to deliver a brief psychological

Retraction

Retracted: Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] A. Juneja, H. Turabieh, H. Upadhyay, Z. K. Bitsue, V. T. Hoang, and K. T. Trung, "Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4602072, 8 pages, 2022.

Research Article

Optimization of Mental Health-Related Critical Barriers in IoT-Based Teaching Methodology

Abhinav Juneja ¹, Hamza Turabieh ², Hemant Upadhyay ³, Zelalem Kiros Bitsue ⁴,
Vinh Truong Hoang ⁵ and Kiet Tran Trung ⁵

¹KIET Group of Institutions, Delhi NCR, Ghaziabad, India

²Department of Information Technology, College of Computers and Information Technology, P.O. Box 11099, Taif 21944, Taif University, Saudi Arabia

³BMIET, Sonapat, India

⁴US AHO, Addis Ababa, Ethiopia

⁵Ho Chi Minh City Open University, Ho Chi Minh City, Vietnam

Correspondence should be addressed to Zelalem Kiros Bitsue; bitsue.zelalem29@gmail.com

Received 4 February 2022; Revised 28 February 2022; Accepted 11 March 2022; Published 1 April 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Abhinav Juneja 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.

Online learning has changed all elements of teaching of entire learning structure from primary to university level all around the world so that the challenges of online teaching are required to be optimized. The prominent objective of this manuscript is to optimize the issues of online teaching-learning in online education. Twelve issues of online teaching-learning are shortlisted by performing deep reviewing of the literature and grouping into three categories: “Students’ issues,” “Common issues,” and “Teachers’ issues” using the opinions of expert people. The analytical hierarchy process method is chosen for ranking of issues of online teaching. The findings can become effective in planning to get solution of the challenges of online teaching. These challenges of online teaching may lead to fragmental illness mentally over a long period of time. Because social media platforms may become an efficient tool for incorporating into online education, social media is a vital aspect of online learning. Over time, social media use may have an effect on the human brain in one way or another. The given work’s exploration of online teaching-learning challenges could lead to a social media-based examination of mental illness.

1. Introduction

The flexible nature of online classes has proven online teaching a comprehensive element in education [1]. Online classes are a learning experience as a mode of instruction and their characteristics have met the requirements of various learners [2]. Online teaching people are supposed to have a specific set of skills [3]. Education organizations are in a prime part in enhancing the effectiveness of online learning by assisting teachers, scholars, and content refinement [4]. The union government has focused on ICT at tertiary level. For better execution of academic shift (from traditional pedagogy to online system), consequences of shift are required to be addressed [5].

Online learning is going to be instrumental for the college and university studies [6]. Considerable work has been done to check the feasibility of online classes for various parameters [7]. Several papers proposed the need of supporting academic organizations for improvement in the results of online learning [8]. Literature is deficient in case of dealing with the challenges of empowerment of online instructors and integrating techniques in methodological inquiry. Moreover, the past related work on online learning is hardly vocal about the critical optimization of competency-based teacher learning pattern [9]. Teachers build social media groups for each set of students to stream live sessions on social media platforms in order to answer questions and manage assignments. Peers might use social media platforms

as an announcement board for a group of students. The use of digital content in conjunction with social media groups and channels may be beneficial. Social media is a powerful tool for aesthetically appealingly presenting a sequence of photographs or graphics.

- (a) Therefore, it is required to address the challenges and issues of online teaching.
- (b) Thus, the prime objectives of the current work are the exploration of challenges and issues of online teaching in Indian perspective.
- (c) Literature survey and experts' opinions are utilized to choose challenges and issues of online teaching in Indian perspective.
- (d) Then, AHP method is chosen to analyze the challenges and issues of online teaching in Indian perspective.
- (e) A set of issues of online teaching-learning has been shortlisted in three categories.
- (f) These critical challenges of online teaching from three different categories have been compared and ranked.
- (g) The findings can become effective in planning to get solution of the challenges of online teaching.
- (h) These challenges of online teaching may lead to fragmental illness mentally over a long period of time.
- (i) Since social media has the potential to be an effective method for inculcating into online learning, it is an important part of online learning.
- (j) Social media platforms are more or less addictive, and they have been linked to psychosocial problems and partial mental disease.

AHP was proposed by Saaty in 1977. This is being applied these days very frequently to compare possible options. The priorities calculated by the AHP method can be used to rank the factors. This consistency-based method is easy, robust, reliable, and mathematical.

The paper has been settled as follows: critical reviewing from concerned literature has been done in the next section. Challenges and issues of online teaching are recognized in Section 3. The proposed method has been discussed in Section 4. The calculation has been shown in Section 5. The concluding statement and scope of future work have been stated in Sections 6 and 7, respectively.

2. Review of Relevant Literature

Liang and Chen studied issues, potential, and challenges of online education for stakeholders while designing online activities [10]. Tai focused on the interaction in the teaching and highlighted some of the challenges and issues of online scholars as opposed to off-line scholars with the expressions of the face and gesture to further support interaction with staff [11].

Bawane and Spector reviewed the transformation learn theory [3]. Taylor explained that "competency and standards

driven" work in online learning as techno-centric approach has developed one-size-fits-all concept for online teachers [12]. Valli considered online teachers who provide knowledge with theory and research for analysis and regular improvement in the level of teaching [13].

Kebritchi et al. concluded that academic organizations require to deal with the issues in online teaching and boost the efficiency of online learning [4]. Martin reviewed hundreds of articles from several online sources to much concentration on the student domain [14]. The policymakers are supposed to ensure the access of workable communication tools and better digital educational exposure and boost technology-assisted education for learners to manage and make the education structure better [5].

There are various statistical methods, which may be used for finding the relational priority of the influencing factors. The analytic hierarchy process (AHP) is an efficient method of handling the complicated decision problems when relevant data are tedious to be analyzed [15]. The AHP methodology, originally deliberated through Saaty, makes the multi-criteria decision analysis, a hierarchy for the choice [16]. AHP methodology converts the decision issue into a hierarchical model, and then, further it outlines a goal for ranking and determines various criteria and subcriteria [17]. The feature of AHP to contemplate criteria and options makes it an acceptable methodology for the industrial applications [18]. The analytic hierarchy methodology is utilized in various applications such as political science, finance, and technology [19]. AHP is beneficial for getting an optimization value with different indicators [20]. A standard objective of survey analysis, ranging from few experts to hundreds of interviewed individuals in AHP, is to gather data representative of a population, and determining the size of the sample is very crucial in this method. Table 1 briefs few of the recent research efforts to address challenges and issues of online learning and relevant use of AHP methodology to rank the parameters.

The literature review indicates that challenges of online teaching are important from the point of view of academics, but a complicated work to be done. Thus, it is necessary to recognize issues of online teaching. Section 3 has the list and details of challenges and issues of online teaching.

3. Identification of Challenges and Issues of Online Teaching

Various databases (papers published in journals/presented in conferences having "Online Teaching" and "Online Classes" keywords) were reviewed, and then, twelve types of issues were categorized into three categories as per expert's opinions. The key issues include "Technical difficulties with online teaching tools," "Over-reliance on the educator," "Students alienation," "LMS complications," "Lack of real face to face interaction," "Unstable/intermittent network connection issues," "Data privacy/security," "Time-consuming," "Staying connected with passive students," "Creating/editing/sharing online teaching content," "Fostering an affective online learning climate," and "Laboratory demonstration." The identified critical barriers are

TABLE 1: Latest research of challenges and issues of online teaching.

Author	Objective and outcome
[4]	They studied about the prime issues and tactics that influence the quality of higher online education. They checked the literature that suggested to deal the challenges for online instructors They highlighted challenges confronting online teaching
[10]	(i) Quality assurance and standards (ii) Commitment versus innovation (iii) Copyright and intellectual property

categorized into three categories as “Students’ issues,” “Common issues,” and “Teachers’ issues.”

3.1. Students’ Issues. Online teaching should pay special emphasis on student interaction [21]. The scholars’ connection with the study management system mainly influences the student experience in case of online education [22]. Kebritchi et al. identified three types of issues in online education as issues related to teachers and students [4]. The issues related to teachers are in the four specific categories of “Technical difficulties with online teaching tools,” “Over-reliance on the educator,” “Students alienation,” and “LMS complications.”

3.1.1. Technical Difficulties with Online Teaching Tools (TDWOTT). The students lack sufficiency of time and fire of eagerness to learn newer things for online learning. The students show dissatisfaction with the technical complications installed by academic organizations for online mode, and as such, students devote much time on learning newer technologies.

3.1.2. Over-Reliance on the Educator (OROE). Most online deliveries, keeping the academic study matter in a general structure in the teaching tool, make effort to develop a comprehensive academic system that the scholars assume from course-wise and module-wise. Excessive dependency on online education becomes the issue for both instructor and learner.

3.1.3. Students Alienation (SA). Students feel very disconnected and isolated in the mode of online teaching, which influences learning. It is probably the consequence like sense of transactional gap in online mode and deficiency of communication cues (face, voice, etc.).

3.1.4. LMS Complications (LMSC). Technological complications may be because of the poor coordination of operation of the hardware and software in better online teaching.

3.2. Common Issues. There are few issues in online learning, which are concerned with students and teachers. The issues related to teachers are in the four specific categories of “Lack of real face to face interaction,” “Unstable/intermittent network connection issues,” “Data privacy/security,” and “Time-consuming.”

3.2.1. Lack of Real Face to Face Interaction (LORF2FI). The durable learning requires the capability for adaption to new era challenges. Effective feedback techniques are demand of hour.

3.2.2. Unstable/Intermittent Network Connection Issues (UINCI). Since online learning requires students and teachers to have access to technology, the challenges of technological accessibility cannot be ignored. Learners find the online classes as a discriminatory way by rendering them disproportionate with their mates regarding online learning technologies. Access to newer technical implications for online learning activities is hurdled by slow data speed and unstable network connectivity.

3.2.3. Data Privacy/Security (DPS). In the fast application of newer technology, educational institutes do gather huge personal info on learners and teachers and this has been creating sufficient troubling questions about data privacy.

3.2.4. Time-Consuming (TC). Even after a lot of merits related to online teaching videos, issue of long-duration videos for teaching also exists. Learners’ opinion on online study material is “bulky, cumbersome, and time-consuming.” The duration of any online video is not directly proportional to the fraction of videos watched by learners.

3.3. Teachers’ Issues. Among several ones, the pedagogical role is the most important [3]. The guide’s part in the online learning needs much comprehensive as instructor work with pedagogical issues dealing with different disciplines and technologies [23]. It is notable that teachers do structure discourse of topics and dedicate huge time support for scholars in online learning [24]. The issues related to teachers are in the four specific categories of “Staying connected with passive students”; “Creating/editing/sharing an online teaching content”; “Fostering an affective online learning climate”; and “Laboratory demonstration.”

3.3.1. Staying Connected with Passive Students (SCWPS). There are different types of learners. Passive learners acquire the information but do not show interest to look for the practical applications. It is required to deliver in such a fashion that may indulge non-active students in online classes to assist to get all possible advantages.

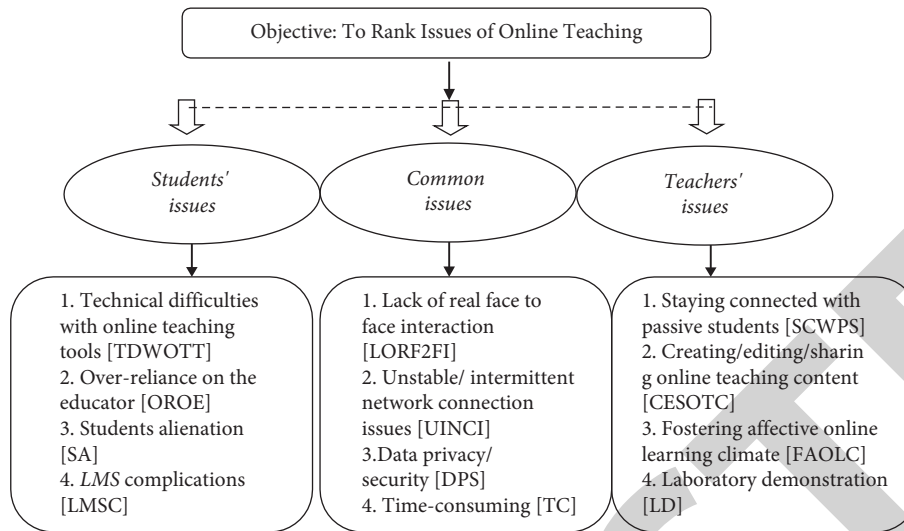


FIGURE 1: AHP-based hierarchical model to analyze issues of online teaching.

3.3.2. *Creating/Editing/Sharing an Online Teaching Content (CESOTC)*. Simplification of assignment is required for better understanding of learners. Teachers face issues with developing teaching matter content on LMS. Course redesign for online classes is about the technical complication of designing new teaching material and establishing favorable learning climate.

3.3.3. *Fostering an Affective Online Learning Climate (FAOLC)*. For creating impressive online education environments, a systemized method is needed so that evaluation of the impact of education can be analyzed.

3.3.4. *Laboratory Demonstration (LD)*. Teachers express huge emphasis on the laboratory activities for the students in online studies. Teachers demand the demonstration technologies in practice.

4. Methodology

To obtain the prime aim of our research work, which is the optimization of challenges of online teaching, 12 issues of online teaching are subcategorized into 2 levels (Students' issues, Common issues, Teachers' issues) as mentioned in Figure 1. The figure displays the overall ranking corresponding to complete hierarchical structure of issues taken in the current research work.

Optimization of twelve issues of online teaching is taken into three categories as Students' issues, Common issues, Teachers' issues. Analytical hierarchical process (AHA) is a hierarchy-based process; therefore, it is the best-suited multi-criteria decision-making technique for our study with a two-level categorized structure. A Web-based survey was also done with inputs from 50 individuals using Google Form to take the unbiased inputs of stakeholders to understand the priorities of the related people for various issues as potential challenges. This survey response has been duly used as a weight in the ranking of issues. Thereafter, the AHP

technique has been employed in ranking these essential barriers.

4.1. *Analytical Hierarchy Process*. AHP technique calculates the consistency employing a consistency index. It permits customers to analyze the relational weight age of numerous options in context with given parameters. The analytic hierarchy process (AHP) is a math and psychology-based system for organizing and analyzing complicated decisions. The AHP has got used in many domains involving selections, during which prioritization or foretelling is required. The analytic hierarchy process (AHP) is a process that measures through pairwise comparisons and relies on expert opinion to derive priority scales. AHP permits the evaluation of comparative bias of several criteria counter to certain criteria spontaneously. It is one of the most extensively used tools for making multiple-criteria decisions. The AHP is being practiced in government policymaking, R&D, academic activities, business decisions, defense, and other many more verticals wherein decisions are made taking into account choice, preference, or prediction [25]. AHP is a versatile method for simplifying the complex MCDM issues in an organized manner by disintegrating a compound decision problem into an ordered array of relational decision members (evaluation criteria, sub-criteria, and substitutions).

The analytic hierarchy process on the basis of variable weight age methodology may adequately deal with the fixed weight ages of conventional analytic hierarchy process [26]. The analytic hierarchy process has three fundamentals such as problem decomposition, comparison judgment, and synthesis of comparative influence or ranks [27–29]. Many systems employ the AHP algorithm approach to make tough decisions. The AHP method might be considered when making decisions because data processing utilizing the AHP method will give you the greatest advice for making a big decision [30]. AHP is a versatile tool that may be used alone or in combination with other tools to solve construction decision-making issues [17, 31].

TABLE 2: Comparison scale in AHP (Foteinopoulos et al.).

Impact of importance	Classification	Description
1	Equally significant	Two actions endorse likewise for an objective
3	Feeble significance of one over another	Familiarity and evaluation partially weigh an action to the other
5	Indispensable significant	Familiarity and evaluation strongly weigh an action to the other
7	Established significance	An action is predominantly being biased and its authority is validated
9	Complete significance	This is the maximum feasible degree of conformance
2, 4, 6, 8	Intermediary tenets	In case of a need for some negotiation

TABLE 3: RI values of n criteria (Saaty).

n	1	2	3	4	5
RI	0.00	0.00	0.58	0.90	1.12

TABLE 4: Pairwise comparison matrix of criteria.

Criteria	Students' issues	Common issues	Teachers' issues	Priority matrix	Rank
Students' issues	1	2	5	0.580	1st
Common issues	0.5	1	3	0.317	2nd
Teachers' issues	0.2	0.333	1	0.110	3rd

Maximum eigenvalue = 3.006 and C.I. = 0.003.

TABLE 5: Pairwise comparison matrix of students' issues.

Students' issues	TDWOTT	OROE	SA	LMSC	Priority matrix	Rank
TDWOTT	1	7	9	3	0.606	1st
OROE	0.142	1	4	0.500	0.125	3rd
SA	0.111	0.250	1	0.200	0.047	4th
LMSC	0.333	2	5	1	0.220	2nd

Maximum eigenvalue = 4.114 and C.I. = 0.038.

For the optimization of the challenges and issues of online teaching using the AHP approach, the identified 12 issues have been compared and evaluated using analytical hierarchical process method and further ranked in terms of priorities [32]. After identification of the hierarchical structure of AHP, further the AHP can be invoked through the following sequence of steps [33–35].

- (a) Selection of criteria (n number of criteria).
- (b) Thorough evaluation of the relational value via a sequence of pairwise associations corresponding to every decision condition employing a 9 pointer rule (1–9) to indicate their observations. The significance of values from 1 to 9 is tabulated in Table 2.
- (c) Building an array of pairwise comparative square matrices ($n \times n$) corresponding to every element in decision-making.
- (d) The decision components are compared pairwise by employing $n (n - 1)$ judgments.
- (e) There is an auto-assignment of the reciprocals corresponding to every pairwise comparison.
- (f) Normalization of the generated comparison matrix is performed.
- (g) Hierarchical combination is obtained to generate the weights of the criteria.

- (h) Eigenvalues λ can be calculated as the ratio of weighted sum value and criteria weight for each row of consistency matrix.
- (i) Principal eigenvalue is the average of all eigenvalues. Here, we calculated the principal eigenvalue from the individual eigenvalues using the following equation:

Principal Eigen Value,

$$\lambda_{\max} = \sum \frac{\lambda}{n} \tag{1}$$

- (j) The calculation of consistency index (AHP calculation) can be done using the following equation:

$$\text{Consistency Index CI} = \frac{(\lambda_{\max} - n)}{n - 1} \tag{2}$$

- (k) Random index (RI) is the value attained through arbitrarily generated pairwise comparison matrix. The significant quantum values of random index are given in Table 3.
- (l) The consistency ratio (AHP calculation) is the ratio of consistency index and random index and is given as follows:

$$\text{Consistency Ratio (CR)} = \frac{CI}{RI} \tag{3}$$

TABLE 6: Pairwise comparison matrix of common issues.

Common issues	LORF2FI	UINCI	DPS	TC	Priority matrix	Rank
LORF2FI	1	2	9	7	0.531	1st
UINCI	0.5	1	8	5	0.338	2nd
DPS	0.111	0.125	1	2	0.072	3rd
TC	0.142	0.200	0.5	1	0.059	4th

All four diagonal elements of Table 6 are unity value, and values in all other elements are on the basis of inputs. Maximum eigenvalue = 4.182 and C.I. = 0.061.

TABLE 7: Pair wise comparison matrix of teachers' issues.

Teachers' issues	SCWPS	CESOTC	FAOLC	LD	Priority matrix	Rank
SCWPS	1	2	0.5	0.25	0.134	3rd
CESOTC	0.5	1	0.2	0.142	0.066	4th
FAOLC	2	5	1	0.5	0.284	2nd
LD	4	7	2	1	0.518	1st

Maximum eigenvalue = 4.014 and C.I. = 0.00466.

TABLE 8: Computation of overall weight of issues of online teaching.

Category name	Weight of categories	Rank	Identified issues of online teaching	Local weight of issues	Overall weight of issues	Overall rank of issues
Students' issues	0.580	1st	Technical difficulties with online teaching tools (TDWOTT)	0.606	0.3514	1st
			Over-reliance on the educator (OROE)	0.125	0.0725	5th
			Students alienation (SA)	0.047	0.0273	8th
			LMS complications (LMSC)	0.220	0.1276	3rd
Common issues	0.317	2nd	Lack of real face to face interaction (LORF2FI)	0.531	0.1683	2nd
			Unstable/intermittent network connection issue (UINCI)	0.338	0.1071	4th
			Data privacy/security (DPS)	0.072	0.0228	9th
			Time-consuming (TC)	0.059	0.0187	10th
Teachers' issues	0.110	3rd	Staying connected with passive students (SCWPS)	0.134	0.0148	11th
			Creating/editing/sharing an online teaching content (CESOTC)	0.066	0.0073	12th
			Fostering an affective online learning climate (FAOLC)	0.284	0.0312	7th
			Laboratory demonstration (LD)	0.518	0.0569	6th

TABLE 9: Comparison of this study with some recent contributions.

Paper	Outcome	Present work
[5]	Their study found the challenges and issues in online teaching as the unstable network connection, intermittent signal issues, and offline conduction of classes, a lack of motivation	Our study identified the twelve key issues in online education and ranked them using MCDM technique, on the basis of a Web-based survey with inputs from fifty stakeholder individuals
[22]	They examined some of the challenges between online students and teachers. They suggested to develop flexible learning activities about course topics	Critical challenges and issues of online teaching from three different categories (Students' issues, Common issues, and Teachers' issues) have been compared by analytical hierarchy process (AHP) methodology

The consistency ratio must be, at most, 10% to maintain consistency in calculated judgments. A consistent ratio (CR) value of 0.1 or less is acceptable, else termed as inconsistent [17].

5. Results

On the basis of the data gathered by Web-based survey regarding challenges and issues of online teaching, matrices

have been formulated and further calculation for calculating priorities is done by the concept of AHP. The formation of the pairwise comparison of the different categories of issues of online teaching is presented in Table 4. Table 4 has diagonal elements unity, and other elements are on the basis of survey values.

Table 4 results show that "Students' issues (0.580)" was the most important category of issues of online teaching followed by "Common issues (0.317)" and "Teachers' issues

(0.110).” Further, various issues of online teaching have been ranked for each category. In Table 5, the issues under category “Students’ issues” of online teaching had been checked for hierarchy.

“Technical difficulties with online teaching tools (0.606)” had been reported most important issue under category “Students’ issues” of online teaching, followed by “LMS complication (0.220),” “Over-reliance on the educator (0.125),” and “Students alienation (0.047)” in Table 5. In Table 6, issues under category “Common issues” had been checked for hierarchy.

“Lack of real face to face interaction (0.531)” had been reported most important issue under category “Common issues” of online teaching, followed by “Unstable/intermittent network connection issues (0.338),” “Data privacy/security (0.072),” and “Time-consuming (0.047)” as shown in Table 6. The issues under category “Teachers’ issues” are checked to hierarchy scale below.

Table 7 shows that “Laboratory demonstration (0.518)” had been noted as the most effective issue in category “Teachers’ issues” of online teaching, followed by “Fostering an affective online learning climate (0.284),” “Staying connected with passive students (0.134),” and “Creating/editing/sharing an online teaching content (0.066).”

The consistency ratio calculated values are in the permissible range for various pairwise comparison matrices shown in Table 4 to Table 7.

We have evaluated the overall weight of every issue with the help of local weight of the issue and calculated the product by respective category’s weight. After computing, overall weights of issues are tabulated in Table 8, showing that “Technical difficulties with online teaching tools,” “Lack of real face to face interaction,” and “LMS complications” are presented top three issues on the basis of overall weight age of issues.

Category ranking and overall ranking of all the issues are done in Table 8. Firstly, local weights of all twelve factors have been computed (as shown in the fifth column of Table 8), and then, the weight of their relevant category has been multiplied to get the overall weights of the factors (as shown in the sixth column of Table 8). Table 8 demonstrates that “Creating/editing/sharing an online teaching content,” “Staying connected with passive students,” and “Time-consuming issue” are presented bottom three issues based on the overall ranking of issues.

The current work has been fairly compared with the contribution from other researchers in the similar domain. Though there is no success as on date to curb the virus, some of the researchers have done wonderful efforts in uncovering challenges and issues of online teaching, which may be considered to be helpful for policymakers. The outcomes are in Table 9.

6. Conclusions

This study has shown a MCDM method for optimization of challenges and issues of online teaching using the AHP technique, with twelve issues (“Technical difficulties with online teaching tools,” “Over-reliance on the educator,”

“Students alienation,” “LMS complications” “Lack of real face to face interaction,” “Unstable/intermittent network connection issues,” “Data privacy/security,” “Time-consuming,” “Staying connected with passive students,” “Creating/editing/sharing online teaching content,” “Fostering an affective online learning climate,” and “Laboratory demonstration”). A total of twelve critical challenges and issues of online teaching from three different categories have been compared and ranked. The current work has explored the possible issues factors, which can become a key to address the challenges of online teaching. It is obvious from the findings of the work that factors of *Students’ issues*, which are influenced by factors such as “Technical difficulties with online teaching tools” and “LMS complications,” prove to be the most critical challenges of online teaching. These observations may relate the mental maladies because of social media linked with online mode of teaching and learning.

7. Scope for Future Work

In our study, we have applied only one technique but some other methods such as DEMATEL, TOPSIS, best-worst method, and fuzzy AHP, which can be instrumental for similar issues. Further results by applying many techniques can be compared.

Data Availability

The data will be available from the author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Taif University Researchers Supporting Project Number TURSP-2020/125, Taif University, Taif, Saudi Arabia.

References

- [1] C. Li and B. Irby, “An Overview of online education: attractiveness, benefits, challenges, concerns, and recommendations,” *College Student Journal, Part A*, vol. 42, pp. 449–458, 2008.
- [2] R. Clerehan, J. Turnbull, T. Moore, A. Brown, and J. Tuovinen, “Transforming learning support: an online resource centre for a diverse student population,” *Educational Media International*, vol. 40, no. 1-2, pp. 15–32, 2003.
- [3] J. Bawane and J. M. Spector, “Prioritization of online instructor roles: implications for competency-based teacher education programs,” *Distance Education*, vol. 30, no. 3, pp. 383–397, 2009.
- [4] M. Kebritchi, A. Lipschuetz, and L. Santiago, “Issues and challenges for teaching successful online courses in higher education,” *Journal of Educational Technology Systems*, vol. 46, no. 1, pp. 4–29, 2017.
- [5] L. Mishra, T. Gupta, and A. Shree, “Online teaching-learning in higher education during lockdown period of COVID-19 pandemic,” *International Journal of Educational Research Open*, vol. 1, Article ID 100012, 2020.

Retraction

Retracted: Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] G. Murugesan, T. I. Ahmed, M. Shabaz et al., "Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8555489, 10 pages, 2022.

Research Article

Assessment of Mental Workload by Visual Motor Activity among Control Group and Patient Suffering from Depressive Disorder

G. Murugesan ¹, Tousief Irshad Ahmed ², Mohammad Shabaz ³, Jyoti Bhola ⁴,
Batyrkhan Omarov ^{5,6,7}, R. Swaminathan ⁸, F. Sammy ⁹, and Sharmin Akter Sumi ¹⁰

¹Department of Computer Science and Engineering, St. Joseph's College of Engineering, Chennai 600119, India

²Department of Clinical Biochemistry, Sher-i-Kashmir Institute of Medical Sciences, Soura, Srinagar, J&K, India

³Model Institute of Engineering and Technology, Jammu, J&K, India

⁴Electronics & Communication Engineering Department, National Institute of Technology, Hamirpur, India

⁵Al-Farabi Kazakh National University, Almaty, Kazakhstan

⁶International University of Tourism and Hospitality, Turkistan, Kazakhstan

⁷Suleiman Demirel University, Kaskelen, Kazakhstan

⁸Saveetha School of Engineering, Chennai, Tamil Nadu, India

⁹Department of Information Technology, Dambi Dollo University, Dembi Dolo, Welega, Ethiopia

¹⁰Department of Anatomy, Bangabandhu Sheikh Mujib Medical University, Dhaka, Bangladesh

Correspondence should be addressed to Tousief Irshad Ahmed; khagankhan@gmail.com and F. Sammy; sammy@dadu.edu.et

Received 3 February 2022; Revised 1 March 2022; Accepted 8 March 2022; Published 31 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 G. Murugesan 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.

Major depressive disorder (MDD) is a mood state that is not usually associated with vision problems. Recent research has found that the inhibitory neurotransmitter GABA levels in the occipital brain have dropped. *Aim.* The aim of the research is to evaluate mental workload by single channel electroencephalogram (EEG) approach through visual-motor activity and comparison of parameter among depressive disorder patient and in control group. *Method.* Two tests of a visual-motor task similar to reflect drawings were performed in this study to compare the visual information processing of patients with depression to that of a placebo group. The current study looks into the accuracy of monitoring cognitive burden with single-channel portable EEG equipment. *Results.* The alteration of frontal brain movement in reaction to fluctuations in cognitive burden stages generated through various vasomotor function was examined. By applying a computerised oculomotor activity analogous to reflector image diagram, we found that the complexity of the path to be drawn was more important than the real time required accomplishing the job in determining perceived difficulty in depressive disorder patients. The overall perceived difficulty of the exercise is positively linked with EEG activity measured from the motor cortex region at the start of every experiment test. The average rating for task completion for depression patients and in control group observed and no statistical significance association reported between rating scale and time spent on each trial ($p = 1.43$) for control group while the normalised perceived difficulty rating had 0.512, 0.623, and 0.821 correlations with the length of the pathway, the integer of inclination in the pathway, and the time spent to complete every experiment test, respectively ($p < 0.0001$) among depression patients. The findings imply that alterations in comparative cognitive burden levels during an oculomotor activity considerably modify frontal EEG spectrum. *Conclusion.* Patients with depression perceived the optical illusion in the arrays as weaker, resulting in a little bigger disparity than individuals who were not diagnosed with depression. This discovery provided light on the prospect of adopting a user-friendly mobile EEG technology to assess mental workload in everyday life.

1. Introduction

The human cognitive system is of interest to researchers working on computational applications for a variety of reasons across many fields. For more than three decades,

researchers have been seeking to overcome this challenge to understand human cognitive capability [1, 2]. In major depressive disorder, abnormalities in motor activity are linked to other characteristics such as a lack of interest in daily chores (MDD). As a consequence, patients with

current MDD have lower motor activity [3, 4], and treatment response has been linked to increased motor activity [5, 6]. Here, we also brief about other mental illness disorders.

- (i) Anxiety: unlike typical feelings of apprehension or anxiety, anxiety disorders are characterized by excessive fear or anxiety. The most common mental ailments are anxiety disorders. Anxiety sufferers may strive to avoid circumstances that aggravate or cause their symptoms. Workplace productivity, schoolwork, and personal relationships could all be harmed. Generalized anxiety disorder, panic disorder, particular phobias, agoraphobia, social anxiety disorder, and separation anxiety disorder are all examples of anxiety disorders.
- (ii) Panic disorder: the most prevalent symptom of panic disorder is recurrent panic attacks, which are characterized by an overwhelming combination of physical and psychological anguish. Many of these symptoms, such as trembling or shaking, palpitations, pounding heart or high heart rate, perspiration, feeling alienated, and fear of losing control, might occur simultaneously during an attack. [7].
- (iii) Posttraumatic stress disorder (PTSD): people who have been threatened with death, sexual violence, or serious injury, or who have experienced or witnessed a traumatic incident such as a natural disaster, a catastrophic accident, a terrorist attack, war/combat, or rape, may develop posttraumatic stress disorder (PTSD). People with PTSD have strong, unsettling thoughts and feelings about the traumatic event that last long after it has happened. They may have flashbacks or dreams about the event, as well as feelings of sadness, fear, or anger, as well as a sense of being distant or estranged from people [8].
- (iv) Dissociative Disorders: dissociative disorders impact memory, identity, emotion, perception, conduct, and sense of self. Dissociative symptoms can impact every element of a person's mental health. Detachment or feeling as if one is outside one's body, as well as memory loss or amnesia, are examples of dissociative symptoms. Dissociative illnesses are typically linked to traumatic experiences in the past [9].

Neuropsychological problems are common symptoms in people suffering from chronic fatigue syndrome (CFS). The goal of the study was to compare the cognitive abilities of CFS patients to those of two illness comparator groups (nonmelancholic depression and acute infection), as well as healthy controls [10, 11]. However, quantitative measurements of motor activity have become more common in recent years. On the one hand, miniaturized accelerative devices with great reliability and validity have been developed as a result of technological advancements [12]. Recall questionnaires, on the other hand, were found to have only limited validity in metaanalyses [13]. Women are diagnosed with depression at a higher incidence than men, which may

be due to women's proclivity to seek treatment as soon as possible [14, 15]. In accordance with this, latest long-term telehealth strategies use objectively assessed cognitive movement as a replacement for depressed clinical signs.

The relationship between mental and physical health is mentioned here. Chronic physical illnesses get worsened by poor mental health. People who are suffering from major mental diseases are more likely to develop chronic physical ailments. Those patients who have severe physical disorders are more prone to develop mental problems [16, 17]. The connection among psychological and physical health is well established [18]. While new technologies are being developed to observe numerous indices of physical health with wearable and wireless devices, such as pulse, blood pressure, and body temperature [19, 20], the alternatives for monitoring cognitive burden are far more limited. The difficulty of reliably measuring and quantifying mental states is one of the key obstacles.

A transportable EEG system with several functional prototypes has recently been suggested. The preponderance of these recommended mobile EEG devices offers consumers a wireless and feasible option for monitoring their real-time brain processes while maintaining the excellent transient response of data gathered by conventional laboratory EEG equipment [21]. Mobile electroencephalography (mobile EEG) is the next neuroscientific device for studying real-time activity in the brain that is fairly affordable, noninvasive, and has portability. It takes an advantage of cutting-edge hardware, as well as the long-established advantages of traditional EEG and recent signal processing advancements [22, 23]. The EEG signals obtained through such mobile equipment and applications have been convenient in assessing and determining the mental conditions of users [24]. As a result, the purpose of this study was to see if a single-channel dry sensor mobile EEG device could be used to assess cognitive burden. In addition to the conventional wet EEG system, which emerges with various tech challenges such as skin and electrode readiness, user discomforts, and lengthy preparation time, the dry sensor EEG system's ergonomic design and high flexibility enable us to evaluate the psychological condition of users during an unrestrained innate activity, in a physical world establishing well outside research lab.

In major depressive disorder (MDD), both quantitative and qualitative cognitive evaluations are changed, but there is minimal correlation between them especially in those with residual mood symptom everyday memory and related cognitive concerns expressed by people should prompt further research into sleep quality and (or) additional health related complications, as these conditions are known to impact mental enactment [25, 26]. Sleep, for instance, may have mental confusion, disorientation, or delirium effects [27], and the lack of good quality sleeping hours impacts a variety of cognitive domains [28]. Depression is not a natural element of getting older and should never be overlooked or neglected. Unfortunately, geriatric depression is commonly misunderstood and neglected, but they may be hesitant to seek treatment. A few of the depression symptoms in the elderly may differ or be less evident, which could be memory

problems or personality shift aches, and aches all throughout the body [29, 30]. Although there is evidence that visual processing is disrupted in MDD, it is unclear whether the loss is in retinal or cortical processing (or both). A retinal deficiency would imply more localised alterations in visual information processing, whereas a cortical deficit could indicate a more general abnormality. Electrophysiological and behavioural examinations of contrast perception in major depressive disorder (MDD) have produced mixed results, and most previous behavioural research has examined contrast detection at the threshold. From a practical perspective, visual contrast tests are a quick, easy, and noninvasive procedure that could be refined further to serve as biomarkers for aberrant visual information processing in depression. This discovery provided light on the prospect of adopting a user-friendly mobile EEG technology to assess mental workload in everyday life.

1.1. Importance of the Study. Two of the most prevalent persistent problems among MDD patients who achieve clinical remission are cognitive dysfunction and functional impairments. Our findings include a mental workload assessment of MDD patients and a control group, as well as moderate impact sizes in the neurocognitive domains of processing speed, attention, executive function, learning, and memory.

1.2. Objective. The objective of the research is to evaluate mental workload by a single channel EEG approach, brain rhythms through visual-motor activity, and comparison of parameter among depressive disorder patient and in control group.

2. Method

2.1. Participants. For this investigation, a total of twenty participants were enrolled. Ten individuals have healthy or actually corrected eyesight and have never had a neurodegenerative or psychiatric problem. An average age of control group is 28.3 ± 3.2 years. While other ten participants have average age of 27.3 ± 2.8 years. The beck depression inventory (BDI) [28] was conducted before pretesting at a psychologist's institute, and these subjects were enrolled based on their results. For the depressive disorder patients, BDI score above 30 was considered (Table 1). The BDI toll was employed since it is a frequently used depression assessment measure in research and clinical practise. Participants with elevated levels of depression and nondepressed controls will be easily identified using this assessment. For validity and reliability, the BDI has been standardised in several countries and cultures [31, 32].

Ethical clearance: the data collecting format and experimental methodologies to prospectively acquire thorough information on acute stroke patients in order to conduct this study were approved by the Institutional Committee as well as a partner hospital. Each patient was required to complete an informed consent form. The entire data gathering procedure was carried out in compliance with the methodology

TABLE 1: Beck depression inventory score (BDI).

Levels of depression as per beck depression inventory	Total score
These ups and downs are considered normal	1–10
Mild mood disturbance	11–16
Borderline clinical depression	17–20
Moderate depression	21–30
Severe depression	31–40
Extreme depression	Over 40

and standards that had been accepted. On request, a full form will be provided.

Healthy individuals are administered through a phone screening session.

2.1.1. Inclusion Criteria. Both control and depressive patients had to be highly right handed (a rating of 34 on Chapman & Chapman's 1987 39-point scale). Age was <18 years for both groups. For the depressive disorder patients, BDI score above 30 was considered.

2.1.2. Exclusion Criteria

- (i) For Control group: participants have no background of head injury with failure of consciousness lasting more than 10 minutes, cerebrovascular accident, seizure disorders, electroconvulsive therapy, usage existing psychiatric drugs, or frequent suicidal behaviour (though respondents with present mental and rehabilitation therapy were consider). The healthy controls had no current or prior history of mental illness, and they said that they had no relatives with Axis I disorders such as anxiety, panic disorder, dissociative disorders, or posttraumatic stress disorder.
- (ii) BDI score was >30, as well as other comorbid conditions and mental illnesses were acute suicidality, depressive disorder with psychotic symptoms, bipolar disorder, organic psychiatric disorders, substance-abuse disorders, schizophrenia, schizoaffective disorders, and borderline personality disorder for the depressive disorder group.

2.2. Data Trial. Ten trials of computerised visual motor activity were completed by all subjects. Milner was the first to utilise the mirror drawing challenge to see how cognitive deficits affected learning and developing motor skill [33]. In order to complete the tasks correctly, respondents were had to develop a unique set of visual-motor connections (i.e., moving their hands in the opposite direction as depicted on the computer monitor) and suppress the well-learned relationship between visual and motor coordination [34]. Participants in this study were instructed to accomplish a computerised activity comparable to oculomotor activity and reflector image diagrams that was designed in MATrix LABORatory (MATLAB) using the PsychToolbox [35, 36]. Figure 1 shows the MATrix LABORatory single-buffered drawing model.

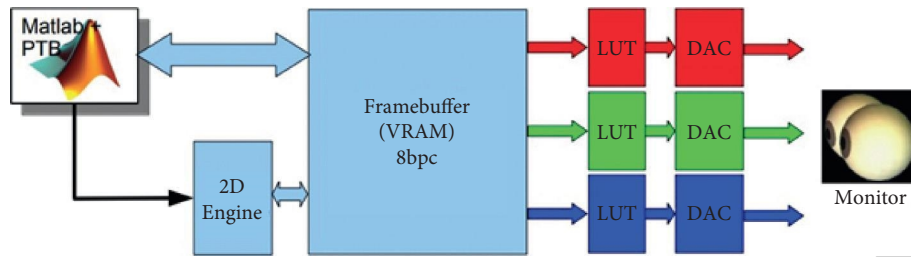


FIGURE 1: MATrix LABoratory single-buffered drawing model.

Every session required participants to use a mouse to sketch the boundary of the given pictures, as shown in Figure 2. The mouse's left and right movements were reversed by the programme. The participants were requested to hold down the left mouse key during the tracing, and the tracing was discontinued when they let off of the key. Prior to signal data acquisition, all subjects were given 10 pre-testing chance to acquaint themselves for predesigned assignment. Studied subjects were instructed to finish the reproduction or draw over in less than 300–360 seconds in every signal data acquisition test. All subjects were instructed to trace as precisely as possible within the boundary and to return to the border at the same point where the tracing left it. On a psychometric response scale for oculomotor activity of 1 to 7, participants were instructed to assess the complexity of the tracing, with 1 being “not difficult at all” and 7 representing “extremely tough.”

2.3. Data Collection. The draw over assignment and the single channel-EEG data acquisition were managed by two independent computers. Before each experiment, the clocks were synchronised. A 32" liquid crystal display (LCD) display was used to show the sketching activity to the subjects, which was situated 30" from their foreheads.

The MATLAB programme accumulated behavioural data such as timeframe (the time from the very first mouse pointer motion until the last observed motion), completion rate (the percentage of draw over activity which was accomplished by subjects), preciseness (the fraction of draw over activity that the studied participants prepared within the border of the displayed picture), and interpretive rating of task complexity along with the true tracing route. The NeuroSky Mind Wave Mobile headset was used to capture single-channel EEG data from the volunteers' foreheads at a frequency range of 480 Hz.

The Mind Wave Mobile headset is an EEG-monitoring device that safely detects brainwave patterns and tracks individuals' attention levels as they interact with various apps. It detects brain signals, filters out background noise and electrical interference, and transforms to digital power.

During the experiment, EEG data were wirelessly sent and recorded on the data gathering compute (Figure 3).

2.4. Practice and Evaluation of EEG Single. Real-time EEG data were captured at 512 Hz, and the frequencies within 0.5 and 45 Hz were recovered. The dataset was then subjected to

a wavelet--based filter to eliminate eye blink and movement artefacts.

The eye-blinks and movement of the eye balls are a common difficulty encountered during the clinical recording of electroencephalogram (EEG) signals. Changes in the electric fields are around the eyes, and hence over the scalp, are caused by eye blinks. As a result, EEG recordings are frequently highly distorted, making interpretation difficult. Here, we applied a stationary wavelet transform (SWT) to the corrupted EEG to correct for the presence of the ocular artefact (OA) or movement artefacts (MA) [37].

The commencement and departure times of mouse pointer motion were used to fragment continuous EEG into epochs. For all preprocessed EEG segments, short-time Fourier transformations with 50 percent overlapping 2s hamming windows were calculated. Individual individuals' average power spectra were then calculated across segments for each trial they completed [35, 38].

The hamming window is a reduction created by a raised cosine with nonzero endpoints that is tailored to minimise the side lobe closest to it. It was suggested for smoothing the time domain truncated auto covariance function. The hamming window is most commonly seen in signal processing literature, where it is one of many windowing methods for smoothing values. It is also called as a tapering function or apodization (that indicates “removing the foot,” i.e., smoothing irregularities at the start and finish of the sampled signal).

The frequency band is depicted in Table 2.

3. Results

We have included 10 healthy control participants (7 male, 3 female) to match the remaining 10 depressive patients (7 male, 3 female) so that there was no considerable difference in age or sex between the two groups. All participants were passed through a sequences of investigation tests that measured perspicacity, ability of the visual system to distinguish objects from the background, and intellectual power to see whether there were any additional characteristics that could affect task performance. The Freiburg visual acuity and contrast test (FrACT) was used to assess the participants' visual acuity and contrast sensitivity at a distance of 200 cm [39]. The Sloan Letters have ten options, the Tumbling *E* has four, and the Landolt-C has four or eight. In a nutshell, 4 directions are less likely to be confused; however, the predicting probability is larger, necessitating more trials; 8 places allow for a faster determination of visual

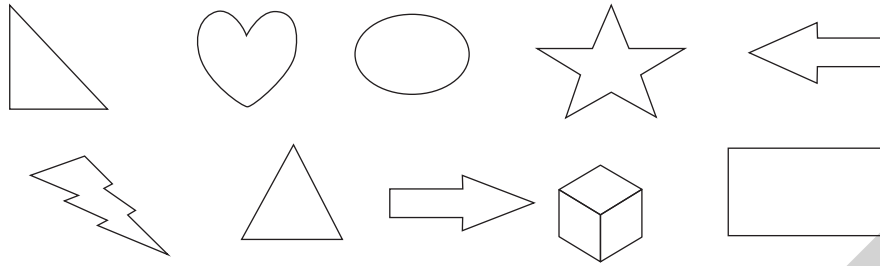


FIGURE 2: Ten shapes depicted in the draw over activity. Test 1–5 (first line-from left to right) and second line 6–10 (from left to right).

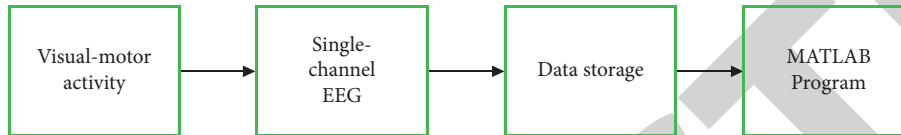


FIGURE 3: Data collection workflow.

TABLE 2: Frequency band.

Type of band	Hz
Delta	Up to 4
Theta	4–8
Alpha	(i) Lower alpha 8–11 (ii) Upper alpha 11–14
Beta	(i) Lower beta 15–30 (ii) Upper alpha 31–36
Gamma	(i) Lower gamma 37–45 (ii) Upper gamma 46–50

acuity (fewer trials) due to the lower predicting probability. The final result should be identical, but the number of runs must be carefully selected. Wilkinson [40] measured pre-morbid IQ, and researchers used the reading subtest of Othe wide range achievement test (WRAT). Exclusion criteria included incapability to demonstrate normal or corrected-to-normal vision (visual acuity score >1.0), inability to rectify low-contrast stimulation (2 percent contrast or less), as measured by a contrast sensitivity threshold score of >50, and a WRAT percentile score of less than 50%. Subjects who did not fit the above-mentioned diagnostic criteria were also eliminated from the study (Table 3).

The average rating for task completion difficulty rating for depression patients’ and in control group is shown in Figure 4. And the trial completion time is shown in Figure 5. Trials 4,6, and 9 were the most challenging job for most participants, whereas Trials 3,7, and 10 were the easiest, and others were considering under moderate category. The normalised perceived difficulty rating had 0.512, 0.623, and 0.821 correlations with the length of the pathway, the integer of inclination in the pathway, and the total duration to complete every test observed, respectively ($p < 0.001$), among depression patients. While the rating was 0.182, 0.236, and 0.382 that were reported among control group, and there was no statistical significance association reported between rating scale and time spent on each trial ($p = 1.43$). The number of severe changes in path direction resembles to the integer of curves for the pathway reported by patients’ group. The normalised perceived complexity and the mean

frequency with which each participant finished each trial had no clear linear relationship. Tracing accuracy has also been proven to be independent to operating velocity and period in recent testing.

It can be seen that Figure 5 represents time spent on each trial by the studied participants. Overall, maximum time taken by patients’ group is as follows: highest time was taken on task 4 (260 seconds), 6 (280 seconds), and 9 (290 seconds). In comparison with their counterparts (participants from the control group), for trials 4, 6, and 9, less time taken, 180 seconds, 190 seconds, and 200 seconds, respectively. X-axis denotes trails. Y-axis denotes task completion time in seconds in Figure 5. For both groups, Table 4 shows the relationship between average band power and each trial observation. The mean EEG power in the alpha-rhythm of upper cutoff, alpha- rhythm of lower cutoff, and beta-rhythm of lower bands of the complete experiment indicated a significantly reciprocal correlation with individual participants’ experienced severity in the depressed patients’ group, but no such relationship was seen in the power in other EEG bands. Drawing precision was completely irrelevant to EEG spectrum strengths at the frontal site throughout the study. Additionally, there was a considerable relationship between the upper and lower alpha band power ($p < 0.01$) and upper beta band power ($p < 0.05$) and a strong correlation between task performance accuracy and perceived problems.

The link among data gathered from distinct time-frame sections and total observed (incertitude) complexity was analysed, as indicated in Table 5, despite the fact that the power spectrum of the EEG changed throughout time intervals. While looking into the depressive disorder patients’ group, the first 30 minutes in cooperation of theta-rhythm in frontal and alpha power rhythm showed significant associations with incertitude complexity of every task, but their relevance faded over time, especially theta power. While in the control group, no significant association in studied EEG characteristics was observed in the time frame. Only the first 30 seconds of each trial’s EEG activity are indicative of the subjects’ perceived difficulty during the visual-motor task;

TABLE 3: Demographic data for healthy control and depressive disorder groups.

	Depressive disorder patients	Control group	<i>p</i> value
Sex	Male -7, Female -3	Male -7, Female -3	1.00
Age of participants (years)	27.3 ± 2.8	28.3 ± 3.2	0.25
Perspicacity (twenty-twenty vision)	1.62 (0.51)	1.31 (0.36)	0.60
Visual perception	162.5 (56.7)	170.5 (47.6)	0.75
Wide range achievement test (WRAT) percentile	88.7 (12.1)	82.4 (11.7)	0.82
Beck depression inventory score	35.6 ± 3.7	18.5 ± 1.8	0.001

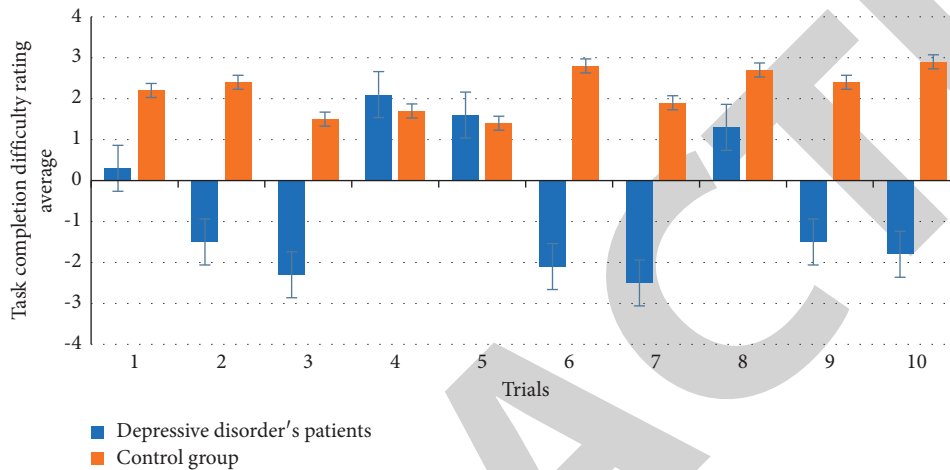


FIGURE 4: Average task completion difficulty ratings.

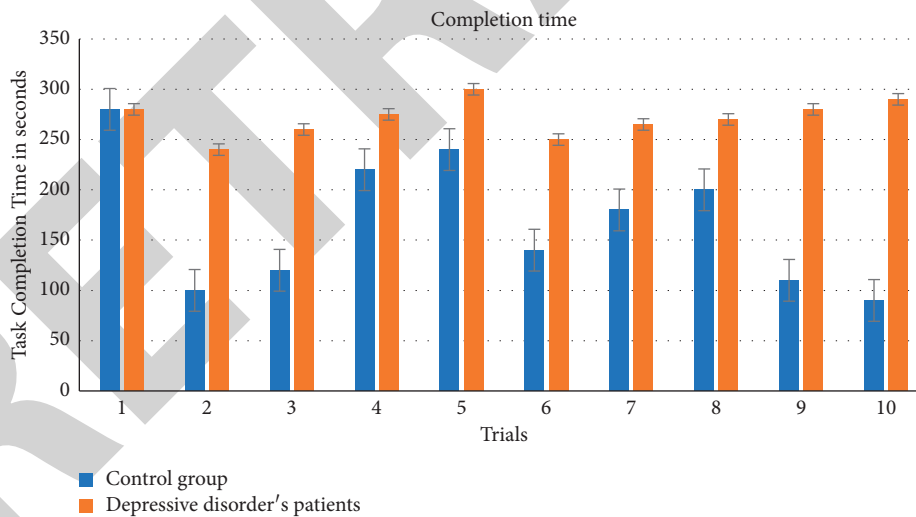


FIGURE 5: Average time taken to complete each trial.

data obtained after that are likely due to other spatiotemporal variables introduced to the individuals or created through them.

Figure 6 shows the fraction modifications in mean power spectrum density compared to total average activity in the initial 30 seconds of the further challenging tests. The fourth test has greatest complexity. Test 5 was easier to complete apart from Tests 6 and 9. Trials 4 and 9 statistically have no effect on the average idiosyncratic difficulty grading for all tests in the experimentations.

4. Discussion

Detecting the variations in brain activity and movement that underpin mental illnesses is critical for better understanding the onset of these disorders and developing effective treatments. As a result, the researchers believe that more research into the brain's altered processing of visual information as a result of depression is required. It might be good to evaluate and expand the usage of perception tests as investigation techniques and prospective methods of

TABLE 4: Associations among mean frequency strengths and outcomes in every experiment test observed among patients' group and control group.

Band	Pearson correlation with mean frequency strengths			
	Idiosyncratic difficulties in depressive disorder patients' group	Idiosyncratic difficulties in control group	Performance accuracy in depressive disorder patients' group	Performance accuracy in control group
Delta	0.12	0.16	-0.0056	0.127
Theta	0.19	0.15	-0.0712	0.238
Lower alpha	0.11*	0.18	0.0348*	0.762
Upper alpha	0.08**	0.12	-0.0176**	0.487
Lower beta	0.05	0.018	0.0659	0.892
Upper beta	0.016*	0.13	0.0521	0.592
Lower gamma	0.025	0.030	0.0439	0.490
Upper gamma	0.07	0.021	0.0287	0.703

* $p < 0.05$; ** $p < 0.01$.

TABLE 5: Associations among participants' idiosyncratic difficulties and EEG characteristics in specific time structure among both groups.

EEG characteristics	Correlation with idiosyncratic difficulties			
	Depressive disorder patients' group		Control group	
	First 30 s	Last 30 s	First 30 s	Last 30 s
Delta	0.134	0.123	0.101	0.074
Theta	0.238*	0.276*	0.123	0.085
Lower alpha	0.324**	0.389**	0.125	0.032
Upper alpha	0.437**	0.463**	0.006	0.043
Lower beta	0.147	0.234	0.076	0.078
Upper beta	0.101	0.143	0.082	0.097
Lower gamma	0.187	0.189	0.043	0.054
Upper gamma	0.054	0.190	0.050	0.043

* $p < 0.05$; ** $p < 0.01$.

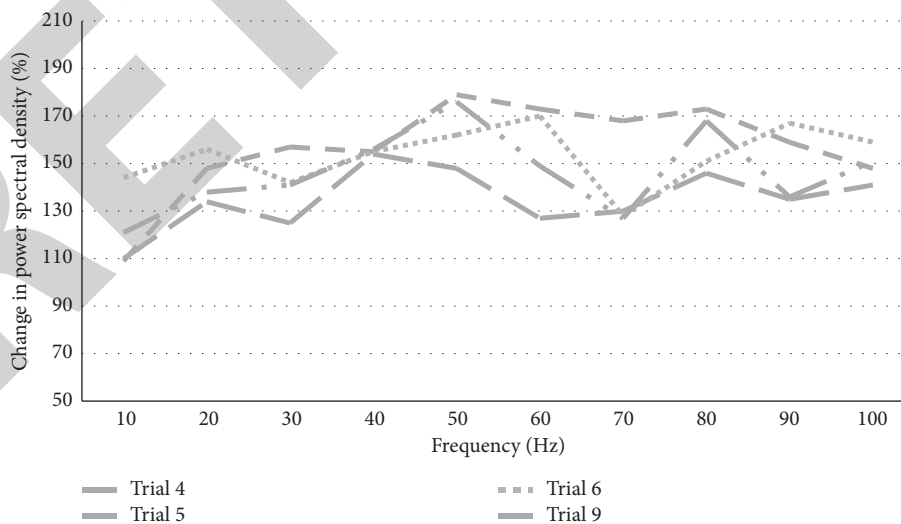


FIGURE 6: Modifications in fraction power spectrum for the most difficult trials seen in the group of depressed patients.

diagnosing electronic data processing abnormalities in sufferer of the mental disorders. Perception tests, for instance, could be used like a supplement to other methods of evaluating the efficacy of various medications as the

treatment continues. However, because the experiential changes remain modest and demonstrated specifically when matching among sets, depression cannot be recognised by assessing visual perception [41].

Taking averages across trials and then across people is standard in EEG data processing, but multidimensional characterisation of independent variability is rarely documented [42]. Feedback averaging presumes that variations between single trials are caused by noisy sums of artefactual and/or task-independent EEG activity that can be simply averaged away to reveal single, fixed (real) brain interaction patterns in both time domain (ERP) and frequency domain (ERSP) evaluate. Examining EEG and other dynamic brain imaging data on a test-by-test basis appears to be an important step toward a deeper comprehension of electrical brain dynamics [43, 44].

In this paper, authors present a novel technique for characterising test-to-test changeability in the interval realm. The results of which recommend that in single-trial data, the activities of a spatially compressed setup of procedures with coarsely corresponding average dynamics demonstrate robust amplitude fluctuations and multiple dynamic approaches. This study discovered a substantial link among the spectral properties of single-channel frontal EEG and the task's mental effort requirements in normal control group and depressive disorder patients. According to the findings, a large increase in mental workload caused a constant escalation in EEG activities in the higher alpha band rhythm (between 11 and 14 Hz) in case of patients suffering from the depression.

Mental emotions can affect the EEG recorded from various parts of the brain. Negative feelings have been connected to brain waves in the right frontal lobe, while happy feelings have been connected to waves in the left frontal lobe [45, 46]. Asymmetry in brain electrical activity could indicate depression or the potential for depression. According to research, the right frontal cortex activity in depressed people is higher than the left frontal cortex activity [47]. Brzezicka et al. (2017) conducted research on controlled and depressed patients who completed a reasoning task and then underwent 5 minutes of electroencephalography recording, with the results showing that the frontal alpha asymmetry index in the depressed group was characterized by greater variance than in the control group [48].

In the first 30 seconds of each trial, the link between frontal EEG activity and mental effort is strong in case of the control group. While depressive participants' EEG features showed significant variation, other temporal events that occurred beyond the first 30 seconds may have influenced this outcome. Other temporal events that occurred beyond the first 30 seconds may have influenced this outcome.

Our findings also demonstrated that the frequency of acute particular movements in the trace, rather than the real time required to complete the line of work, influenced the mental exertion level associated with the activity. There has been a lot of research into the relationship between theta and alpha signals and difficult and effortful activities [12, 25, 39, 40, 49, 50]. Furthermore, during a high-load activity, Gevins and Smith discovered that frontal theta power increased whereas parietal alpha power decreased [25, 40]. According to our findings, both frontal theta and alpha power have positive associations with mental exertion. The alpha power disparity between our results and those of Gevins and Smith is most likely due to the recording locations, Fz and Pz.

From a practical perspective, visual contrast tests are a quick, easy, and noninvasive procedure that could be refined further to serve as biomarkers for aberrant visual information processing in depression. This discovery provided light on the prospect of adopting a user-friendly mobile EEG technology to assess mental workload in everyday life.

5. Conclusion

In conclusion, we have used some of the most regularly used EEG variables to demonstrate the EEG association of cognitive burden in a rather basic motor movement.

EEG asymmetry differences observed in depressive disorder patients' group in theta and alpha band showed significant associations with uncertainty complexity of given tasks and reduced relative left frontal asymmetry exhibit during the oculomotor activity.

More complex encoding and feature extraction techniques are needed to improve the effectiveness of the algorithm. Further study employing longitudinal and cross-modal designs is needed to assess the reliability, accuracy, and responsiveness of this proposed technique for determining diagnostic contrast increment thresholds in depression and other mental diseases such as schizophrenia. Because of the nature and neuronal activation pattern distinctions between motor and cognitive tasks, a distinct set of EEG characteristics may be detected.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] S. Vyas and D. Bhargava, "Big data analytics and cognitive computing in smart health systems," in *Smart Health Systems Smart Health Systems*, pp. 87–100, Springer, Singapore, 2021.
- [2] M. K. Loganathan, B. Mishra, C. M. Tan, T. Kongsvik, and R. N. Rai, "Multi-criteria decision making (MCDM) for the selection of Li-ion batteries used in electric vehicles (EVs)," *Materials Today Proceedings*, vol. 41, pp. 1073–1077, 2021.
- [3] J. O. Berle, E. R. Hauge, K. J. Oedegaard, F. Holsten, and O. B. Fasmer, "Actigraphic registration of motor activity reveals a more structured behavioural pattern in schizophrenia than in major depression," *BMC Research Notes*, vol. 3, no. 1, p. 149, 2010.
- [4] M. Faurholt-Jepsen, S. Brage, M. Vinberg et al., "Differences in psychomotor activity in patients suffering from unipolar and bipolar affective disorder in the remitted or mild/moderate depressive state," *Journal of Affective Disorders*, vol. 141, no. 2–3, pp. 457–463, 2012.
- [5] D. Todder, S. Caliskan, and B. T. Baune, "Longitudinal changes of day-time and night-time gross motor activity in clinical responders and non-responders of major depression," *World Journal of Biological Psychiatry*, vol. 10, no. 4, pp. 276–284, 2009.

- [6] S. Chopra, G. Dhiman, A. Sharma, M. Shabaz, P. Shukla, and M. Arora, "Taxonomy of adaptive neuro-fuzzy inference system in modern engineering sciences," in *Computational Intelligence and Neuroscience*, S. H. Ahmed, Ed., vol. 2021, pp. 1–14, Hindawi Limited, 2021.
- [7] A. J. Gelenberg, M. P. Freeman, J. C. Markowitz et al., "American Psychiatric Association practice guidelines for the treatment of patients with major depressive disorder," *American Journal of Psychiatry*, vol. 167, no. Suppl 10, pp. 9–118, 2010.
- [8] The American Psychiatric Association (Apa), "What is posttraumatic stress disorder?," 2020, <https://www.psychiatry.org/patients-families/ptsd/what-is-ptsd>.
- [9] The American Psychiatric Association (APA), "What Are Dissociative Disorders?," 2018, <https://www.psychiatry.org/patients-families/dissociative-disorders/what-are-dissociative-disorders>.
- [10] M. Sehgal and D. Bhargava, "Knowledge mining: an approach using comparison of data cleansing tools," *Journal of Information and Optimization Sciences*, vol. 39, no. 1, pp. 337–343, 2018.
- [11] S. Saralch, V. Jagota, D. Pathak, and V. Singh, "Response surface methodology based analysis of the impact of nanoclay addition on the wear resistance of polypropylene," *The European Physical Journal - Applied Physics*, vol. 86, no. 1, p. 10401, 2019.
- [12] T. J. Trull and U. Ebner-Priemer, "Ambulatory assessment," *Annual Review of Clinical Psychology*, vol. 9, no. 1, pp. 151–176, 2013.
- [13] S. A. Prince, K. B. Adamo, M. Hamel, J. Hardt, S. Connor Gorber, and M. Tremblay, "A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 5, no. 1, p. 56, 2008.
- [14] J. Bhola and S. Soni, "A study on research issues and challenges in WSN," in *International Conference on Wireless Communications, Signal Processing and Networking (WiSP-NET)*, pp. 1667–1671, IEEE, 2016.
- [15] S. Tang and M. Shabaz, "A new face image recognition algorithm based on cerebellum-basal ganglia mechanism," *Journal of Healthcare Engineering*, 2021.
- [16] X. Huang, V. Jagota, E. Espinoza-Muñoz, and J. Flores-Albornoz, "Tourist hot spots prediction model based on optimized neural network algorithm," *International Journal of System Assurance Engineering and Management*, 2021.
- [17] M. K. Loganathan, P. Goswami, and B. Bhagawati, "Failure evaluation and analysis of mechatronics-based production systems during design stage using structural modeling," in *Applied Mechanics and Materials*, vol. 852, pp. 799–805, Trans Tech Publications Ltd, 2016.
- [18] R. A. Dienstbier, "Arousal and physiological toughness: implications for mental and physical health," *Psychological Review*, vol. 96, no. 1, pp. 84–100, 1989.
- [19] M. Li, V. Rozgic, G. Thattai et al., "Multimodal physical activity recognition by fusing temporal and cepstral information," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 369–380, 2010.
- [20] A. Pantelopoulos and N. G. Bourbakis, "A survey on wearable sensor-based systems for health monitoring and prognosis," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 40, no. 1, pp. 1–12, 2009.
- [21] L. D. Liao, C. Y. Chen, I. J. Wang et al., "Gaming control using a wearable and wireless EEG-based brain-computer interface device with novel dry foam-based sensors," *Journal of Neuroengineering and Rehabilitation*, vol. 9, no. 1, pp. 5–12, 2012.
- [22] A. H. Jabbar, S. O. Mezan, S. M. A. Absi, and M. A. Agam, "Assessment of anticholinesterase effect of polyvinylpyrrolidone/silver nanocomposite biosynthesized by Pandanus atrocarpus extract," *Materials Today Proceedings*, vol. 42, pp. 2578–2583, 2021.
- [23] A. Kumar, V. Jagota, R. Q. Shawl et al., "Wire EDM process parameter optimization for D2 steel," *Materials Today Proceedings*, vol. 37, pp. 2478–2482, 2021.
- [24] A. Luo and T. J. Sullivan, "A user-friendly SSVEP-based brain-computer interface using a time-domain classifier," *Journal of Neural Engineering*, vol. 7, no. 2, Article ID 026010, 2010.
- [25] S. L. Naismith, W. A. Longley, E. M. Scott, and I. B. Hickie, "Disability in major depression related to self-rated and objectively-measured cognitive deficits: a preliminary study," *BMC Psychiatry*, vol. 7, no. 1, pp. 32–37, 2007.
- [26] J. W. Murrough, B. Iacoviello, A. Neumeister, D. S. Charney, and D. V. Iosifescu, "Cognitive dysfunction in depression: neurocircuitry and new therapeutic strategies," *Neurobiology of Learning and Memory*, vol. 96, no. 4, pp. 553–563, 2011.
- [27] M. P. Walker, "Cognitive consequences of sleep and sleep loss," *Sleep Medicine*, vol. 9, no. Suppl 1, pp. S29–S34, 2008.
- [28] N. Goel, H. Rao, J. S. Durmer, and D. F. Dinges, "September) Neurocognitive consequences of sleep deprivation," in *Seminars in Neurology*, vol. 29, no. No. 04, pp. 320–339, © Thieme Medical Publishers, 2009.
- [29] J. Bhola and S. Soni, "Information theory-based defense mechanism against DDOS attacks for WSN," in *Lecture Notes in Electrical Engineering Advances in VLSI, Communication, and Signal Processing*, pp. 667–678, Springer, Singapore, 2021.
- [30] A. Sharma and R. Kumar, "Computation of the reliable and quickest data path for healthcare services by using service-level agreements and energy constraints," *Arabian Journal for Science and Engineering*, Springer Science & Business Media BV, vol. 44, no. 11, , 2019.
- [31] B. M. Byrne, S. M. Stewart, and P. W. H. Lee, "Validating the Beck depression inventory-II for Hong Kong community adolescents," *International Journal of Testing*, vol. 4, no. 3, pp. 199–216, 2004.
- [32] J. C. Cole, I. Grossman, C. Prilliman, and E. Hunsaker, "Multimethod validation of the Beck depression inventory-II and grossman-cole depression inventory with an inpatient sample," *Psychological Reports*, vol. 93, no. 3_suppl, pp. 1115–1129, 2003.
- [33] A. T. Beck, R. A. Steer, and M. G. Carbin, "Psychometric properties of the Beck depression inventory: twenty-five years of evaluation," *Clinical Psychology Review*, vol. 8, no. 1, pp. 77–100, 1988.
- [34] B. Milner, "Memory and the medial temporal regions of the brain," *Biology of memory*, vol. 23, pp. 31–59, 1970.
- [35] J. D. E. Gabrieli, S. Corkin, S. F. Mickel, and J. H. Growdon, "Intact acquisition and long-term retention of mirror-tracing skill in Alzheimer's disease and in global amnesia," *Behavioral Neuroscience*, vol. 107, no. 6, pp. 899–910, 1993.
- [36] M. Kleiner, D. Brainard, and D. Pelli, "What's new in Psychtoolbox-3?" *Perception*, vol. 36, pp. 1.1–16, 2007.
- [37] T. Zikov, S. Bibian, G. A. Dumont, M. Huzmezan, and C. R. Ries, "October). A wavelet based de-noising technique for ocular artifact correction of the electroencephalogram," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*

Retraction

Retracted: A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment

Computational Intelligence and Neuroscience

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 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] J. Cen and Y. Li, "A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1354233, 8 pages, 2022.

Research Article

A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment

Junjie Cen ¹ and Yongbo Li ²

¹College of Computer Science and Technology, Henan Institute of Technology, Xinxiang, Henan 453002, China

²College of Computer and Information Engineering, Henan Normal University, Xinxiang, Henan 453002, China

Correspondence should be addressed to Junjie Cen; cen@hait.edu.cn

Received 13 January 2022; Revised 22 February 2022; Accepted 26 February 2022; Published 28 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Junjie Cen and Yongbo Li. 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.

Aiming at the lack of feature extraction ability of rumor detection methods based on the deep learning model, this study proposes a rumor detection method based on deep learning in social network big data environment. Firstly, the scheme of combining API interface and third-party crawler program is adopted to obtain Weibo rumor information from the Weibo “false Weibo information” public page, so as to obtain the Weibo dataset containing rumor information and nonrumor information. Secondly, the distributed word vector is used to encode text words, and the hierarchical Softmax and negative sampling are used to improve the training efficiency. Finally, a classification and detection model based on the combination of semantic features and statistical features is constructed, the memory function of Multi-BiLSTM is used to explore the dependency between data, and the statistical features are combined with semantic features to expand the feature space in rumor detection and describe the distribution of data in the feature space to a greater extent. Experiments show that when the word vector dimension is 300, compared with the compared literature, the accuracy of the proposed method is improved by 4.232% and 1.478%, respectively, and the F1 value of the proposed method is improved by 5.011% and 1.795%, respectively. The proposed method can better extract data features and has better rumor detection ability.

1. Introduction

With the development of time and the progress of science and technology, Internet technology has gradually entered thousands of households. As one of Web2.0 (such as Weibo, twitter, wechat, and other platforms), social networks have also developed at an amazing speed [1–3]. Among them, Weibo has quickly won people’s favor with its fast and convenient content form and information interaction, which can be widely spread in the society, and has a profound impact on people’s living conditions, daily travel, and even all aspects [4–7].

The amount of information on the Internet is very complex, and every user can access an astonishing number of articles or posts through the Internet. At the same time, due to the low threshold for the use of social networks and the lack of an effective way of audit and supervision, it

provides convenience for criminals. Criminals can use the characteristics of human nature to spread false information and create fear to achieve their own purposes [8–10]. In addition, ordinary people may also publish their hearsay and shadowy information to social media for verification, which will lead to the generation and spread of rumors [11–14]. The widely spread rumors are usually very confusing. For the general public, due to the lack of relevant knowledge, they often can only judge the information obtained by relying on their own feelings or past experience and cannot effectively judge the information, which may inadvertently aggravate the spread of rumors. Moreover, there are a large number of unidentified and even contradictory information on the network, which also makes users more prone to anxiety and confusion. With the help of online social media, rumors will spread rapidly [15–18]. These spreading rumors will greatly mislead the people become a potential unstable factor in

society and affect the stable operation of economy and society.

The automatic detection of network rumors refers to the comprehensive use of Internet, statistics, machine learning, communication, psychology, and other knowledge to let the machine detect rumors. Compared with the time-consuming and laborious characteristics of manual methods, the automatic detection of rumors can detect network hot events in large quantities in real time, so as to effectively deal with the massive information appearing on the Internet every day [19–21]. In recent years, with the development of big data and artificial intelligence technology, rumor detection ability has also been greatly improved. Most research studies on rumors' detection in social networks are based on deep learning methods [22]. At present, the accuracy of automatic rumor detection cannot reach the manual standard, but this is an important breakthrough in purifying cyberspace.

In order to detect rumors in social networks as soon as possible, Chen et al. [23] proposed an unsupervised learning model combining recurrent neural network and variational autoencoder to learn the network behavior of social network users. By using the significant difference between normal data and abnormal data in the dimensionality reduction process, the error between the output and the target value is compared with the specified threshold to judge whether it is a rumor. Guo et al. [24] proposed a rumor detection model based on hierarchical neural network (HSA-BLSTM) combined with social information. Firstly, a hierarchical two-way long short-term memory model representing learning is established, and then, the weight of the vector is adjusted through the attention mechanism. Wang et al. [25] use multihead self-attention mechanism to detect rumors, extract features from context information, and obtain text local features by CNN. The application of attention mechanism to rumor detection improves the ability of rumor detection. According to the account information of rumor mongering users published in the Weibo rumor refutation, Sun et al. [26] use Weibo API to collect the information of rumor mongering users and use the Weibo information of rumor mongering users, user friend information, and Weibo information published by user friends for Weibo rumor detection. Yongheng Chen et al. [27] proposed a rumor early detection multitask learning (RL-MT-RED) based on reinforcement learning, which describes the closely related rumor detection and position classification problems as a multitask learning and joint learning. RL-MT-RED integrates reinforcement learning to control multitask learning and realizes the dynamic setting of trusted checkpoints. Li et al. [28] proposed a rumor tracking integration model based on deep reinforcement learning (RL-ERT), which aggregates multiple components through the weight adjustment strategy network and uses specific social characteristics to improve the performance of the model. Lin et al. [29] proposed a deep sequence context model for Weibo rumor detection, which considers two important factors of rumor: falsity and influence. Firstly, in order to learn falseness, the hypothesis of word independence is cancelled, and the long-term and short-term memory units are used to capture the two-way sequence context information in the

content. Secondly, in order to understand the impact, the deep-seated contextual information is combined with social characteristics to understand the relationship between content and social characteristics. Although these methods have achieved certain results, most of the data feature extraction is not sufficient, which belongs to the shallow extraction of rumor recognition elements, so it cannot further improve the accuracy.

Based on the above analysis, the current research mainly manually establishes semantic features from the text content of rumors, which has the problem of insufficient feature extraction ability. Therefore, a rumor detection method based on deep learning in the social network big data environment is proposed. The main improvements of this study are as follows: using distributed word vector to encode text words, which improves the operation efficiency in the case of a large number of corpus. In the optimization training stage, hierarchical Softmax and negative sampling are used to improve the training efficiency. Specifically, this study constructs a multi-BiLSTM network and statistical feature fusion model to deeply extract text features and improve the classification efficiency. The attention mechanism is introduced to adjust the weight of the vector, which improves the efficiency and accuracy of task processing. Experiments show that the proposed method can better extract data features and has better rumor detection ability.

2. Rumor Detection Process Based on Deep Learning

At present, rumor detection is mainly based on traditional machine learning. It usually establishes features manually from the three aspects of rumor text content, communication structure, and credibility. The manually established features are single and the process of establishing features is time-consuming. To solve these problems, this study proposes a Weibo rumor detection model based on deep learning, which automatically extracts the characteristics of rumor by training neural network through rumor Weibo data, so as to realize the detection of Weibo rumor. The overall method flow is shown in Figure 1. The idea of this study is to obtain rumor and nonrumor Weibo data from Weibo platform, establish corpus, filter spam comments, and carry out word segmentation, word vectorization, and other operations. Secondly, the text vector is input into the neural network. In order to get better experimental results, a rumor detection model based on multi-BiLSTM network is proposed in this study. Finally, a comparative experiment is set up to verify the effectiveness of the proposed method in rumor detection.

3. Weibo Data Collection and Text Word Vector Training

3.1. Construction of Weibo Corpus. To use the massive data of Weibo to study rumors, it is necessary to do a good job in data collection and cleaning. Using the combination of API and third-party crawler program, the Weibo dataset contains rumor information and nonrumor information.

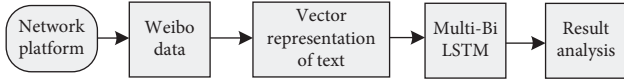


FIGURE 1: Flowchart of rumor detection based on deep learning.

The Weibo open platform provides more than 200 API interfaces, which can download Weibo content, comments, likes, forwards, and other related information. Users can use OAuth2.0 to collect relevant information and build a local database after obtaining authorization. The authorization process of OAuth protocol is shown in Figure 2.

The web crawler program can automatically collect web page information according to the network protocol and crawl from the initial URL. The crawler program uses search engine and web page hyperlink to obtain web page information. This study uses simulated Login technology to obtain deeper information on the page until the crawled data meet the requirements. At present, it is not easy to collect a large amount of Weibo text data, and there are few public text corpora for Weibo. Therefore, when solving the problem of Weibo rumor detection, we usually need to collect our own research corpus. This study selects the scheme of combining API and third-party crawler program to obtain data. The Weibo dataset contains rumor information and nonrumor information.

Sina company directly publishes false Weibo information in the Weibo community management center, so it can directly climb the rumor text on the center page to build a rumor Weibo dataset. This method not only is simple and fast but also avoids the problem of manual annotation and reduces a lot of manpower and time. This study mainly collects 10385 pieces of rumor Weibo information published by the Weibo management center from March 1, 2019, to November 31, 2021. At the same time, it collects the name of the whistleblower Weibo, the link of the whistleblower Weibo, the Weibo name of the whistleblower, and the rumor comments. Then, crawl the normal Weibo, as well as relevant comments and forwarding. Finally, 10385 rumor Weibo and 10759 nonrumor Weibo were collected, respectively, the proportion of training set and test was 9:1, and the relevant comments were 158952.

3.2. Vector Representation of Text. The word vector generated in the form of one hot is too sparse and has too many invalid positions, which will greatly consume the amount of calculation and the physical memory of the computer in practical application. This study uses distributed word vector to encode words. Word2vec mainly uses two models: Continuous Bag of Words Model (CBOW) and Skip-Gram model. Figures 3 and 4 show the basic structures of these two models. The input of the model is the effective vector, and then, it will pass through a projection layer. Finally, the data of the output layer are used for softmax regression to obtain the output vector.

The training goal of the CBOW model is to predict the occurrence probability $p(w_i | w_{i-k}, \dots, w_{i-1}, w_{i+1}, \dots, w_{i+k})$ of the target word according to the context information of the

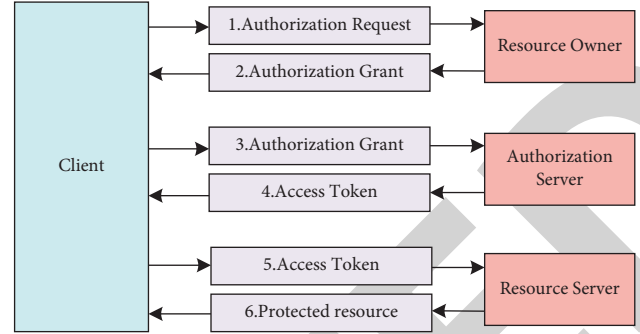


FIGURE 2: Authorization process of OAuth agreement.

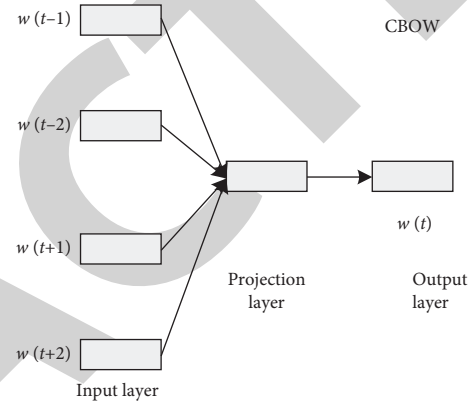


FIGURE 3: CBOW model.

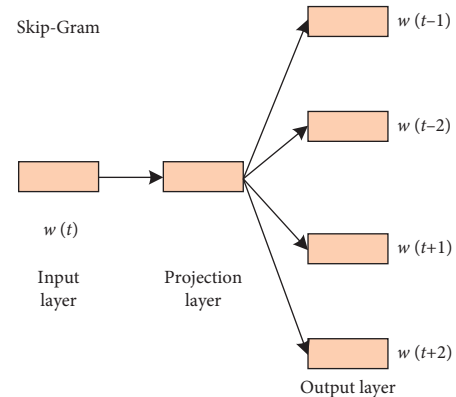


FIGURE 4: Skip-gram model.

target word. The Skip-Gram model, on the contrary to CBOW, uses the target word to predict the context word in its context, that is, the input is the word vector of a specific word, while the output is the context corresponding to a specific word, and the output size is related to the window size set in advance. CBOW is more suitable for small corpora, while the Skip-Gram model performs better in large data corpora.

3.3. Optimization Training. Word2vec implements the optimization algorithm of word vector representation in theory. However, in practical application, the Chinese

corpus is very large, and the number of words reaches millions. In the above two models, the whole vocabulary needs to be calculated, which will cause great difficulties in training and calculation. This study uses hierarchical Softmax and negative sampling to speed up the training speed and improve the training efficiency.

3.3.1. Hierarchical Softmax. The essence of hierarchical softmax is to turn the N -classification problem into $\log(N)$ quadratic classification. In this optimization method, instead of using the traditional DNN network training parameters, Huffman tree is used to replace the neurons in the hidden and output layer. The leaf node of Huffman tree plays the role of output neuron. The leaf node is the size of vocabulary, and the internal node of Huffman tree plays the role of the hidden layer. Taking the CBOW model as an example, the output layer corresponds to a Huffman tree, which takes the words appearing in the corpus as the leaf node and the word frequency as the weight of the node. For any word in dictionary D , there must be a unique path from the root node to the word in the Huffman tree. Each branch on the path can be regarded as a binary classification problem. Each classification produces a probability. Multiplying these probabilities can obtain the target probability. Therefore, the output layer is changed from the original model single layer to the Huffman tree.

3.3.2. Negative Sampling. In the corpus, positive samples are the data samples represented by the current word and its context word, and some context words and current words are randomly generated to form negative samples. Negative sampling is the process of randomly generating context words. When solving the objective function, considering that the central word and its negative sample cannot appear in the training data window at the same time, the probability distribution formula can be understood as the joint probability of the probability that the current word and the negative sample word are not observed in this window. The final optimization objectives are defined as follows:

$$\log p(w_t|c_t) = \log s(w_t, c_t) + \sum_{k=1}^K \log(1 - s(w_k, c_t)), \quad (1)$$

where w_t is the current vocabulary, c_t is the context, and w_k is a random negative sample. Thus, the change from calculating the whole vocabulary to calculating only K negative samples is realized, and the time complexity is reduced.

4. Classification and Detection Model Based on the Combination of Semantic Features and Statistical Features

The traditional word embedding method realizes the shallow extraction of text features, but the task of rumor detection requires the deep extraction of text features, so as to improve the classification efficiency. With the popularity of deep learning in recent years, the use of deep network to extract rumor features has achieved excellent results, among which

the RNN network represented by LSTM is the most common. In this study, multi-BiLSTM network and statistical feature fusion are selected as the network structure of feature extraction. On the one hand, text data have natural serialization characteristics, and LSTM can retain the location information of text. On the other hand, general social media information belongs to typical short text. LSTM uses a gating mechanism to effectively solve the problems of less context information and unclear semantics in short text.

In order to solve the problems of shortage of single microblog content, lack of semantics and inability to integrate global features by simple semantic classification. At the same time, as the current distributed expression of words can bring more semantic information and the training and use of word vectors are becoming more and more standardized, this section proposes based on multifeature fusion. That is, the deep neural network model integrating statistical features and semantic features. The proposed model based on multi-BiLSTM and statistical feature fusion is shown in Figure 5. This study adopts a deep learning model based on the combination of semantic features and statistical features. The first part extracts the semantic features of Weibo. Firstly, the input Weibo data are transformed into time-series data. Multi-BiLSTM has memory function and can find the dependency between data. Therefore, the semantic features that change with time can be learned through the multi-BiLSTM model. This study introduces the attention mechanism to connect different modules and give the weight value to each module in the global scope. The more important the module with high weight value is in feature selection, the lower the value and proportion of the module with low weight value in semantic feature representation. The second part is used to extract the social features based on user information, the social features of communication content and the social features of Weibo content, vectorize the three types of features, splice the vectorized features into a one-dimensional vector, and then map them into a vector with the same dimension as the semantic features through the full connection layer. Finally, the semantic features and statistical features are spliced, and the classification results are obtained through the Softmax layer.

The attention mechanism is introduced into the model because the amount of information contained in the news content is different at different times in the process of news communication. This is related to the communication mode of language. In the early stage of Weibo release, netizens may carry out special communication because of the impossibility or interest of rumors. However, as more and more netizens participate in the discussion of the event, they will doubt the authenticity of the event. At this time, there will be similar "really?" and "impossible?" and other questionable content appears and will also look for official certification. With the development of the event, the news that is finally certified as a rumor will be forwarded with the attitude of popular science. At this time, there will be "this is false" and "deceptive" and other contents. Therefore, in the process of language communication, the importance of message content is different at different time points. For example, in the early stage of forwarding, people only forwarded the interest

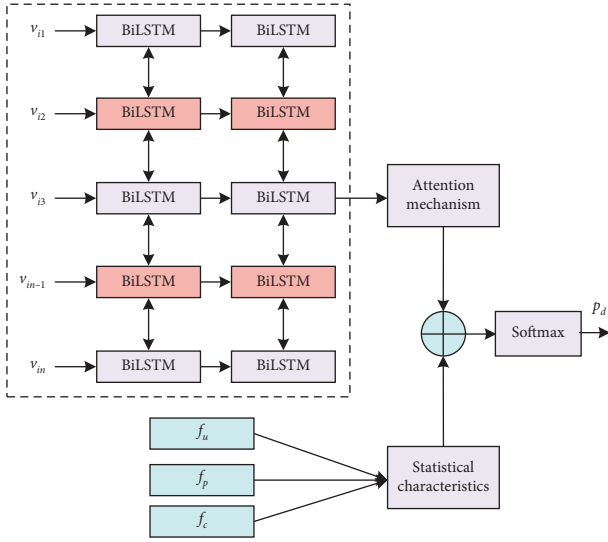


FIGURE 5: Multi-BiLSTM and statistical characteristics fusion model.

of the original Weibo, and there is no questioning attitude in the rumor characteristics. Therefore, the message weight of this part can be appropriately reduced. In the middle or late stage of new use communication, the emergence of rumor pattern words brings more semantic information to the forwarding comment content. At this time, the message content can be given a higher weight.

In the model of multi-BiLSTM combined with statistical features, the input of multi-BiLSTM is still represented by paragraph vector divided by time. On the one hand, if each Weibo is input into the model in the form of vector, it is difficult for the model to learn dependent information. On the other hand, the complexity of the model brings difficulties to training and calculation. Therefore, it is necessary to simplify the steps of the model. At the same time, in order to get more semantic information, events are divided into subevents. The length of the model is equal to the number of sub events, and the input of the model is the vector expression of subevents or paragraphs. For a given event Q , it is divided into N subevents; then, the event representation is $Q = (Q_1, Q_2, \dots, Q_N)$, and the subevent Q_i is a one-dimensional vector with the length of word vector multiplied by the number of words and then multiplied by the number of messages. Each subevent Q_i is used as the input of multi-BiLSTM to obtain the output vector y_i . After a Weibo event is all input into the network, the output sequence $y_i = \{y_1, y_2, \dots, y_N\}$ is obtained.

Considering the time dependence of LSTM, an attention mechanism is introduced into the model. Taking the output of the hidden layer as the input of the attention layer, the formula of the attention layer is as follows:

$$e_i = \tanh(W_w y_i + b_w), \quad (2)$$

where W_w and b_w represent the weight matrix and bias, respectively.

The output vector y_i of the attention mechanism is passed through a single-layer perceptron to obtain the

intermediate vector e_i of the hidden layer state, and then, the paragraph level vector v_w is introduced to determine the weight α_i of the output vector y_i by calculating the similar values of the intermediate vectors e_i and v_w of the hidden layer state. The weight α_i is implemented by a softmax function:

$$\alpha_i = \frac{\exp(e_i^T v_w)}{\sum_t \exp(e_t^T v_w)}. \quad (3)$$

The output sequence is given different weights, and the semantic feature m_E of the last event realized by weighted summation is as follows:

$$m_E = \sum_t \alpha_t y_t. \quad (4)$$

When extracting statistical features, in order to obtain the same amount of information as semantic features, the statistical features are mapped into feature vectors with the same dimension as the output m_E through the full connection layer. The formula is as follows:

$$s = f(W_s \cdot f_s + b_s), \quad (5)$$

$$f_s = f_u \oplus f_p \oplus f_c, \quad (6)$$

where \oplus indicates series operation, f_u , f_p , and f_c , respectively, represent social features based on user information, social features based on communication content, and social features based on Weibo content.

Finally, the mapped statistical feature vector s and semantic feature vector are connected in series, and the two classification results are made through the Softmax layer. The prediction results are as follows:

$$p_d = \text{Softmax}\{m_E \oplus s\}. \quad (7)$$

The cross-entropy loss function is adopted in this model:

$$L = \sum_{d \in D} [y_d \log p_d + (1 - y_d) \log (1 - p_d)], \quad (8)$$

where d is the sample, D represents the sample dataset, y_h represents the true value of the sample, and p_d is the predicted value of the sample. In the two classification results, 0 represents nonrumor and 1 represents rumor.

5. Experiment and Analysis

5.1. Simulation Environment Setting and Evaluation Index. The experiment environment adopts Linux system, and the environment configuration is shown in Table 1.

This study uses accuracy, precision, recall, and F1 score to evaluate the performance of the algorithm. The accuracy indicates the number of correctly classified samples, but when the positive and negative classes of the dataset are unbalanced, this index cannot accurately reflect the performance of the model, so it needs to be judged together with other indexes. Recall and precision will affect each other. Usually, one is high and the other is relatively low. In order to deal with this situation, the harmonic parameter F1 score

TABLE 1: Experimental environment configuration.

Items	Description
OS	Linux ubuntu 16.04
GPU	NVIDIA GeForce GTX 1080Ti
RAM	32G 1867MHZ DDR3

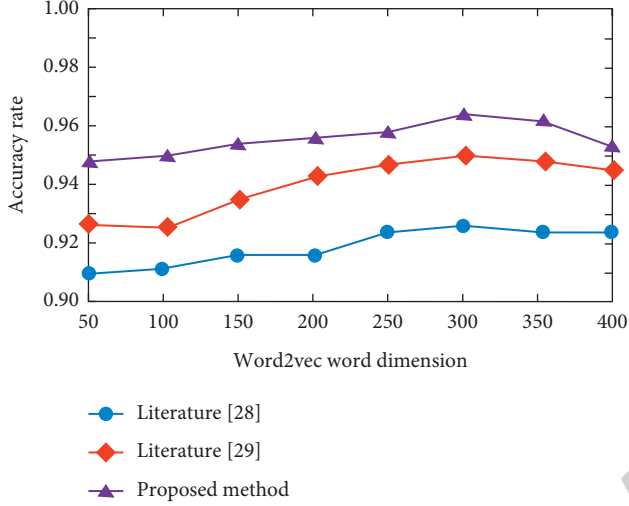


FIGURE 6: Comparison of accuracy of each model under Word2vec word vector.

of recall and precision is introduced. The calculation method of these four indicators is shown in formulas (9) to (12):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (12)$$

For the rumor detection task, TP indicates that the tag is rumor and the classification result is rumor, TN indicates that the tag is nonrumor and the classification result is nonrumor, FP indicates that the tag is nonrumor, but the classification result is rumor, and FN indicates that the tag is rumor, but the classification result is nonrumor.

5.2. Comparison of Bid Evaluation Index Results of Each Model. In order to verify the effectiveness of the model, the proposed method is compared with the methods in [28, 29]. The original Weibo content and their comments were used for training, and the dataset was divided according to the ratio of 9:1 as training data and test data, respectively. The variation of accuracy, precision, recall, and F1 value with word2vec word vector is shown in Figure 6–9.

It can be seen from Figure 6–9 that the proposed model has greatly improved the accuracy, precision, recall, and F1

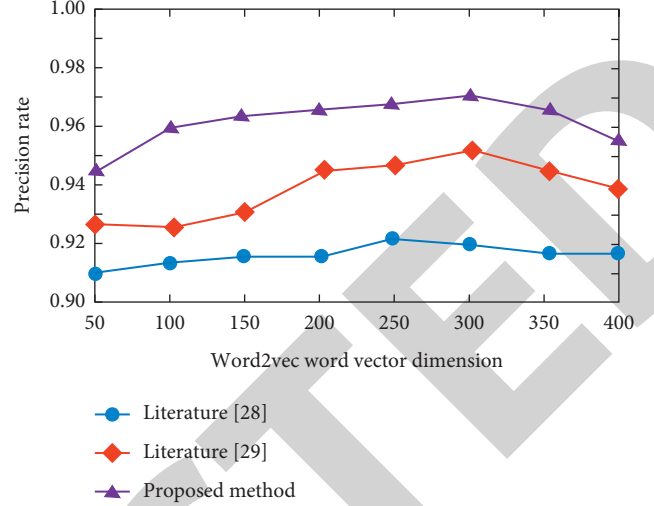


FIGURE 7: Comparison of precision of each model under Word2vec word vector.

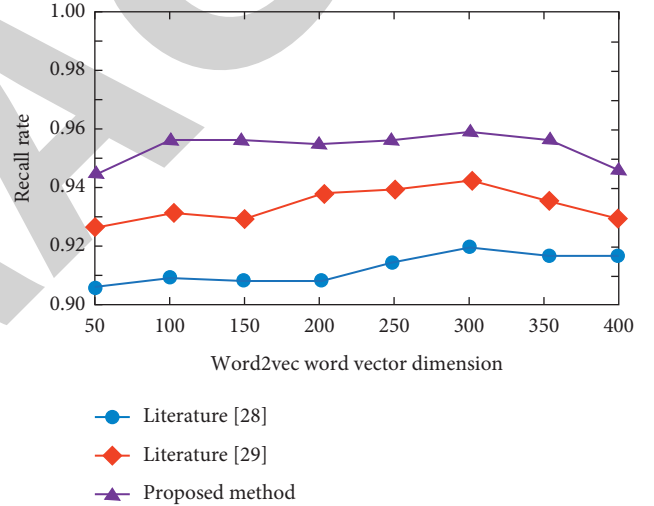


FIGURE 8: Comparison of recall of each model under Word2vec word vector.

value compared with the methods in [28, 29]. When the word vector dimension is 300, the accuracy of the proposed method is 0.962, which is 4.113% and 1.477% higher than 0.924 in [28] and 0.948 in [29], respectively. The precision of the proposed method is 0.971, which is 5.773% and 1.996% higher than 0.918 in [28] and 0.952 in [29], respectively. The recall rate of the proposed method is 0.957, which is increased by 4.135% and 1.592% respectively compared with 0.919 in [28] and 0.942 in [29]. The F1 value of the proposed method is 0.964, which is increased by 5.011% and 1.795%, respectively, compared with 0.918 in [28] and 0.947 in [29]. The proposed method has advantages in the four indicators; this is because the proposed method integrates statistical features and semantic features. As a global feature in rumor detection, statistical features distinguish rumor and non-rumor attributes from the overall situation. However, the statistical feature is only the statistics of attributes, and the

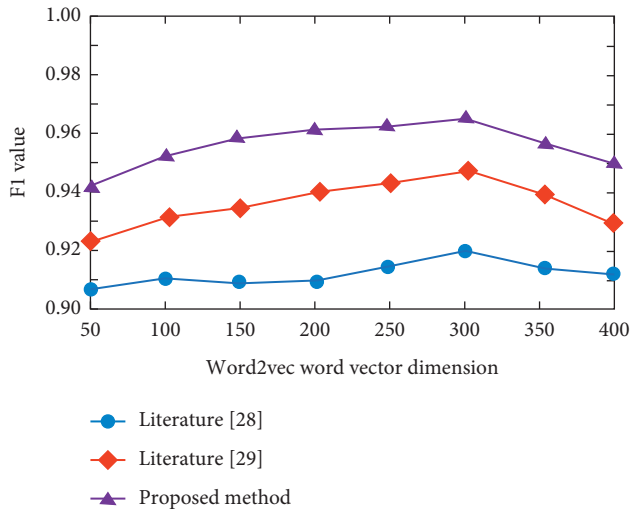


FIGURE 9: Comparison of F1 value of each model under Word2vec word vector.

text semantics cannot be obtained. The text content can only be determined by special symbols or formats. Therefore, the combination of statistical features and semantic features expands the feature space in rumor detection and describes the distribution of data in the feature space to a greater extent, so as to improve the classification performance of the network. While literature [28] method and literature [29] method do not extract text features at a deep level. Based on the four experimental results, it can be seen that the proposed method has reached an advanced level in rumor detection, which shows the effectiveness of the model.

6. Conclusion

Aiming at the lack of feature extraction ability of rumor detection methods based on the deep learning model, this study proposes a rumor detection method based on deep learning in social network big data environment. The distributed word vector is used to encode text words, which improves the operation efficiency of the system in the case of a large number of corpus. In the optimization training stage, hierarchical softmax and negative sampling are used to improve the training efficiency. The multi-BiLSTM network and statistical feature fusion are constructed as the network structure of feature extraction to deeply extract text features and improve the classification efficiency. The attention mechanism is introduced to adjust the weight of the vector, which improves the efficiency and accuracy of task processing.

A problem in the existing research is that most of the existing datasets are for special events in a certain period of time, and the construction of data sets also depends on manual crawling and labeling. The dataset used in the experiment cannot be compared with the amount of information on the huge social network. In the follow-up work, we can try to use transfer learning to improve the generalization ability of the model.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] N. Rani, P. Das, and A. K. Bhardwaj, "Rumor, misinformation among web: a contemporary review of rumor detection techniques during different web waves," *Concurrency and Computation: Practice and Experience*, vol. 12, no. 3, pp. 2012–2021, 2021.
- [2] K. Tu, C. Chen, C. Hou, J. Yuan, J. Li, and X. Yuan, "Rumor2vec: a rumor detection framework with joint text and propagation structure representation learning," *Information Sciences*, vol. 560, no. 1, pp. 137–151, 2021.
- [3] H. Zhang, S. Qian, Q. Fang, and C. Xu, "Multimodal disentangled domain adaption for social media event rumor detection," *IEEE Transactions on Multimedia*, vol. 3, no. 5, pp. 1002–1010, 2020.
- [4] Y. Wang, L. Zheng, and J. Zuo, "Online rumor propagation of social media on NIMBY conflict: temporal patterns, frameworks and rumor-mongers," *Environmental Impact Assessment Review*, vol. 9, no. 1, pp. 106–114, 2021.
- [5] A. Ebrahimi Fard, M. Mohammadi, Y. Chen, and B. Van de Walle, "Computational rumor detection without non-rumor: a one-class classification approach," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 5, pp. 830–846, 2019.
- [6] S. A. Alkhodair, S. H. H. Ding, B. C. M. Fung, and J. Liu, "Detecting breaking news rumors of emerging topics in social media," *Information Processing & Management*, vol. 57, no. 2, pp. 102018–102109, 2020.
- [7] K. Sato, J. Wang, and Z. Cheng, "Credibility evaluation of Twitter-based event detection by a mixing analysis of heterogeneous data," *IEEE Access*, vol. 7, no. 6, pp. 1095–1106, 2018.
- [8] X. Chen, D. Zhu, D. Lin, and D. Cao, "Rumor knowledge embedding based data augmentation for imbalanced rumor detection," *Information Sciences*, vol. 580, no. 5, pp. 352–370, 2021.
- [9] F. Xu, V. S. Sheng, and M. Wang, "Near real-time topic-driven rumor detection in source microblogs," *Knowledge-Based Systems*, vol. 207, no. 6380, pp. 106391–106399, 2020.
- [10] C. Song, C. Yang, H. Chen, C. Tu, Z. Liu, and M. Sun, "CED: credible early detection of social media rumors," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 8, pp. 3035–3047, 2021.
- [11] L. Lin and Z. Chen, "Social rumor detection based on multilayer transformer encoding blocks," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 6, pp. 1–15, 2020.
- [12] T. T. Nguyen, T. T. Nguyen, T. T. Nguyen, B. Vo, J. Jo, and Q. V. H. Nguyen, "JUDO: just-in-time rumour detection in streaming social platforms," *Information Sciences*, vol. 570, no. 1, pp. 70–93, 2021.
- [13] Z. Wu, D. Pi, J. Chen, M. Xie, and J. Cao, "Rumor detection based on propagation graph neural network with attention

Retraction

Retracted: Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Kamal, A. R. Pratap, M. Naved et al., "Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5261942, 8 pages, 2022.

Research Article

Machine Learning and Image Processing Enabled Evolutionary Framework for Brain MRI Analysis for Alzheimer's Disease Detection

Mustafa Kamal ¹, A. Raghuvira Pratap ², Mohd Naved ³, Abu Sarwar Zamani ⁴,
P. Nancy ⁵, Mahyudin Ritonga ⁶, Surendra Kumar Shukla ⁷, and F. Sammy ⁸

¹Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Dammam 32256, Saudi Arabia

²Department of Computer Science and Engineering, V. R. Siddhartha Engineering College, Andhra Pradesh, Vijayawada, India

³Amity International Business School (AIBS), Amity University, Noida, UP, India

⁴Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia

⁵Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, India

⁶Universitas Muhammadiyah Sumatera Barat, Padang, Indonesia

⁷Department of Computer Science & Engineering Graphic Era Deemed to be University, Dehradun, Uttarakhand, India

⁸Department of Information Technology, Dambi Dollo University, Dembi Dolo, Welega, Ethiopia

Correspondence should be addressed to F. Sammy; sammy@dadu.edu.et

Received 2 February 2022; Revised 27 February 2022; Accepted 8 March 2022; Published 28 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Mustafa Kamal 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.

Alzheimer's disease is characterized by the presence of abnormal protein bundles in the brain tissue, but experts are not yet sure what is causing the condition. To find a cure or aversion, researchers need to know more than just that there are protein differences from the usual; they also need to know how these brain nerves form so that a remedy may be discovered. Machine learning is the study of computational approaches for enhancing performance on a specific task through the process of learning. This article presents an Alzheimer's disease detection framework consisting of image denoising of an MRI input data set using an adaptive mean filter, preprocessing using histogram equalization, and feature extraction by Haar wavelet transform. Classification is performed using LS-SVM-RBF, SVM, KNN, and random forest classifier. An adaptive mean filter removes noise from the existing MRI images. Image quality is enhanced by histogram equalization. Experimental results are compared using parameters such as accuracy, sensitivity, specificity, precision, and recall.

1. Introduction

An Alzheimer's patient's memory, cognition, and conduct are all affected by this brain illness. Symptoms normally begin to appear gradually and steadily worsen over time, until they become so severe that they interfere with daily activities. There are many different types of dementia, and Alzheimer's is the most common, but it is also the most

dangerous form of memory loss [1]. It is not a normal part of ageing, despite the fact that growing age is the greatest recognized risk factor. Most of the patients of Alzheimer's disease are older than 65 years [2]. The disease of Alzheimer's is not just confined to the elderly. As Alzheimer's disease progresses over time, so does the disease's severity [3]. It is a sickness that worsens with time since it is degenerative. It is easy to carry on a conversation and react to

your surroundings when memory loss is minimal in its early stages. The creation of biomarkers is necessary for Alzheimer's disease diagnosis [4, 5].

Alzheimer's disease [6, 7] is distinguished by the development of aberrant protein bundles in brain tissue, but experts are unsure of what causes the disorder. Researchers need to know more than simply that there are protein deviations from the norm in order to find a cure or aversion; they also need to know how these brain nerves arise in order to find a solution. Alzheimer's disease progresses for reasons that experts do not fully comprehend. Only a few distinct elements are permitted. We become increasingly vulnerable to Alzheimer's disease as we become older. Many cases of this condition have a family history.

Intelligent behaviour is strongly reliant on one's ability to learn new information [8]. Machine learning is the study of computing approaches for improving performance on a specific task via the learning process [9]. Cognitive, technological, or theoretical [9] goals may be the focus of machine learning research. The goal of cognition is to simulate some aspect of human learning. An example of a technical goal is to automate the acquisition of knowledge for knowledge-based systems. Theoretical analysis takes into account, for example, the scope and constraints of learning processes. Similar to artificial intelligence, machine learning is an interdisciplinary field. In the field of machine learning, for example, statistics are a common tool.

There is an enormous amount of data generated by modern medicine that is maintained in medical records. The physician's interpretation of the results of MRIs, ECGs, blood sugar, blood pressure, cholesterol levels, and other clinical data are examples of medical data. Scientific decision-making for disease diagnosis and therapy is becoming more and more dependent on data mining. This issue can be solved using machine learning in the medical field [10].

According to the physician's medical understanding, the patient's symptoms are assigned to one of several illness groups. The classification models for various neurodegenerative illnesses are hence the learning problem in this research.

This article presents an Alzheimer's disease detection framework consisting of image denoising of an MRI input data set using an adaptive mean filter, preprocessing using histogram equalization, feature extraction by Haar wavelet transform, and classification performed using LS-SVM-RBF, SVM, KNN, and a random forest classifier. The algorithms are compared on the basis of several performance metrics such as accuracy, specificity, sensitivity, and recall.

2. Related Work

2.1. Literature Survey of Preprocessing of MRI Images. Median filtering was proposed [11] to remove noise from photographs, such as salt and pepper and Poisson noise. Because of this, the output intensity of a median filter is determined by sliding an image window along its length, and then calculating its median intensity value by summing the values of all pixels within that window. In addition, the median filter preserves an image's edges while reducing

random noise. All pixels have their values locked to the median value of the pixels closest to them. In order to remove these disturbances, a filter is utilized, and then the bounding box technique is used to locate the tumor.

It was discovered [12] that the evolution of order statistics filters allows for a simple yet effective method of reducing noise in medical images. Median and mean filtering are combined to get the pixel value of an image with no noise, as seen in this example. Additionally, it can be used to reduce picture artefacts like Rician noise.

An anisotropic filter was developed by [13] to eliminate background noise and safeguard the image's edge points. Using this method, both filtering and stitching in real time can be carried out. Diffusion constant selection is dependent on the noise gradient, and it smoothes the signal by removing background noises that are introduced during filtering with the appropriate threshold value.

MRI image enhancement is dependent on the modified tracking algorithm, histogram equalization, and center weighted median (CWM) filter developed by [14]. Two methods are used in this procedure. Utilizing the updated tracking technique to remove film artefacts, labels, and the skull region is the initial step, followed by using the histogram equalization and CWM filter approaches independently to enhance images.

For example, Bayesian denoising bootstraps itself by optimising an information theory metric using the expectation maximization (EM) algorithm in order to estimate priors. The NLM, a parametric filter for random noise elimination, was developed by [15] to remove the random noise present in multicomponent MRI images. The spatially averaged of identical pixels using information from all the image components was used to carry out the denoising process.

2.2. Literature Survey of Feature Extraction Methods. The authors [16] emphasized the need of incorporating not only textual data, but also scan-derived picture visual features and doctor-provided input. Features can have coefficients matching an image spectral transform. Brain pictures can be described using LBPs (local binary patterns) and DCT (direct component transformation). Early detection of Alzheimer's disease can be improved by using visual picture similarity. It demonstrates the accuracy of brain image classification based on user feedback. The photos are then compared using a variety of classification methods.

MRI images were processed to extract both ROI and HOG features, which were then mapped onto the ROI space in order to make them comparable and to confirm the higher similarity value. When it comes to AD diagnosis, a support vector machine is trained on both mapped HOG features as well as actual ROI data. Using ROI features to map HOG characteristics to, we are aiming to provide complementary information so that the features from various perspectives may be not only compared but also understood.

2.3. Literature Survey of Classification Methods. Classifiers based on the Naive Bayes theorem have a rough assumption of feature independence, which makes them

part of the usual family of probabilistic classifiers. It has been shown that the Bayesian network decision model introduced by the researchers [17] outperforms other popular classifiers. Although [18] introduced a multifold Bayesian Kernalization technique that can discriminate AD from NC with improved accuracy, they found unsatisfactory results in the diagnosis of MCI-converter.

A hyperplane generated by an SVM in a high- or indefinite-dimensional space may be used for classification, regression, and other applications. If you have a few training samples, SVMs are generally utilized to address pattern classification issues because of their ability to minimize generalization errors.

A SVM was used [19] as a feature selection criterion and a classifier in MRI data for the diagnosis of Alzheimer's disease and obtained an accuracy of 86% and a specificity of 92%. Schmitter et al. used SVM to examine two different VBM algorithms: the free surfer and an in-house technique. Traditional whole-brain VBM techniques are comparable to MorphoBox in terms of efficacy.

Authors [20] tested two methods for separating older people from those with Alzheimer's disease (AD) and MCI (MRI). Every subject was first filtered and then normalized, and then twelve features were extracted using K-nearest neighbor (KNN) and support vector machine (SVM). In order to choose the best characteristics for accurately identifying classes, two classification methods, permutations and combinations, were used to assess each feature once it had been selected. Using SVM polynomial order three, we were able to achieve an average accuracy of 97.92 percent, and even better, researchers were able to achieve an accuracy of 95.833 percent when using KNN. The categorization accuracy of the three clinical groups was found to be comparable.

To extract the master features of the images, the authors [21] used a rapid discrete wavelet transform (DWT), and then used principal component analysis (PCA) to examine the master features (PCA). Five different decision models each receive a different subset size of the main feature vectors. The models that are included in the classification models include the J48 decision tree and KNN, random forest (RF), and LS-SVM with polynomials and radial basis kernels.

Extraction of picture characteristics was done using a method known as fiber-tract modeling, and the SVM classifier was applied to distinguish AD from NC. The SVM classifier was found to be accurate (86.2% accuracy), sensitive (88.0% accuracy), as well as specific (89% specificity).

Voxel-based morphometric and Fisher criterion were used by [22] for feature selection and reduction over the entire brain, which was then followed by SVM for classification. For independent diagnosis, whole brain approaches have shown a high level of discriminative power.

Authors [23] investigated the multifeature combination correlation technologies and improved the support vector machine recursive feature elimination (SVM-RFE) algorithm using the covariance technique. The effectiveness of the newly developed method is demonstrated by the comparison studies conducted on the accessible ADNI database.

It also indicates that combining many features is superior to using a single feature alone.

3. Methodologies

The Alzheimer's disease detection methodology consists of image denoising of the MRI input data set using an adaptive mean filter, preprocessing using histogram equalization, feature extraction by Haar wavelet transform, and classification is performed using LS-SVM-RBF, SVM, KNN, and a random forest classifier. The block diagram is shown in Figure 1.

To remove the noise that makes an image look better, the adaptive median filter (AMF) algorithms [24] are used a lot. To determine which pixels in an image are impacted by impulse noise, the AMF algorithm performs spatial processing like this. "Impulse noise" is a term used to describe the appearance of a large number of pixels that are not spatially aligned. As a result, noise pixels are hidden from view by using the median value of the pixels in their immediate vicinity that passed the noise labeling test.

The simplest wavelet transform is the Haar wavelet transformation [25]. The Haar transform is a mathematical process that connects Haar wavelets. The Haar transform is a sampling operation that is employed in all wavelet transforms. The Haar transform removes half of a signal's length. Both the first and second examples are running averages, but the first is a running average by comparison.

Histogram equalization usually makes images look more contrast, especially if the important data in the image is shown in close contrast to the color of the background. To make the histogram more even, you can make this change. This lets the parts of the picture that do not have as much contrast get more contrast. Histogram equalization does this by spreading out the intensity values that are most common. This is a good way to make images that have both bright and dark parts. A big advantage of this method is that it is easy to use and cannot be changed. There are two ways in which you can get your original histogram back if you know how to equalize the histogram: the calculation does not require a lot of computing power [26].

A support vector machine (SVM) is a supervised learning model and algorithms that look at the data for classification and regression analysis. Given a set of training examples, the SVM training algorithm creates a model that categorizes new examples into one of two categories, resulting in a nonprobabilistic binary linear classifier. The data points in the SVM model, which represents examples in space, are partitioned into various categories using as wide a gap as possible. To solve linear equations and to find a training model for classification, LS-SVM is an addition to SVM. There are two types of SVMs: one for quadratic equations and the other for linear equations. For less money, you can use an LS-SVM classifier. To utilize LS-SVM, you only need to solve a set of linear equations to understand how it works in comparison to SVM. Few parameters are required for LS-SVM to work. SVM uses the RBF kernel, a radial basis function [27].

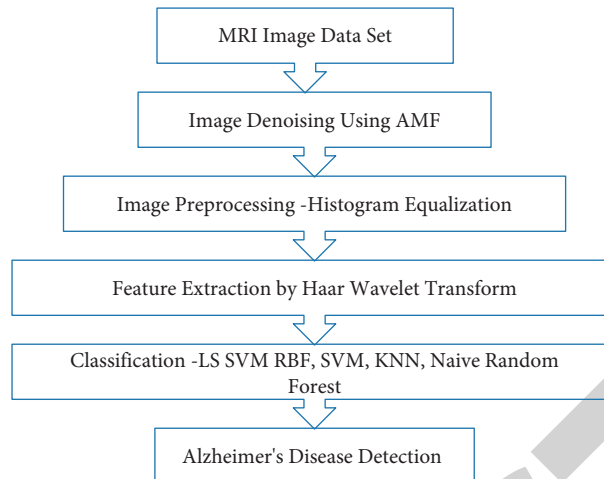


FIGURE 1: A framework for Alzheimer's disease detection.

KNN is one of the most well-known and useful non-parametric classifiers for pattern recognition and machine learning. The kNN algorithm is called the kNN algorithm. This "lazy learning" classifier is still very popular and used in many places because of two main reasons. People use kNN because Bayes' error rate limits how bad it can be at classifying things [9]. Second, the kNN algorithm is very easy to use because it is very simple. The kNN classifier is good because it can compete even if you do not learn how to use it. In the presence of learning, the classifier has a high level of accuracy in both supervised and unsupervised learning frameworks. There are a lot of interesting things about the kNN classifier, but the most interesting thing is that it can be improved by making changes to different parts of the algorithm without changing its main principle. Some of these structural changes may help the classifier be more accurate: the first step is to choose the right distance and similarity measures, along with some strategic changes in how scores are assigned to training classes. This is followed by choosing the dynamic number of nearest neighbors for each test point, and then making sure the algorithm is more robust to the presence of these outliers in the dataset. Finally, multicriteria decision making is used to make the algorithm more flexible and robust to the presence of outliers in the dataset. For a KNN classifier, there are two components. Identifying how far the unknown image differs from each of the images used in the training process is the first step in this process. The second step is to identify which of the practice images is most likely to be a real-world test picture. Objects are classified and their distances are measured using the Euclidean distance. The most popular approach to calculate how far apart two places are from one another is using the Euclidean distance [28].

Classification and regression tasks may both benefit from the use of random forest, which was first proposed by the author [29]. A large number of decision trees are generated during the training phase, and the results of each tree are predicted using regression methods. For prediction purposes, it has a low variance and links the various aspects of the data quickly. People were first unimpressed with random

forest classification since it is difficult to understand. But it has performed better in the prediction task [30].

4. Result and Discussion

In this experimental study, the OASIS [31] data set was used. This data set consists of a total of 416 samples. Image enhancement is machine learning algorithms such as LS-SVM-RBF, SVM, KNN, and random forest are used for classification. Classification results are based on four classes: Alzheimer's disease, Huntington disease, mild Alzheimer's disease, and normal MRI images. Total 100 images are randomly selected, 25 images for each category. Sample images are shown in Figures 2 and 3.

Five parameters such as accuracy, sensitivity, specificity, precision, and recall are used in this study to compare performance of different algorithms.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

where

TP = true positive

TN = true negative

FP = false positive

FN = false negative

The confusion matrix of LS-SVM-RBF, SVM, KNN, and random forest is shown in Table 1.

Figures 4–8 show the accuracy, sensitivity, specificity, precision, and recall of LS-SVM-RBF, SVM, KNN, and random forest for Alzheimer's disease detection. The accuracy of LS-SVM-RBF is higher than that of the other classifiers. The KNN algorithm outperforms the other classifiers in terms of sensitivity and recall. The specificity of

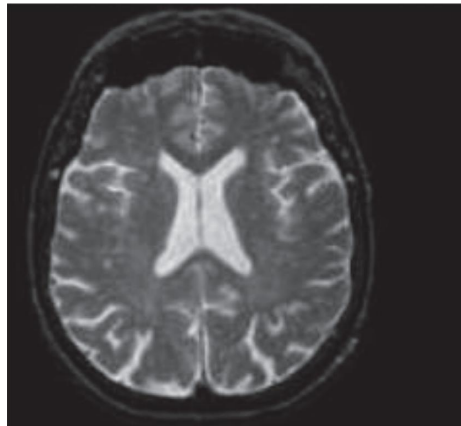


FIGURE 2: Normal brain MRI image.

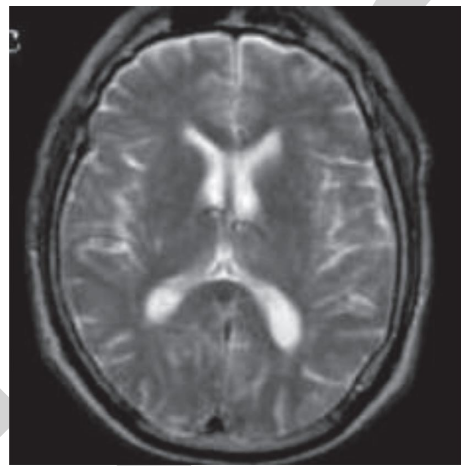


FIGURE 3: Alzheimer's disease MRI Image.

TABLE 1: Confusion matrix of machine learning algorithms.

Parameter	RF	KNN	SVM	LS-SVM-RBF
TP	53	56	56	58
TN	32	36	37	39
FP	8	4	3	2
FN	7	4	4	1

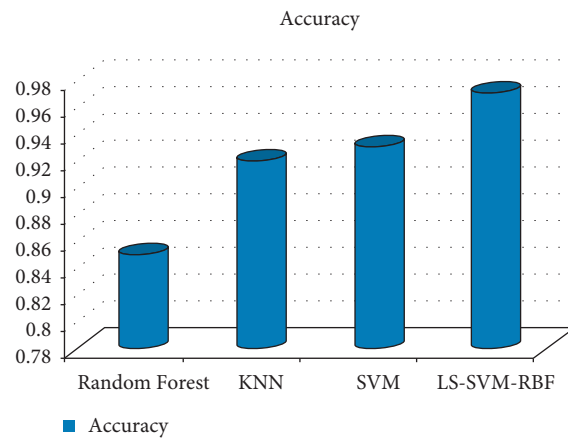


FIGURE 4: Accuracy of classifiers for Alzheimer 's disease detection with Haar wavelet transform feature extraction.

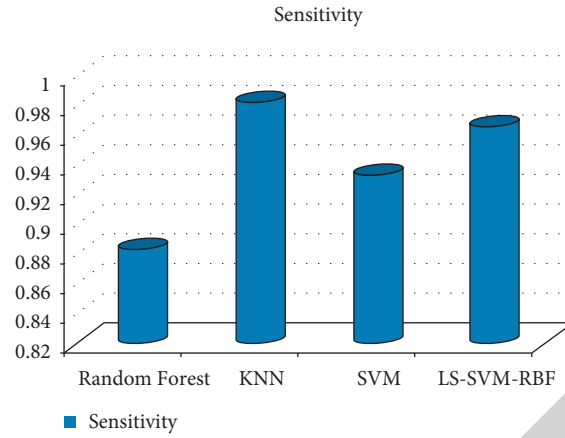


FIGURE 5: Sensitivity of classifiers for Alzheimer's disease detection with Haar wavelet transform feature extraction.

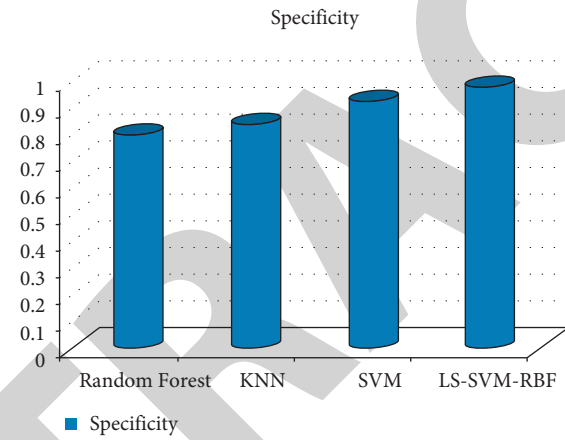


FIGURE 6: Specificity of classifiers for Alzheimer's disease detection with Haar wavelet transform feature extraction.

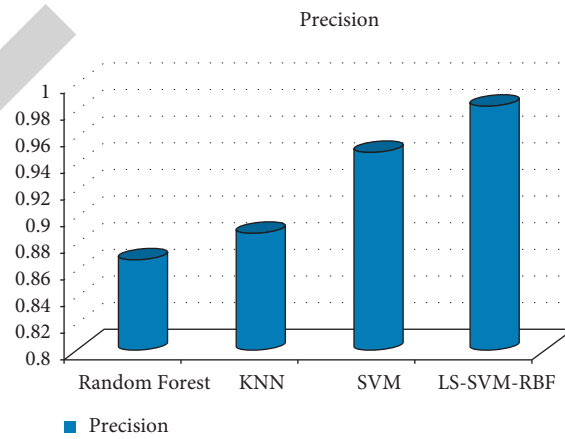


FIGURE 7: Precision of classifiers for Alzheimer's disease detection with Haar wavelet transform feature extraction.

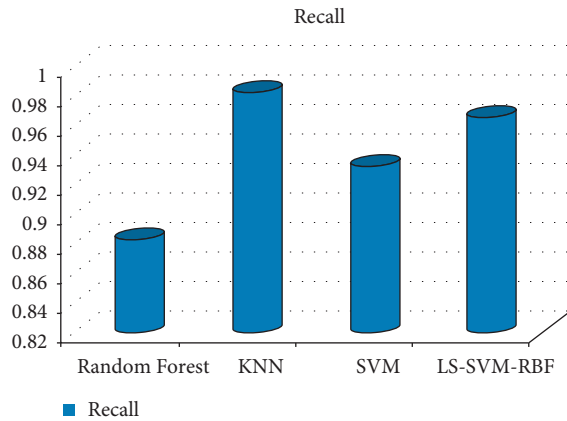


FIGURE 8: Recall of classifiers for Alzheimer's disease detection with Haar wavelet transform feature extraction.

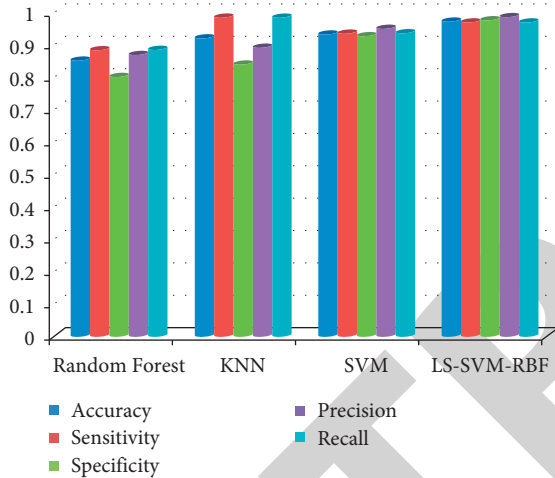


FIGURE 9: Accuracy, specificity, sensitivity, precision, and recall of classifiers for Alzheimer's disease detection with Haar wavelet transform feature extraction.

LS-SVM-RBF is higher than that of the other classifiers. Figure 9 depicts the overall comparison.

5. Conclusion

Alzheimer's disease is distinguished by the formation of abnormal protein bundles in brain tissue, but experts are unsure what causes the condition. To find a cure or aversion, researchers need to know more than just that there are protein deviations from the norm; they also need to know how these brain nerves arise in order to find a solution. Machine learning is the study of computing techniques for improving performance on a given task through the process of learning. This paper describes a framework for detecting Alzheimer's disease that includes image demising of an MRI input data set with an adaptive mean filter, preprocessing with histogram equalization, feature extraction with the Haar wavelet transform, and classification with the LS-SVM-RBF, SVM, KNN, and random forest classifiers. The adaptive mean filter is used to remove noise from preexisting MRI

images. Image quality is improved by histogram equalization. To compare experimental results, the measures of accuracy, sensitivity, specificity, precision, and recall are used. LS-SVM-RBF outperforms the other classifiers in terms of accuracy. In terms of sensitivity and recall, the KNN method outperforms the other classifiers. LS-SVM-RBF has a higher specificity than the other classifiers.

Data Availability

The data are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D Deep Learning on Medical Images: A Review," pp. 1–13, 2020, <https://arxiv.org/abs/2004.00218>.
- [2] T. Jo, K. Nho, and A. J. Saykin, "Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data," *Frontiers in Aging Neuroscience*, vol. 11, 2019.
- [3] J. Wen, E. Thibeau-Sutre, M. Diaz-Melo et al., "Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation," *Medical Image Analysis*, vol. 63, Article ID 101694, 2020.
- [4] E. Altinkaya, K. Polat, and B. Barakli, "Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: a comprehensive review," *Journal of the Institute of Electronics and Computer*, vol. 1, pp. 39–53, 2020.
- [5] P. C. Physicians, "Alzheimer's disease facts and figures," *Alzheimer's and Dementia*, vol. 16, no. 3, pp. 391–460, 2020.
- [6] Y. Yang, X. Li, P. Wang, Y. Xia, and Q. Ye, "Multi-Source transfer learning via ensemble approach for initial diagnosis of Alzheimer's disease," *IEEE J Transl Eng Heal Med*, vol. 8, pp. 1–10, 2020.
- [7] S. Basaia, F. Agosta, L. Wagner et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks," *NeuroImage: Clinic*, vol. 21, Article ID 101645, 2019.
- [8] D. Pan, A. Zeng, L. Jia, Y. Huang, T. Frizzell, and X. Song, "Early detection of Alzheimer's disease using magnetic resonance imaging: a novel approach combining convolutional neural networks and ensemble learning," *Frontiers in Neuroscience*, vol. 14, pp. 1–19, 2020.
- [9] J. Ruiz, M. Mahmud, M. Modasshir, and M. Shamim Kaiser, "3D DenseNet ensemble in 4-way classification of Alzheimer's disease In Brain Informatics," *Cham: Springer International Publishing*, pp. 85–96, New York, NY, USA, 2020.
- [10] A. Simon, M. Deo, V. Selvam, and R. Babu, "An overview of machine learning and its applications," *International Journal of Electrical Sciences & Engineering*, vol. 1, pp. 22–24, 2016.
- [11] Priyanka and S. Balwinder, "An improvement in brain tumor detection using segmentation and bounding box," *International Journal of Computer Science and Mobile Computing*, vol. 2, pp. 239–246, 2013.

Retraction

Retracted: Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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. Nasrullah and A. Jalali, "Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9404242, 6 pages, 2022.

Research Article

Detection of Types of Mental Illness through the Social Network Using Ensembled Deep Learning Model

Syed Nasrullah ¹ and Asadullah Jalali ²

¹Department of Information Systems, College of Computer Engineering & Sciences, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia

²American University of Afghanistan, STM (Science Technology Mathematics), Kabul, Afghanistan

Correspondence should be addressed to Asadullah Jalali; ajalali@auaf.edu.af

Received 22 January 2022; Accepted 9 March 2022; Published 26 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Syed Nasrullah and Asadullah Jalali. 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 today's era, social networking platforms are widely used to share emotions. These types of emotions are often analyzed to predict the user's behavior. In this paper, these types of sentiments are classified to predict the mental illness of the user using the ensembled deep learning model. The Reddit social networking platform is used for the analysis, and the ensembling deep learning model is implemented through convolutional neural network and the recurrent neural network. In this work, multiclass classification is performed for predicting mental illness such as anxiety vs. nonanxiety, bipolar vs. nonbipolar, dementia vs. nondementia, and psychotic vs. nonpsychotic. The performance parameters used for evaluating the models are accuracy, precision, recall, and F1 score. The proposed ensemble model used for performing the multiclass classification has performed better than the other models, with an accuracy greater than 92% in predicting the class.

1. Introduction

Many people use social media to express their emotions and thoughts [1]. Users frequently share their mental health difficulties or illnesses anonymously on numerous social media platforms or online social health forums [2]. It is possible to express sympathy for individuals with comparable symptoms by joining an online health community. Users also frequently utilize social media to gather information about their symptoms in an attempt at self-diagnosis. Numerous academics have used social media to study people's emotional states and mental illnesses, including depression, anxiety, and schizophrenia. According to the study, recent research gathered tweets from people who had reportedly been diagnosed with depression. Linguistic Inquiry and Word Count (LIWC) was a tool used to analyze and track changes in social interaction on Twitter using the collected messages [3]. To better understand the risk of postpartum depression among Facebook users, another study used specialized psychometric measures to compare the levels of postpartum depression between the pre-and

postnatal periods. Reece et al. [4] analyzed image data to predict social network users' depression in their research. Instagram images were analyzed with face identification and colorimetric analysis. Exploring user posts and N-gram language modeling and vector embedding techniques helped a prior study predict who might develop anxiety illnesses in the future [5].

Several prior research found that social media data efficiently monitored or detected users' moods or prospective mental health issues. This study's objective is to build a deep learning model that can detect various mental health problems in users by accumulating social media-related mental health-related data such as depression, anxiety, BPD, schizophrenia, and autism. Groups (or "subreddits") such as r/depression and r/bipolar were employed to acquire data from Reddit users. We collected data from six subreddits that deal with mental health issues, including bipolar, psychiatric disorders (such as schizophrenia and bipolar disorder), and autism, to determine if someone is depressed or anxious [6]. The subreddits that have been previously researched were utilized in this analysis. Mental health-

related subreddits were discovered among the most popular subreddits using a statistical methodology such as a semi-supervised method and an assessment procedure by experts. Subreddits such as R/depression and all others related to depression are examples of this.

Reddit's mental health subreddits helped us determine if individual posts could be classified as relevant forms of mental illness. The symptoms of depression and bipolar disease are so similar that patients with particular mental disorders may not recognize their most appropriate diagnosis due to the difficulty in diagnosing bipolar disorder. At first glance, we felt that people were seeking help by opening up generic stories about themselves by using phrases such as "mental health," "mental ailment," or "mental state." Initially, individuals are more likely to communicate with others on Reddit in a general health-related channel, but they fail to recognize the unique health issues discussed there. Consequently, we are trying to identify social media users who may be suffering from mental health problems [7]. This investigation aims to answer the following research question.

According to the World Health Organization (WHO) Mental Disorders Fact Sheet in January 2020, more than 264 million individuals of all ages suffer from depression [8]. COVID-19's spread has also been connected to increased mental health disorders. It's only a matter of time before more people get the treatment they need for depression. Suicide is a leading cause of death among people aged 15 to 29 due to a shortage of resources for those who have a mental illness. Every year, more than 80,000 individuals die due to this dearth of medical care. There has been a steady rise in major depressive episodes among adolescents in the United States over the past decade. Mental illness is predicted to have a global economic impact of more than 5 trillion dollars by the year 2030, according to projections [9]. It's quite evident that preventative and intervention approaches are in demand.

Researchers have found that a growing number of people are using social media sites such as Twitter and Facebook to express themselves and connect with others in real-time [10]. As a result, a massive amount of social data is generated, containing traces of useful information about people's preferences, moods, and activities. About 57 percent of the world's population is predicted to be using social media by the year 2021, according to Hootsuite. Over the past year, the number of unique users worldwide has increased by 520 million, or more than 13 percent on an annual basis. According to Conway and colleagues, people's health monitoring and mental health applications have used social media data as a data source. Because of their usefulness as medical decision support systems and the difficulty of diagnosing mental disorders using traditional approaches that rely primarily on surveys and interviews, online screening tools that can analyze indications of mental disorders are valuable. As a result, new screening tools for mental disease can be developed by predicting well-known symptoms from social media data. Several studies have proven, for example, that language patterns can serve as an indicator of a person's mental health state utilizing machine learning techniques.

The following is a general outline of how deep learning techniques are used in research: There are three steps involved in this process: creating a file of various types of mental illness diseases, fitting a model based on that data, and testing estimation accuracy concerning that data. An absence has hindered research on applied mental health issues in the number of datasets available. Crisis informatics, rapid diagnosis, action, and effective treatment can benefit from a solution to this problem. Because recognizing risk factors for mental illness often necessitates quick action, creating successful intervention programs is hampered by these gaps in time. Mentally ill individuals may also have less cooperation with the research team. Textual elements (especially tweets) dominate the reviews, but many other types of social data can investigate psychological signals [11]. For example, visual qualities have received relatively little attention. According to certain studies, photographs are the most popular content posted on social media platforms, despite their enormous volume. Many concerns remain unresolved concerning the social data on mental illness, a relatively new field of study.

The paper's organization is as follows: Section 1 describes the introduction, and section 2 covers the literature survey. Then, in section 3, proposed methodology is explained, and section 4 shows the result. Finally, the last section describes the conclusion and future work, followed by the reference section.

2. Literature Survey

Data collection methods in the field of social media-based mental illness prediction can be divided into two categories: (1) collecting data directly from users with their permission, using surveys and data collection tools, or (2) extracting data from public posts through APIs (application programming interfaces). First, researchers used crowdsourcing platforms or data donation websites such as OurDataHelps to post research information and invite users to participate in the application by filling out questionnaires and consenting to collect their social data. Two of the most commonly used questionnaires for measuring depression are the Center for Epidemiologic Studies Depression Scale (CESD) and the Beck Depression Inventory (BDI). To detect suicidal ideation, the Suicide Probability Scale (SPS) and the Satisfaction with Life Scale (SWLS) can be useful instruments [12]. Since social media platforms' APIs allow developers to access public data, the second approach collects related posts by using keywords or regular expressions associated with those keywords or phrases. Self-reported diagnosis can be defined as the use of regular expressions such as "I was diagnosed with (disorder)" for a variety of mental health issues in the postretrieval process. APIs lead to a collection of posts that need to be analyzed before they can be used in further research. Suicide ideation denial, discussion of other people's suicide, or news or reports about suicide are considered irrelevant and removed from the collection if they contain relevant keywords/phrases. Only posts containing hypothetical statements, rebuttals, or quotations are considered positive samples in the self-report diagnosis process.

A common approach to validating regular expressions is with human annotators, which many researchers suggest, given the issues we discussed in the introduction. Predefined datasets, such as those from the myPersonality project, the Computational Linguistics and Clinical Psychology (CLPsych) [13], and the eRisk workshops [14], can also be used to collect data, as can other methods of collecting test scores and social data. Before the actual analysis, most studies preprocessed the collected data by removing stop words, tweets, hashtags, and URLs as well as lowercase characters [15]. ASCII-converted emojis were also used to facilitate future analysis of the data. Selecting only the most relevant features after feature extraction reduces training time, stimulates interpretation, and increases the likelihood that the model will generalise and avoid overfitting [16]. Support vector machines (SVMs) and a variety of kernels, such as linear and radial basis functions (RBFs), as well as different types of regression, such as linear, log-linear, and logistic, as well as naive Bayes, decision trees, and random forests, are the most widely used machine learning methods for predicting mental disorders [17]. People with depression and suicide-related psychiatric stressors have also been studied using deep learning approaches [18]. Even though feature normalisation and parameter tuning were not well documented in most previous studies, the evaluation mechanism is used to assess the classification model's reliability after prediction. Several metrics and visualisation tools in the literature can be used to evaluate the performance of proposed models, such as classification accuracy, confusion matrix, precision, recall, F1 score, and receiver operating characteristics curve (ROC).

According to a few studies, machine learning methods can identify relevant features [19]. The study of new features that can be used to train a classifier and lead to discovering hidden patterns and relationships in the data is the primary focus of this research, not the development of learning methods. When it comes to mental illness, most research relies on textual content and linguistic patterns. NLP methods, such as those developed by Coppersmith et al. [20], can gain insights into specific mental health disorders, such as PTSD; depression; bipolar disorder; and seasonal affective disorder (SAD). They also emphasised the strong connection between mental disorders and language patterns, such as first-person pronouns, anger words, and other negative emotions. Linguistic Inquiry and Word Count (LIWC) analysis contributed to these findings. With the help of dictionaries that cover various psychological categories, it has been manually built by psychologists. For instance, this technique could extract personal pronouns and positive/negative emotions from the text. Also, popular sentiment analysis tools, SentiStrength [21] and OpinionFinder [22], were used in selected research papers to measure the sentiment of textual expressions [23]. Latent Dirichlet Allocation (LDA) is a topic modeling technique that has been used in several efforts to uncover latent topics from user postings [24].

Despite this, most studies focus on textual features, with only a few incorporating image analysis techniques. For example, Kang et al. used colour compositions and SIFT

descriptors as visual features to extract emotional meanings from a Twitter image [25]. Using the image's hue, saturation, and brightness, Reece et al. [4] can predict Instagram users' signs of depression. As Guntuku et al. [26] have shown, using the VGG-Net [27] image classifier can also be used to predict depression in more recent studies.

In the above discussed paper, the focus was on multiple convolution neural network and these networks takes much amount of time during the analysis. So, this proposed work has used the limited number of convolution neural network for the extraction of the features and later on recurrent neural network is used for performing the classification.

3. Proposed Methodology

In this work, 4 different types of classification have been performed based on the mental illness. The four diseases taken for the classification are anxiety_vs_non_anxiety, bipolar_vs_non_bipolar, dementia_vs_nondementia, and the psychotic_vs_non_psychotic. These four diseases are classified using the deep learning model based on the Ensemble deep learning models. The novelty of this work is to apply the ensemble approach using the LSTM model and the convolutional neural network for performing the classification of different types of mental illness. Let us discuss the proposed methodology in detail.

3.1. Dataset. The dataset has been collected from the reddit using NLTK python library. The dataset consists of 301746 records with labels such as anxiety_vs_non_anxiety, bipolar_vs_non_bipolar, depression_vs_non_depression, and the psychotic_vs_non_psychotic. The dataset will be processed through different process such as data cleaning and lemmatization. The most common category of mental health disorders in America impacts approximately 40 million adults 18 and older. Most parts of the dataset are taken from the research paper [6].

3.2. Ensemble Deep Learning Models. In this work, two recurrent (bidirectional LSTM) and one convolutional neural network is used for performing the classification. In the first stage, dataset is preprocessed and then, the preprocessed data are passed through multiple deep learning models such as recurrent neural network and convolutional neural network. Figure 1 shows the flowchart of the proposed work, and the proposed methodology is discussed in detail.

3.2.1. Word Embedding through Continuous Bag of Words Model (CBOW). In natural language processing, Word2vec is widely regarded as a breakthrough. As a result, it is widely accepted and widely used. An embedding technique called Word2vec is utilized to turn the dataset's text into vector form, which a computer can recognize. There is a vector for every word in your dataset, and the size of the vectors varies with word length. For the CBOW model, context is considered and used as an input. Afterwards, it tries to forecast words that are relevant to the context. When it comes to this,

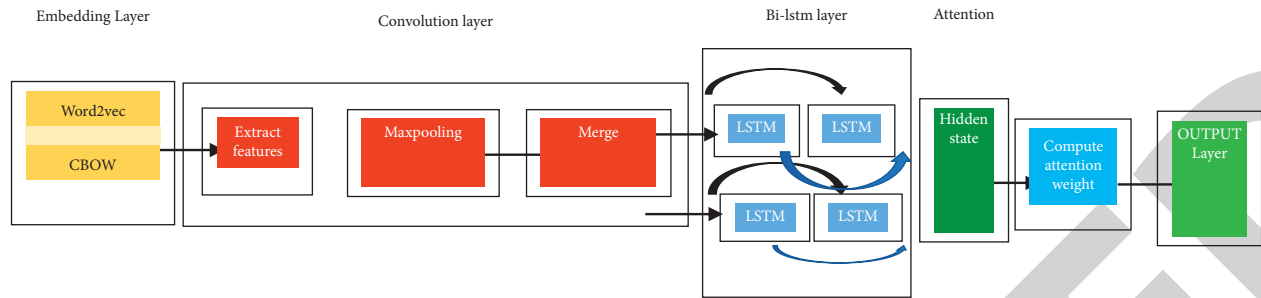


FIGURE 1: Flowchart of the proposed work.

let us look at an example. In this example, the neural network receives the word “depression” as an input and processes the sentence, “Help with rejection, depression, suicide issues.” Predicting the word “depression” is what we are aiming for with this exercise. Input words will be encoded with one-hot encoding, and the error rates of the encoded target word will be measured. Predicting the result based on the term with the least error is easier if we do this.

The above diagram depicts the CBOW model’s architecture. Understanding the context of surrounding words helps the model anticipate a target word. This line, “Help with rejection, depression, suicide issues,” is a good example. This sentence is broken down into word pairs using the approach (context word, target word). The user must customise the window size. For example, if the context word has a 2-word window, the word pairings would be as follows: ([help, with], rejection), ([with, rejection], depression), and ([rejection depression], suicide). When making predictions with this set of word pairs, contextual terms are taken into account. The input layer will consist of four $1 \times W$ input vectors if we employ four context words to predict a single-target word. A $W \times N$ matrix is used to multiply these input vectors in the buried layer. An element-by-element summation of the $1 \times N$ hidden layer output to produce the final activation and output, as shown in Figure 2.

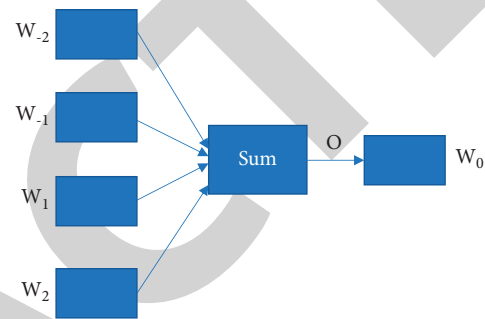


FIGURE 2: Architecture of CBOW word embedding model.

3.2.2. Convolutional Neural Network. In this work, a single CNN model is used for performing the analysis. The input training samples are fed into convolution layers, which extract features. Each convolution layer has a set of filters that aid in extracting features. In general, the complexity of features learned by convolution layers increases as the depth of the CNN model increases. For example, the first convolution layer focuses on simple features, while the last layer focuses on more complex features of training samples. When extracting features, the convolution of a portion of the data sample is considered. Stride length and padding value determine how much data the filter can process at a given time. Before convolution, data samples may or may not be subjected to zero padding. An activation unit called ReLU is used further to process the convolutional output (rectified linear unit). The data is transformed into a nonlinear form by this unit. The ReLU output is clipped to zero when the convolution output is negative. The vanishing gradient problem prevents sigmoid units from being used as

activation units. If the depth of CNN is large, the gradient found at the input layer will have diminished greatly when it reaches the output layer. As a result, the network’s overall output varies slightly. As a result, the convergence rate is either slow or nonexistent. ReLU is preferred to avoid such a situation.

After that, a pooling layer receives the ReLU’s output. During convolution, a pooling layer removes any redundant features that were captured. As a result, the data sample is smaller thanks to this layer. Pooling is based on the assumption that the values of adjacent pixels in an image are nearly identical. Pooling is performed using the average, minimum, and maximum four adjacent pixel values. Using a 2×2 filter, the size of the input image is reduced by half. Before pooling, the input data may or may not be subject to zero padding. In the CNN model, this convolution and pooling layer is repeated. Also, it is repeated 2–4 times for educational purposes. Multilayer neural networks are used to process the output of the convolution and pooling layers. In this case, each neuron unit serves as a feature map for a particular unit. By making the CNN model noise-resistant, the dropout layer reduces overfitting. Layers like these are typically inserted between two layers of a neural network that are already fully connected. As a result, they temporarily disrupted data flow between two completely interconnected systems. This is the same as teaching the model to classify under noisy conditions accurately. As a result, the likelihood of an overfitted model classifying incorrectly decreases. The SoftMax function is used to compute the CNN model’s output. Rather than just ≥ 0.5 for sigmoid output, SoftMax provides the probability of outputs for various classes. The use of SoftMax to find output results based on the most likely

TABLE 1: Comparison table of the deep learning models.

Channel	Class	CNN				Ensemble deep learning model			
		Accuracy (%)	Precision	Recall	F1 score	Accuracy	Precision	Recall	F1 score
Anxiety	Anxiety	77.81	87.54	41.44	56.25	80.13	89.63	50.14	64.30
	Nonanxiety		75.92	96.91	85.14		77.16	98.13	86.39
Bipolar	Bipolar	90.20	87.22	38.02	52.95	92.14	89.13	45.42	60.17
	Nonbipolar		90.40	99.05	94.53		92.56	99.43	95.99
Dementia	Dementia	75.13	89.1	71.75	79.49	80.45	90.46	76.17	82.7
	Nondementia		58.66	82.04	68.41		64.23	86.04	73.55
Psychotic	Psychotic	78.85					83.45	61.23	70.63
	Nonpsychotic						76.73	70.94	73.75

class increases output accuracy. The model is mathematically discussed as follows.

Using L and V as sentence and word vector lengths measures is common. Let $S \in R^{L,V}$ be the sentence matrix. It is necessary to apply a convolutional kernel $C \in R^{L,W}$ to an array of W words before the new feature can be produced using a convolutional operation. There are many ways to generate features such as F_i , from $S[* , i : i + W]$ such as

$$F_i = \sigma \left(\sum (S[* , i : i + W] \circ C) + b \right). \quad (1)$$

In the above equation (1), σ is the nonlinear function and $b \in R$ is a bias term. In this work, ReLu is the nonlinear function and \circ is used to represent the Hadamard product. Hadamard product of two matrices is “To generate a feature map, the convolutional kernel is run on every possible combination of words in the sentence. To identify the most important feature in the feature map, we perform a pairwise max-pooling operation. The pooling operation can be viewed as feature selection in natural language processing. A max-pooling operation is also performed and specifically the feature map, which results from convolution $F = [F_1, F_2, \dots, F_{L-W+1}]$ is the pooling operation’s input. Allowing for a two-fold reduction in input, let’s calculate these two adjacent features in the feature map as follows:

$$M_i = \max(F_{2xi-1}, F_{2xi}). \quad (2)$$

Maximizing the pooling operation yields the following output:

$$M = \left[M_1, M_2, \dots, M_{\frac{L-W+1}{2}} \right], \quad (3)$$

where $M \in R^{(L-W+1/2)}$.

3.2.3. Recurrent Neural Network. For the most part, our classification model is based on an attention-based Bi-LSTM architecture. Each word has a unique correlation with final classification, even though CNNs reduce input features to be used for prediction. We plan to use CNN and Bi-LSTM to their full potential in this study. As a result, the Bi-LSTM is able to encode word dependencies over long distances more effectively. Figure 1 depicts the Bi-LSTM process. Using the CNN stage’s features, it extracts the final hidden layer to generate new features. The information obtained by Bi-LSTM can be viewed as two different textual representations

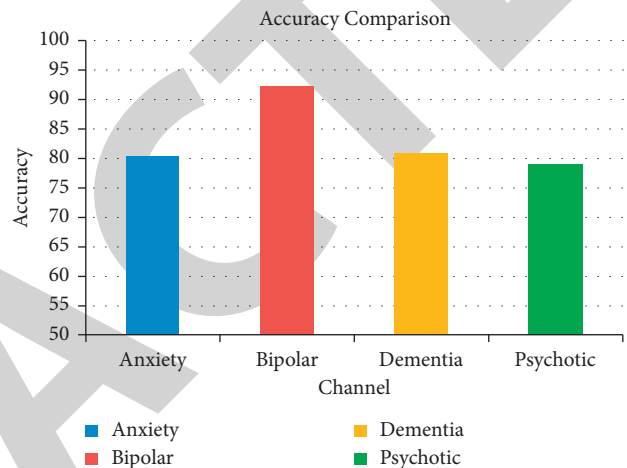


FIGURE 3: Accuracy comparison of the proposed work.

because it has access to both the preceding and subsequent contextual information. A Bi-LSTM model generates a representation of the sequence using the features obtained from the CNN. In the end, the attention layer selects the most correlated features to use for final classification. For “I’m angry. I’m angry at myself,” the attention mechanism assigns more weight to “angry” and less weight to “myself” in order to classify the sentence as depression. The point is to demonstrate that despite the illness’s negative impact, the “angry” attitude remains unfazed. The attention mechanism improves prediction accuracy and reduces the number of learnable weights required by this process. This model’s attention mechanism uses the Bahdanau attention with the following:

$$\text{score}(\text{query}, \text{key}) = V^T \tanh(W_{1\text{key}} + W_{2\text{key}}). \quad (4)$$

4. Results and Discussion

The proposed work is implemented through Python programming language. The parameters used for the evaluation of the performance of the models are accuracy, precision, recall, and F1 score.

Table 1 has shown the comparison of the existing model with the proposed ensemble model. The table has shown that the proposed model has performed better than the existing model based on convolutional neural network.

Retraction

Retracted: Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview

Computational Intelligence and Neuroscience

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 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] M. F. Manzoor, A. Hussain, D. Tazeddinova, A. Abylgazinova, and B. Xu, "Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6531483, 16 pages, 2022.

Review Article

Assessing the Nutritional-Value-Based Therapeutic Potentials and Non-Destructive Approaches for Mulberry Fruit Assessment: An Overview

Muhammad Faisal Manzoor ¹, Abid Hussain,² Diana Tazeddinova,^{3,4}
Aizhan Abylgazina,^{4,5} and Bin Xu ¹

¹School of Food and Biological Engineering, Jiangsu University, Zhenjiang, Jiangsu, China

²Department of Agriculture and Food Technology, Karakoram International University, Gilgit, Pakistan

³Department of Technology and Catering Organization, South Ural State University, Chelyabinsk, Russia

⁴Higher School of Technologies of Food and Processing Productions,
Zhangir Khan West Kazakhstan Agrarian Technical University, Uralsk, Kazakhstan

⁵Scientific-Production Center of Livestock and Veterinary Medicine, Nur-Sultan, Kazakhstan

Correspondence should be addressed to Bin Xu; food_oil@126.com

Received 11 January 2022; Accepted 15 February 2022; Published 24 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Muhammad Faisal Manzoor 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.

Among different fruits, mulberry is the most highlighted natural gift in its superior nutritional and bioactive composition, indispensable for continuing a healthy life. It also acts as a hepatoprotective immunostimulator and improves vision, antimicrobial, anti-cancer agent, anti-stress activity, atherosclerosis, neuroprotective functions, and anti-obesity action. The mulberry fruits also help reduce neurological disorders and mental illness. The main reason for that is the therapeutic potentials present in the nutritional components of the mulberry fruit. The available methods for assessing mulberry fruits are mainly chromatographic based, which are destructive and possess many limitations. However, recently some non-invasive techniques, including chlorophyll fluorescence, image processing, and hyperspectral imaging, were employed to detect various mulberry fruit attributes. The present review attempts to collect and explore available information regarding the nutritional and medicinal importance of mulberry fruit. Besides, non-destructive methods established for the fruit are also elaborated. This work helps encourage many more research works to dig out more hidden information about the essential nutrition of mulberry that can be helpful to resolve many mental-illness-related issues.

1. Introduction

Fruits and vegetables carry health-promoting and bioactive constituents; consequently, consumers' preference has been shifted towards their extensive consumption [1]. Chemical compositions of such food items can protect from various diseases without harming the human body. Mulberry is also a nutritious fruit cultivated 5,000 years ago in China, along with sericulture. Different health-promoting compounds such as moranoline, albafrican, albanol, morusin, kuwanol, calystegine, and hydroxymorcin that can regulate metabolic

activities efficiently have been reported [2, 3]. Explorative studies on mulberry fruit have investigated different health-promoting compounds such as moranoline, albafrican, albanol, morusin, kuwanol, calystegine, and hydroxymorcin that can regulate metabolic activities efficiently [4, 5]. In recent years, some reviews were published based on the health benefits of mulberry fruits [6, 7], and only a few studies discussed analysis methods for the mulberry fruit [8]. Moreover, with the recent global increase in demand for nutrient-dense and high-quality foods, there is a strong emphasis on the non-destructive methods of assessment

with accuracy and high sensitivity for different phytochemicals. Recently, many reviews have discussed the use of non-destructive spectral imaging and spectroscopic techniques for different food and agricultural applications such as exploring the lycopene content, ripening, and maturity of the fruits and quality assessment of alcoholic beverages and spices [9–11].

Therefore, in the current attempt, we discussed non-destructive and rapid methods for the non-invasive investigation of mulberry samples and also elaborated on the importance of mulberry fruit as a nutritional and health-promoting source.

2. Nutritional Significance of Mulberry Fruits

The occurrence of ascorbic acid, carbohydrates, proteins, fats, minerals, and vitamins (thiamine, nicotinic acid, and riboflavin) and their precursors make mulberry fruit the most nutritious agricultural product for consumers [12]. However, a broad range of topographical, climatic, and soil conditions can affect plants' nutritional and chemical status. For example, moisture can range from 80.8 ± 2.81 g/100g FW, ash can vary from 0.6 ± 0.09 g/100g DW, protein can vary from 1.46 ± 0.18 g/100g DW, and the lipid can range from 0.58 ± 0.06 g/100g DW. Likewise, crude fiber (1.2 ± 0.40 g/100g DW), carbohydrate (15.30 ± 1.27 g/100g DW), and energy (72.30 ± 3.25 kcal/100g DW) can also fluctuate [13, 14]. The main sugars are fructose (1.7 – 2.11 g/100 FW) and glucose (1.7 – 2.44 g/100 FW) in mulberry fruit. The pH and the total soluble solids (TSS) in mulberry fruit range from 3.23 to 3.42 and 6.19 to 9.32, respectively [13, 14].

Moreover, the fruit also contains essential minerals elements (both macro and micro), which aids in regulating metabolic mechanisms. Potassium (K), calcium (Ca), magnesium (Mg), and sodium (Na) are essential elements, while the iron (Fe), zinc (Zn), and nickel (Ni) are among the microminerals reported in different studies [14, 15]. Evaluation of essential minerals in different mulberry varieties from different zones confirmed K as the predominant element, ranging from $906.75 \pm 41.49\%$. Trace elements assessed were ranging in a reasonable amount in the fresh matter, but selenium (Se), arsenic (As), and chromium (Cr) are the least dominant elements [16].

In mulberry fruit, 18 different amino acids were also determined, among which 9 were the essential amino acid compulsory for our body. According to the protein ratio, mulberry fruit is close to the sound quality protein foods such as milk and fish. In mulberry essential, amino acid/total amino acid (EAA/TAA) ratio is about 42% [17]. Similarly, in mulberry fruit, the polyunsaturated fatty acid (PUFA) content is higher than monounsaturated and saturated fatty acids. Moreover, behenic acid (C22:0) and palmitoleic acid (C16:1) are essential fatty acids reported in mulberry fruits only. All the reported fatty acids in the fruit are displayed in Table 1.

Additionally, the mulberry fruit also contains riboflavin, thiamin, folate, niacin, and vitamins B-6, A, K, and E. Ascorbic acid is higher, ranging from 36 to 36.4 mg/100g [20]. Similarly, tocopherols were also revealed by Gómez-

Mejía et al. [23] in mulberry fruit, including α -tocopherol (0.73 mg/100 g FW), δ -tocopherol (2.2 mg/100 g FW), and γ -Tocopherol (25 mg/100 g FW). Besides, the fruit also carries organic acids [22], making it a valuable, healthy product, and plays a vital role in sensory properties by imparting sugar and sour taste. Acids are primarily used as an additive, specifically acidulates (malic, tartaric, ascorbic, and citric acids), anti-oxidants (malic, tartaric, and citric acids), or preservatives (benzoic and sorbic acids) [24]. Fruits' base acids have no adverse health effects because during the metabolism they are quickly oxidized. Other vitamins and organic acids present in the fruit are presented in Table 1.

3. Phytochemicals in Mulberry Fruit

Non-nutritive phenolic compounds protect from various dysfunctions with no side effects to the consumer's health [25]. The polyphenolic content reported in mulberry fruit includes flavonoids, fibers, β -carotene, anthocyanins, anthraquinone, glycosides, and oleic acid [2, 7] (Figure 1). Berries are a good source of polyphenolic and indicate the large family, categorized by the structural attribute as phenolic acids, flavonoids, tannins, stilbenes, and lignans, which are indispensable for life [7]. In mulberry fruits, the phenolic contents fluctuate with various cultivars. Beyond the cultivars, the maturity stages of mulberry fruits also have a notable influence on the phenolic values. With the increase in the maturity stage of mulberry fruit, phenolic content also enhances [12, 26].

3.1. Flavonoids. Flavonoids are among a large group of non-nutritive polyphenolic compounds that exhibits anti-oxidant attributes that play a pivotal role in curing oxidative stress-related problems such as atherosclerosis [27]. The variation in flavonoids contents in various breeds of mulberries is notable. Chinese mulberries have higher flavonoid contents (0.0024 mg/kg) than Korean (0.0006 mg/kg) [18]. Quercetin 3-O- β -D-(6''-O-malonyl) glucoside is the most important flavonoid for delivering anti-oxidant properties in the fruit [28]. The kaempferol 3-O-glucoside content was observed in mulberry to be 3.55–47.80 mg/kg of fresh weight in *Morus atropurpurea* cv Taiwanguosuang and Yuefenshen [29]. Besides, mrin (flavonoid) was also revealed to suppress cyclosporine in tissues. Cyclosporine is an effective immune suppressive agent that reduces nitric oxide creation by the activated macrophages [30]. Studies also demonstrated inhibition of the human cytochrome CYP3A process in a pooled human liver microsome system by regular consumption of black mulberry fruit juice [31]. Studies conducted on mice also confirmed anti-stress activity in black mulberry juice due to its valuable phytochemical composition [32].

3.2. Anthocyanins. Anthocyanins are color imparting pigments, widely distributed in agricultural products including fruits, vegetables, flowers, and others [33]. Numerous studies confirmed the presence of health-promoting anthocyanins

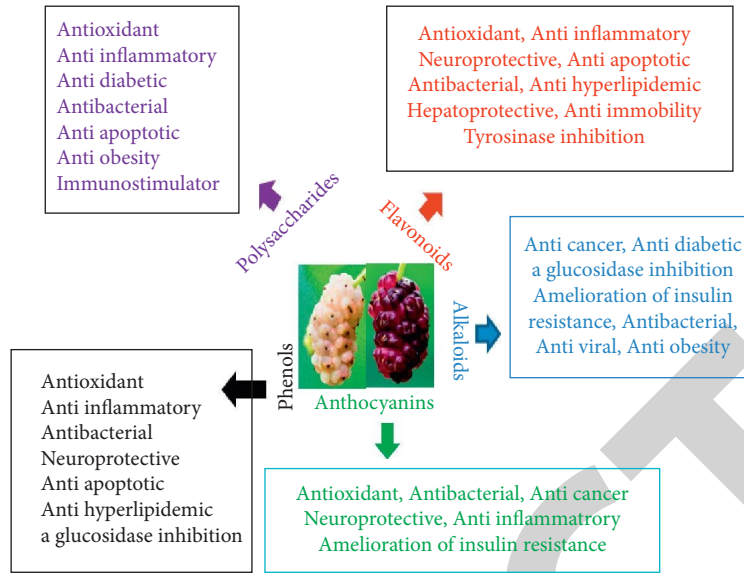


FIGURE 1: Main mulberry functional components and their therapeutic properties.

TABLE 1: Fatty acid composition, vitamins contents, and organic acids profile in mulberry fruit.

Fatty acid		Fatty acid		References
Name	Concentration	Name	Concentration	
Myristic acid	0.47–0.49	Behenic acid	1.3–1.37	[18, 19]
Palmitoleic acid	0.38–0.40	Tetracosanoic acid	1.0–1.04	
Palmitic acid	20–22.26	9-Octadecynoic acid	—	
Heptadecanoic acid	0.26–0.28	10-Nonadecenoic acid	—	
Linoleic acid	26–26.45	Azeloic acid	0.23	
Oleic acid	10.00–10.68	Oxiraneoctanoic acid	0.41	
Stearic acid	8–8.62	11-Eicosenoic acid	—	
Eicosanoic acid	2.10–2.45	Brassicic acid	0.83	
Linolenic acid	0.66	PUFA	74.11–79.52	
MUFA	5.92–6.89	SFA	14.56–19.82	
Vitamin		Vitamin		References
Name	Concentration	Name	Concentration	
Thiamin (B ₁)	0.026–0.029 mg/100 g	Folate DFE ^b	6.00 µg/100 g	[14, 20]
Riboflavin (B ₂)	0.900–0.101 mg/100 g	Vitamin A, RAE ^c	1.00 µg/100 g	
Nicotinic acid	0.700–0.800 mg/100 g	Vitamin A, IU ^a	25 IU/100 g	
Ascorbic acid	36.00–36.40 mg/100 g	Vitamin E (α-tocopherol)	0.80–0.87 mg/100 g	
Niacin	0.600–0.620 mg/100 g	Vitamin K	7.60–7.80 µg/100 g	
Vitamin B-6	0.030–0.050 mg/100 g			
Organic acids		Organic acids		References
Name	Concentration	Name	Concentration	
Malic acid	9.095 mg/g	Citric acid	1.805 mg/g	[21, 22]
Tartaric acid	0.145 mg/g	Oxalic acid	0.660 mg/g	
Fumaric acid	0.213 mg/g	Succinic acid	2.836 mg/g	
Lactic acid	0.662 mg/g	Acetic acid	0.053 mg/g	

^aFatty acids results: Results are described as percentage over the total peak area content of gas chromatography-mass spectrometry analysis. ^bIU = international unit, ^cDFE; dietary folate equivalents, and ^cRAE; retinol activity equivalents.

in mulberry fruit juice. Among them, cyanidin-3-O-glucoside and cyanidin-3-rutinoside are widely distributed (Table 2) [41].

Moreover, cyanidin 3-O-β-D-glucopyranoside isolated from mulberry fruits inhibited the cerebral ischemic damage caused by oxygen-glucose deprivation in PC12 cells [42]. Similarly, anthocyanins from the mulberry fruit (black) were

also reported to prevent Cu-induced peroxidation of liposomes [42]. Meanwhile, the co-oxidation of linoleic acid and β-carotene confirmed that the extracts of mulberry (*Morus nigra*) fruits show safeguard action against peroxidative damage to biomolecules and their membranes [43]. Furthermore, Jiang et al. [44] and Wu et al. [45] concluded that the fruit anthocyanin and water extract could scavenge free

TABLE 2: Composition of flavonoids in mulberry fruit.

Class	Subclass	Compound	Contents	References
Flavonoids	Anthocyanins	Cyanidin 3-O-(6''-O- α -rhamnopyranosyl- β -D-glucopyranoside)	57 mg/g CMA	[34]
		Cyanidin-3-rutinoside	108.78 mg/g MAE	[35]
		Cyanidin-3-glucoside	301.74 mg/g MAE	[36]
		Cyanidin 3-O-(6''-O- α -rhamnopyranosyl- β -D-glucopyranoside)	270 mg/g CMA	[34]
		Cyanidin 3-O- β -D-glucopyranoside	233 mg/g CMA	[34]
		Cyanidin 7-O- β -D-glucopyranoside	33 mg/g CMA	[34]
		Pelargonidin-3- glucoside	NA	[37]
		Pelargonidin-3-rutinoside	NA	[37]
		Petunidin 3-O- β -glucopyranoside	5.1 mg/g CEE	[38]
	Flavanols	Rutin	0.065–7.728 mg/100 g FW	[39]
		Myricetin	0.66–1.18 mg/100 g DW	[40]
		Quercetin	31.88–58.42 mg/100 g DW	[40]
		Quercetin 3-O-glucoside	1.069 mg/100 g FW	[29]
		Quercetin 3-O-rutinoside	2.869 mg/100 g FW	[29]
		Quercetin 3-O-galactoside	0.002 mg/100 g FW	[29]
		Kaempferol	0.24–1.61 mg/100 g DW	[40]
		Kaempferol 3-O-rutinoside	2.00–14.00 mg/100 g DW	[18]
		Kaempferol 3-O-glucoside	1.623 mg/100 g FW	[29]
		Catechin	309.26–750.01 mg/100 g DW	[40]
Flavanols	Epicatechin	8.47–17.12 mg/100 g DW	[40]	
	Epigallocatechin gallate	0.033–0.086 mg/100 g DW	[39]	
	Procyanidin B1	59.64–224.41 mg/100 g DW	[40]	
	Procyanidin B2	1.02–5.66 mg/100 g DW	[40]	

MAE, mulberry anthocyanin extract; NA, not available; CEE, crude ethanol extract; CMA, crude mulberry anthocyanin; DW, dry weight; and FW, frozen weight.

radicals, prevent low-density lipoprotein oxidation, reduce blood lipid, and also be found to prevent atherosclerosis.

3.3. Flavonols and Flavanols. Flavonols and flavanols are flavonoids subgroups, and the structures of these flavonoids are almost the same but different in some positions, such as C-2, C-3, and C-4. In flavonols, a double bond exists between C-2 and C-3 and, in the C ring, the carbonyl group at C-4 compared with flavanols. Mulberry fruit consists of many flavonols such as quercetin, rutin, kaempferol, and myricetin, and the derivatives of kaempferol and quercetin are the main components. Some kaempferol in glycosylated form has been found in some cultivars of mulberry fruit, such as kaempferol 3-O-rutinoside and kaempferol 3-O-glucoside [46, 47]. Usually, the flavanols do not exist as a glycoside naturally. But, in mulberry fruit, epigallocatechin gallate, catechin, procyanidin B1, epicatechin, and procyanidin B2 have been reported (Table 2).

3.4. Phenolic Acids. Mulberry fruit consists of different types of phenolic acids. Benzoic acid and hydroxycinnamic acids are the leading derivatives that represent the phenolic acids in the fruit. Ferulic acid, cinnamic acid, chlorogenic acid, o-coumaric acid, caffeic acid, and p-coumaric acid are the leading derivatives of hydroxycinnamic acid found in the samples. The gallic acid, protocatechuic acid, hydroxybenzoic acid, and vanillic acid are the essential derivatives of benzoic acid in mulberry fruit (Table 3). In mulberry fruit, the most abundant phenolic acid is chlorogenic acid ranging from 5.3 to 17.3 mg/100 g DW [49]. Furthermore, Butkhup et al. [40] reported that cinnamic acid varied from 11.63 to

15.04 mg/100 g DW and gallic acid fluctuated from 7.34 to 23.35 mg/100 g DW, which is the most prominent phenolic acids in different cultivars of the fruit.

3.5. Polysaccharides. Polysaccharides play significant parts in pathological and physiological activities [50–52]. Different polysaccharides were purified from the mulberry fruit with hypoglycemic and anti-oxidant activities using numerous purification methods as presented in Table 4. A glycoprotein extracted from the lyophilized powder and fruit juice of mulberry fruit exhibit good anti-inflammatory and anti-apoptosis agents in rats' primary splenocytes [54].

4. Biosynthesis of Anthocyanin and Phenolic Contents in Mulberry

Anthocyanins are the phenylpropanoid metabolic pathway responsible for exhibiting red, purple, and bluish colors to the mulberry fruits. Its biosynthesis begins with amino acid phenylalanine that is converted by phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumarate-CoA ligase (4CL) to p -coumaroyl-CoA (anthocyanin, flavonols, and lignins precursor). Anthocyanins (cyanidin 3-O-rutinoside) are primarily synthesized by chalcone synthase (CHS), chalcone isomerase (CHI), flavanone-3-hydroxylase (F3H), flavonoid-3'-hydroxylase (F3'H), dihydroflavonol reductase (DFR), anthocyanidin synthase (ANS), anthocyanidin 3-O-glucosyltransferase (3GT), and UDP-rhamnose: anthocyanidin-3-glucoside rhamnosyltransferase (3RT). This biosynthesis is regulated by transcription factors (TFs), including MYB and basic

TABLE 3: Composition of phenolic acid in mulberry fruit.

Class	Subclass	Compound	Content	References
Phenolic acid	Hydroxycinnamic acid	Chlorogenic acid	5.3–17.3 mg/100 g DW	[48]
		Cinnamic acid	11.63–15.04 mg/100 g DW	[40]
		p-Coumaric acid	0.024–0.142 mg/100 g DW	[39]
		o-Coumaric acid	0.015 mg/g FW	[22]
		Ferulic acid	0.057–2.949 mg/100 g DW	[39]
		Caffeic acid	1.06–8.17 mg/100 g DW	[40]
	Benzoic acid	p-Hydroxybenzoic acid	0.028–0.154 mg/100 g DW	[39]
		Protocatechuic acid	0.264–0.794 mg/100 g FW	[39]
		Gallic acid	7.34–23.35 mg/100 g DW	[40]
		Vanillic acid	0.008 mg/g FW	[22]
		Syringic acid	0.049 mg/g FW	[22]

FW, frozen weight and DW, dry weight.

TABLE 4: List of isolated polysaccharides from the mulberry fruit.

Compound name	Molecular weight	Bioactivities	References
FMAP	130	—	[53]
MP	—	Anti-apoptotic and anti-inflammatory	[54]
PMF-1	71.68	—	[55]
PMF-2	84.33	—	
PMF-3	103.17	—	
MFP	—	Hypoglycemic and anti-oxidant	[56]
MFP-1	—	Hypoglycemic and anti-oxidant	
MFP-2	—	Hypoglycemic and anti-oxidant	
MFP-1	7.9, 1.0, 0.7	Hypoglycemic and anti-oxidant	
MFP-2	149, 9.3, 2.6, 1.5	Hypoglycemic and anti-oxidant	[35]
MFP-3	167, 5, 1.5	Hypoglycemic and anti-oxidant	
MFP-4	185, 64.4, 1.5, 0.2	Hypoglycemic and anti-oxidant	[57]
MFP3P	136.6	Hypoglycemic and anti-oxidant	
JS-MP-1	1639	Anti-obesity and immunomodulation	[58]

helix-loop-helix (bHLH) TFs and WD40-repeat proteins also [59]. Cyanidin-3-rutinoside and cyanidin-3-glucoside are the major anthocyanins isolated from mulberry fruits [59].

Among flavonoids, rutin, quercetin, and kaempferol are the major existing ones in mulberry fruits. Some mulberry cultivars have been reported with glycosylated forms of quercetin and kaempferol, including quercetin 3-O-rutinoside, quercetin 3-O-glucoside, quercetin 3-O-galactoside, kaempferol 3-O-glucoside, and kaempferol 3-O-rutinoside [20]. Moreover, rutin was reportedly the most abundant phenolic acid contributing approximately 44.66% of the total phenolic acid content in eleven different mulberry fruit samples [20].

5. Mulberry Homeostasis in Human Guts

The gut response was observed in evaluating anti-oxidant studies (ABTS and FRAP) of mulberry cultivars, whereas, on stimulation of gastrointestinal digestion, white mulberry cultivars exhibited insufficient anti-oxidant activity as compared to their black counterparts. Compared with FRAP assay, white cultivars on digestion possess better anti-oxidant capacity, while the black cultivar showed better results in the non-digested form [60]. Human gut microbiota fermentation effect on the anti-oxidant and phenolic content

was observed, where the black mulberry variety showed higher anti-oxidant activity than white in ABTS scavenging activity, while, for the FRAP assay, ferric reducing capacity showed fluctuations in results collected at 0, 2, 6, 12, and 48 hours. Similarly, the phenolic acid content decreased after fermentation in white while in black initially increased to 960.42 mg gallic acid equivalent (GAE)/100 g FW 24 hours. Still, it then significantly decreased to 543.03 GAE at 48 hours. Anthocyanins were also reportedly increased at 0 h, but with a change in pH, a structural modification was observed during fermentation producing different phenolic metabolites. After in vitro digestion, mulberry anthocyanins degraded owing to the alkaline condition of intestinal digestion. Studies showed that acidic pH was considered more stable for anthocyanins' structural integrity [60, 61]. Studies indicate the potential of digested and fermented mulberry samples in suppressing the outbreak of reactive oxygen species (ROS), as highlighted in Figure 2 [62]. Anthocyanins and catechins also act as a cellular signaling messenger to regulate the anti-oxidant enzymes and activate the Keap1/Nrf2 signaling pathway that could increase the gene expression of anti-oxidant enzymes and resultantly maintain the cellular redox balance [60, 63]. Black mulberry cultivars exhibited higher bioactive compounds with potent in vitro anti-oxidant activity and cellular ROS scavenging activity among the different tested cultivars.

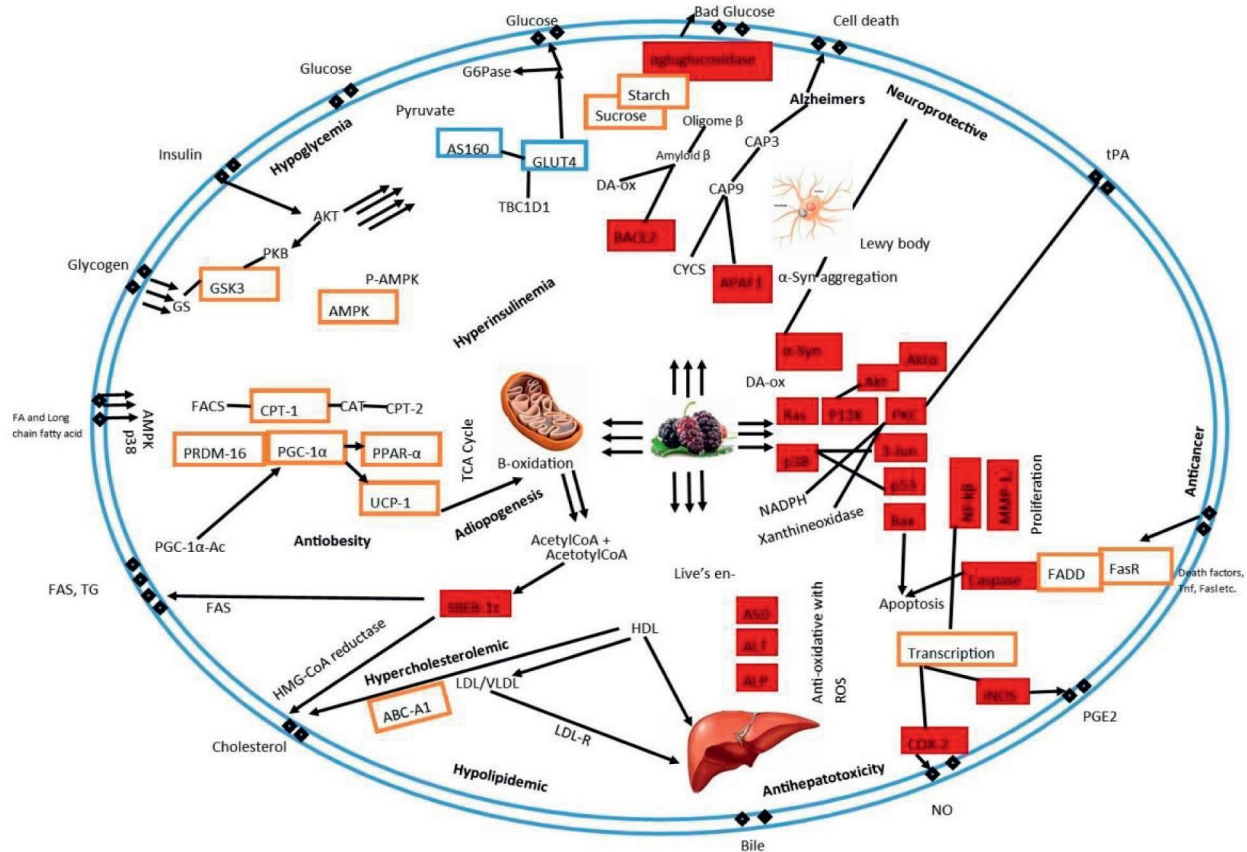


FIGURE 2: The primary health effects mechanism of mulberry's polyphenols designed from available literature. GSH-Px: glutathione peroxidase, CPT-1: carnitine palmitoyltransferase-1, PGC1 α : peroxisome proliferator-activated gamma coactivator 1- α , PPAR α : peroxisome proliferator-activated receptor alpha, UCP1: uncoupling protein 1, PRDM16: PR domain containing 16, HMG-CoA: 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase, SREBP-1c: sterol regulatory element-binding transcription factor 1, LDLR: low-density lipoprotein receptor, AMPK: AMP-activated protein kinase, ABCA1: ATP-binding cassette transporter A1, SR-B1: scavenger receptor class B type 1, GLUT4: glucose transporter type 4, G6Pase: glucose 6-phosphatase, AS160: Akt substrate of 160 kDa, PEPCK: phosphoenolpyruvate carboxykinase, AKT: protein kinase B, FOXO1: fork head box protein O1, GSK-3 β : glycogen synthase kinase-3 β , MMPs: matrix metalloproteinases, NF- κ B: nuclear factor kappa-light-chain-enhancer of activated B cells, ROS: reactive oxygen species, AP-1: activator protein 1, u-PA: urokinase plasminogen activator, p53: phosphoprotein p53, Apaf1: apoptotic peptidase activating factor 1, Bcl-2: b-cell lymphoma 2, Bace2: beta-secretase 2, AST: aspartate aminotransferase, PI3K: phosphatidylinositol-3 kinase, ALT: alanine aminotransferase, COX 2: cyclooxygenase-2, iNOS: inducible nitric oxide synthases, and ALP: alkaline phosphatase. Note: Orange boxes indicate increase and red boxes indicate decrease.

6. Medicinal Value

Due to the health-promoting nutritional composition of mulberry fruit, its applications extend to medicine, economic enhancement, industrial by-products, clinical and domestic fields [32, 42]. Numerous authors stated that mulberry (fruit, roots, and bark) has vital importance in Chinese folk medicine, reported using since 659 AD for the treatment of different ailments such as preventing diabetes, anemia, hypotension, anti-phlogistic, hepatoprotective, diuretic, hypotensive, anti-pyretic, analgesic, expectorant, and also effective against arthritis [18, 64–67]. Furthermore, the fruit and its extracts are useful against epileptic convulsions, mental problems, and hemicranias. The fruit can also prevent asthma, vitiated conditions, rheumatism, and inflammatory issues [42].

Nonetheless, for many years, mulberry fruit juice has also been consumed as a folk medicine for tumors of fauces,

diarrhea, aphtha, flue, cough, dyspepsia, edema, fever, headache, hypertension, and rapid healing of injuries [68]. Moreover, mulberry juice is provided in febrile disorder and malaria to reduce the body temperature because the juice (mulberry) is a natural refrigerant [69]. Similarly, the products (especially syrups and recipes) prepared using mulberry fruits effectively alleviate constipation problems, insomnia, premonitory, and apoplexy dysfunctions [31].

Moreover, the fruit is also effective in treating loss of appetite problems, flatulence, controlling intestinal parasites, and, most importantly, improving the production of body fluids. Different experiments also proved its anti-hyperlipidemic properties, hypertension preventive agent, anti-hyperglycemic, and anti-allergic properties (Table 5) [7, 32, 33, 42, 69].

Chinese pharmacopeia enlists all parts of mulberry (fruit, root bark, stem, and leaves) as a critical constituent in medicinal preparations [84, 85]. Additionally, it poses an

TABLE 5: Therapeutic properties of mulberry fruit.

Therapeutic properties	Intake type	Bioactive compounds	Health effects	References
Hypolipidemic	Mulberry freeze-dried powder	Fatty acids, dietary fiber, phenolics, anthocyanins, flavonoids, and vitamins	A significant decline in serum and liver triglyceride levels, total cholesterol, serum low-density lipoprotein cholesterol, and a decrease in the atherogenic index, while the serum high-density lipoprotein cholesterol was significantly increased	[70]
	Mulberry water extract	Phenolics, anthocyanins, flavonoids, and vitamins	Significant reduction in the levels of low-density lipoprotein, cholesterol, and triglyceride	[71]
Anti-atherosclerotic	Mulberry freeze-dried powder	Anthocyanins	Significant reduction in the low-density lipoprotein cholesterol and total cholesterol	[72]
	Mulberry water extract	Phenolics, anthocyanins, flavonoids, and vitamins	Significant decrease in severe atherosclerosis in the aorta by 42–63%	[71]
	Ethyl acetate-soluble extract	Phenolic compounds (25 different types)	Significant reduction in the glycosylated serum protein and fasting blood glucose and increase anti-oxidant enzymatic activities (SOD, CAT, GSH-Px) in streptozotocin (STZ) induced diabetic mice	[73]
Anti-diabetic	Mulberry fruit polysaccharides extract	Polysaccharides	Significant reduction in fasting blood glucose, oral glucose tolerance test, fasting serum insulin levels, homeostasis model of assessment-insulin resistance, glycated serum protein, and triglycerides	[74]
	Mulberry anthocyanin extract	Anthocyanin	Significant improvement in the dysfunction in diabetic mice and mitigated insulin resistance in HepG2 cells via activation of PI3K/AKT pathways	[75]
Anti-obesity	Ramulus mori polysaccharides extract	Ramulus mori polysacchguoarides	Blood glucose-lowering and metabolism-normalizing roles and also improvement in the function of the pancreas by inhibiting the inflammatory response and attenuating the oxidative stress in pancreas tissue	[76]
	Mulberry water extracts	Gallic acid, chlorogenic acid, rutin, and anthocyanins	Regulation of lipolysis and lipogenesis, which exerted the hypolipidemic and anti-obese effects	[77]
	Mulberry leaf extract and mulberry fruit extract	Cyanidin-3-glucoside, 1-deoxynojirimycin, rutin, and resveratrol	Potential anti-obesity effects through modulation of oxidative stress and obesity-induced inflammation in high fat diet-induced obesity	[78]
Anti-tumor	Mulberry anthocyanins rich extract	Anthocyanin	Targeting c-jun and p38/p53 pathways suppress tumorigenesis and cell survival but produce apoptotic death in AGS cells	[79]
	Extract	Anthocyanin	Potent protective effect on CCl4-induced liver fibrosis in rats	[80]
Hepatoprotective	Mulberry anthocyanin extract	Anthocyanin	Occurrence of the hypolipidemic effects of mulberry anthocyanin extract via inhibition of lipid biosynthesis, phosphorylation of AMPK, and stimulation of lipolysis	[81]
Neuroprotective	Mulberry extract	Cyanidin-3-O-beta-d-glucopyranoside	Neuroprotective effects on the PC12 cells exposed to hydrogen peroxide in vitro and on cerebral ischemic damage in vivo	[82]

TABLE 5: Continued.

Therapeutic properties	Intake type	Bioactive compounds	Health effects	References
Protective against cytotoxicity and oxidative stress	Mulberry juice purification and mulberry marc purification	Total flavonols, total phenolic acids, and anthocyanins contents	Potent anti-oxidant and anti-fatigue properties	[83]

anti-aging effect and imparts positive effects on blood lipid and atherosclerosis [86]. This widely grown fruit also has corrective action against bronchitis, edema, influenza, eye infections, and nosebleeds [85]. Traditionally, mulberry can also be used to treat weakness, fatigue, premature hair falling and graying, urinary problems, tinnitus, dizziness, and hypoglycemic action [42]. In contemporary medicine, mulberry is used to prepare oral syrups, add flavor, or impart color to different drugs [87].

6.1. Anti-Oxidant Potential. Natural anti-oxidants in produce continuously inactivate the reactive species (which damage cells) and keep them in minor amounts, required for normal cell functioning [88]. The in vitro free radical's assays are the most generally employed methods in estimating the anti-oxidant potential of mulberry fruits (Table 6). Generally, the anti-oxidant activity of the whole frozen fruits was 0.21–8.15%, 50.18–86.79%, 16.53–62.83%, and 0.03–38.45 μM ascorbic acid by using a metal chelating ability, DDPH activity, superoxide anion radical scavenging methods, and FRAP activity, respectively [93]. Different experiments have concluded that fruits containing anti-oxidant compounds significantly reduce specific chronic ailments [94]. Berthollide compounds, one of the secondary metabolites, were also detected in mulberry fruits. These bioactive ingredients are free radical scavengers to protect the cell from oxidation [95]. According to another study, the fruit also strengthens the oxidation safeguarding mechanism and inactivates the red-blood-cell-damaging ingredient in diabetes-induced mice [96].

Moreover, research was conducted to compare the anti-oxidant potential of different fruits. Among the tested samples, the mulberry pulp was characterized highest anti-oxidant (ferric reducing the power of 4.11 mmol/100 g wet weight) exhibiting fruit [97]. Furthermore, a spectroscopic assessment also showed the fruit juice contains efficient scavenging characteristics against superoxide, hydroxyl, and nitric acid.

6.2. Immunostimulator. The immune system (IS) is the primary regulatory system managing the body's homeostasis and plays an essential part in developing life from childbirth to death. The IS can be balanced and guarded by using different immunostimulators. Mulberry carries a more significant amount of bioactive flavonoids, particularly anthocyanins and other bioactive compounds that play an essential role in improving the consumer's immunity [98]. *Morus alba* extracts also improved cell-mediated and humoral immunity during experimental animal studies [99].

6.3. Anti-Cancer Agent. Cancer disease is one of the main reasons for the death of both humans and animals [100]. Anthocyanins extracted from the mulberry fruit exhibit inhibitory results on migration and invasion of highly metastatic A549 carcinoma cells (human lung) in a dose-dependent method [64]. The methanolic extract of *M. alba* subdued the production of tumor necrosis factor- α (TNF- α) in macrophages. It inhibited or blocked the production of nitrogen oxide, which was LPS-activated RAW2647 [101]. Mulberry fruit extracted hydroxycinnamic acid derivatives to enhance ROS production by playing as pro-oxidants and destroying the cancer cells [102]. In another study, Huang et al. [79] proposed that mulberry anthocyanins suppressed tumorigenesis and cell survival in AGS gastric cancer xenograft model cells by attacking the c-jun and p38/p53 signaling pathways. Besides, more clinical trials and evaluations verified the curative properties of anthocyanins toward cytotoxic cells, a low-cost and readily accessible source for cancer medication, and decreased cancerous cells [103].

6.4. Hepatoprotector. The liver is one of the essential organs in the human body responsible for nutrients, growth, biochemical pathways, energy supply, and several other basic mechanisms. Hepatotoxins are dangerous elements that can harm the liver [104]. Some specific bioactive compounds in mulberry fruit, such as coumarin, flavonoids, anthocyanins, and stilbenes, were described to own hepatoprotective activity [67]. Furthermore, the hydroalcoholic extract of mulberry was verified to decrease isoniazid-produced hepatotoxicity, a specific enzyme (alanine aminotransferase and aspartate aminotransferase) deficiency ailment [105].

6.5. Atherosclerosis. Atherosclerosis is the deposition of hard yellow plaques of cholesterol in arteries' inner layers, producing heart attack or coronary thrombosis. Investigations on human health proved that dietary intake of natural anti-oxidants inhibits coronary cardiovascular diseases. Oxidation of low-density lipoprotein (LDL) and cholesterol deposition are two essential factors of atherosclerosis. Though, anti-oxidants supplementation could reduce the growth of atherosclerosis ailments [106]. Valuable bioactive compounds such as quercetin and anthocyanins are proclaimed for their shielding effects as anti-oxidant nutrients. Quercetin and its conjugates are principal representatives of the flavonol group of the mulberry; flavonols have potent inhibitory results on oxidative modification of human LDL in vitro [107]. Liu et al. [108] examined the mulberry anthocyanin extract (MAE) and mulberry water extracts (MWE) for the anti-atherosclerosis effect in vitro. The MAE and

TABLE 6: Anti-oxidant potential of mulberry fruit measured through various methods.

Species	ORAC, mmol TE g ⁻¹	ABTS, mg TE 100g ⁻¹	FRAP, mg TE 100g ⁻¹	DPPH, mg TE 100g ⁻¹	Cuprac, mg TE 100g ⁻¹	References
Black mulberry	Nr	2,788.0	1,836.0	946.0	4,046.0	[31]
	Nr	0.68–1.44	0.73–1.69	Nr	Nr	[89]
	Nr	Nr	Nr	11.5–14.5	Nr	[90]
Red mulberry	Nr	0.51–0.73	0.37–0.77	Nr	Nr	[89]
	0.301–1.728	Nr	Nr	Nr	Nr	[91]
	Nr	Nr	Nr	29.19–44.71	Nr	[92]
White mulberry	Nr	Nr	Nr	10.7–12.9	Nr	[90]
	Nr	0.44–1.39	Nr	Nr	Nr	[22]
Different cultivars	Nr	0.0384–0.2073	Nr	0.0362–0.1291	Nr	[18]
	Nr	1.0–325.55	Nr	1.0–160.0	Nr	[29]

Nr: not reported.

MWE scavenged the DPPH radicals and repressed the electrophoretic mobility, the production of thiobarbituric acid reactive substances, and Cu²⁺ induced ApoB fragmentation in oxidation LDL ($p < 0.05$). MAE and MWE also suppress the formation of foam cells and oxidative LDL-induced macrophage death ($p < 0.05$).

6.6. Neuroprotective. According to studies, one of the principal causes of neurodegeneration is caused by free radicals [109]. In mulberry fruit, the occurrence of cyanidin and its 3-O- β -D-glucopyranoside (Cyn3-O- β -D GP) compounds protect consumers against cerebral ischemia [7]. Cyn3-O- β -D GP is a prominent neuroprotective component of mulberry fruit extract [82]. Furthermore, Cyn3-O- β -D GP has free radical scavenging and inflammation suppressing activity and protects the brain from endothelial dysfunction [110]. Similarly, mulberry fruit extract and Cyn3-O- β -D GP can prevent reactive oxygen species production and suppress neuronal disorders. Moreover, in PC12 (oxygen-glucose-deprived) cells, Cyn3-O- β -D GP improves the viability of cells and acts as a neuroprotector against cerebral ischemia.

6.7. Hypolipidemic and Anti-Obesity Action. The anti-obesity activity of mulberries was conducted both in animal and cell models by various mechanisms (Figure 2). Obesity is defined as an elegant fat collection that increases the risk of health. Obesity hurts diabetes, hypercholesterolemia, atherosclerosis, hepatic steatosis, and hyperlipidemia and reduces the number of sugar absorption that ends in body weight. Research on mulberry extract showed the inverse relationship on the melanin-concentrating hormone (MCH) receptor, which is very helpful to reduce body weight [111]. MWE consists of polyphenols such as chlorogenic acid, gallic acid, anthocyanins, and rutin. MWE reduced visceral fat, body weight induced by a high-fat diet accompanied by hypolipidemic effects by lowering cholesterol, serum triacylglycerol, the LDL/HDL ratio, and free fatty acid. Protect the liver from impairment by lowering the hepatic lipids. In a study, MWE significantly raised the receptor α and carnitine palmitoyltransferase-1 of the hepatic peroxisome, while 3-

hydroxy-3-methylglutaryl-coenzyme A reductase and fatty acid synthases enzymes were suppressed. The results indicated that the MWE regulates lipolysis and lipogenesis that eventually imparts hypolipidemic and anti-obese effects [77]. Lately, the mulberry anti-obesity mechanism was revealed. The illustrative mulberry anthocyanin could improve the mitochondrial function via the p38-AMPK-PGC1 α pathway [112]. Moreover, mulberry pelargonidin and cyanidin controlled various obese signs in male C57BL/6 mice, including a high-fat intake [113]. Peng et al. [77] described that after six weeks of feeding with polyphenol-rich extracts of mulberry, the free fatty acids and bodyweight of high-fat intake old male hamsters were decreased.

7. Non-Destructive Techniques and Food Quality Evaluation

With the rapid increase in population and awareness, good quality food is an emerging challenge globally. Therefore, researchers focus on establishing reliable approaches for authenticating agricultural products' quality parameters, including internal and external attributes. Non-destructive powerful spectroscopic techniques have been studied for applications in milk, fish, meat, fruit, vegetable, and beverages [114–118]. The spectral imaging technique is also an accessible option, which combines digital imaging and spectroscopic techniques to provide a powerful analytical device. Such imaging approaches can deliver both spatial and spectral information simultaneously, enabling the detection of the sample down to molecular levels [11, 119].

Non-destructive assessment techniques are the central part of high-quality control functions, and they assist the different established techniques as well. Non-interruptive examination leads to the surface testing of agricultural products without any interfering technique concerning the food quality and appearance. These techniques provide data on food properties such as mechanical, chemical-physical, and structural properties. The employment of a non-destructive assessment is the most suitable way for food processing [120]. Agricultural products possess anti-oxidant attributes due to bioactive compounds such as lycopene, anthocyanins, quercetin, and polyphenols, preventing

cellular oxidation. However, these functional components are highly unstable and can be destroyed by conventional methods such as high-performance liquid chromatography, gas chromatography, thin-layer chromatography, and other techniques. Therefore, to ensure the quality assessment of such nutritious compounds, non-destructive spectroscopic and imaging methods are the best available options.

7.1. Non-Destructive Technique and Mulberry Fruit Assessment

7.1.1. Chlorophyll Fluorescence (CF). CF estimation is a non-interruptive and straightforward tool, widely applied to calculate the degree of pigment changes during ripening stages of different agricultural products [121]. In general, the light absorbed by photosynthetic organisms (using chlorophyll) can undergo three different pathways, whether it can be employed to carry out the photosynthesis process result in heat production, or reemitting fluorescence (red). All the operations occur in the competition; therefore, an increase or decrease in one pathway affects the intensity of others. Moreover, estimation of the total yield of fluorescence (chlorophyll) can provide information about changes in the power of photochemistry and heat generation (Figure 3(a)) [124, 125].

Handheld equipment is cost-effective and readily available in the market. Numerous studies have been reported recently using CF to monitor the ripening stages of fruits, including tomato fruit [124], jujube (*Ziziphus jujuba* Mill.) [126], and tobacco seeds [127].

Furthermore, in another study, CF measurements and red, green, and blue (RGB) intensity were used to investigate sugars, total phenols, ABTS cation, total flavonoid, and DPPH radical scavenging properties during different ripening stages non-destructively. The fitted relationship showed a high correlation relationship between CF and RGB intensities with the tested parameters. A high correlation between CF and tested parameters was estimated from 0.82 to 0.94 during the 4–7 ripening stages, while the correlation between RGB and internal tested parameters (R^2) fluctuated from 0.93 to 0.97 for stages 4–7. The study concluded that the CF and RGB intensity values could non-destructively and rapidly assess the quality of fruits during different ripening stages [128].

7.1.2. Image Processing (IP). In recent years, a combination of machines proved promising in different research areas. For example, machine vision integrated with artificial intelligence delivered the best results for identifying and classification quality attributes of agricultural commodities [123, 129]. Besides, IP and machine vision also aid in the quality control of food items with high accuracy and rapidness and in a non-destructive manner [130, 131]. Numerous research works have been conducted using machine vision efficiently on different fruits and vegetables such as date [132], carrot [133] apples [134], banana [135], potato [136], olive [137], and pomegranate [138]. The main

components of a typical IP and essential steps are presented in Figure 3(b).

Similarly, a study was also designed for grading mulberry fruits based on maturity (ripe, unripe, and overripe) using IP and classification methods. Each segmented sample's color and textural attributes were extracted by employing a correlation-based feature selection subset (CFS) and consistency subset (CONS) as two different feature reduction models. Simultaneously, artificial neural networks (ANN) and support vector machines (SVM) were helpful in the sample classification. ANN classification combined with the CFS subset feature extraction method delivered accuracy of 100%, 100%, and 99.1% and the least mean square error (MSE) calculations of 9.2×10^{-10} , 3.0×10^{-6} , and 2.9×10^{-3} for training, validation, and test sets, respectively. Moreover, the ANN approach integrated with the CONS subset feature extraction approach delivered an acceptable model with the accuracy recorded as 100%, 98.9%, and 98.3%, and MSE resulted as 4.9×10^{-9} , 3.0×10^{-3} , and 3.1×10^{-3} for training, validation, and test sets, respectively, in a study [138].

7.1.3. Hyperspectral Imaging (HSI). A combination of spectral and imaging tools proved promising in numerous food applications recently. The method's sensitivity is due to the generation of three-dimensional data cubes by converting spectral information into spatial data. Hence, HSI can provide a spatial map composed of spectral variations at each pixel [139]. The vibrational attributes of C–H, H–O, C–O, and N–H bonds in the food system can be easily studied by employing the HSI system [140]. Compared to traditional computer vision and human vision, the HSI system has natural advantages that can highlight some of the problematic or impossible features to extract with conventional computer vision systems [141, 142]. With the advancement in optical sensing and imaging approaches, the HSI system has recently become a scientific and effective tool for monitoring and evaluating the quality of fruits and vegetables. The main components of a typical HSI are presented in Figure 3(c).

Due to its high sensitivity and non-destructive nature, HSI was employed to detect different constituents in complex food matrices. For example, visible and near-infrared (Vis-NIR) HSI (covering 400–1,700 nm range) has been investigated for non-destructive pectin polysaccharides detection in Dashi and Guihuami intact mulberry varieties. Also, four types of pectins (DASP, WSP, CSP, and TSP) were stored in the room and at low temperatures to analyze the prediction efficiency. Results revealed easier detection of pectins in Dashi samples than others, and the samples stored at room temperature delivered good prediction compared to low temperature stored samples [143]. Similarly, total anthocyanin values and anti-oxidant attribute was also determined by Huang et al. [144] in mulberry fruit using the Vis-NIR HSI method. The best prediction method for total anthocyanins and anti-oxidant property showed R^2_{val} of 0.959 and 0.995, and RPD was calculated as 4.96 and 14.25, respectively. Moreover, the non-destructive method was also used to monitor various pectins in mulberry fruits at

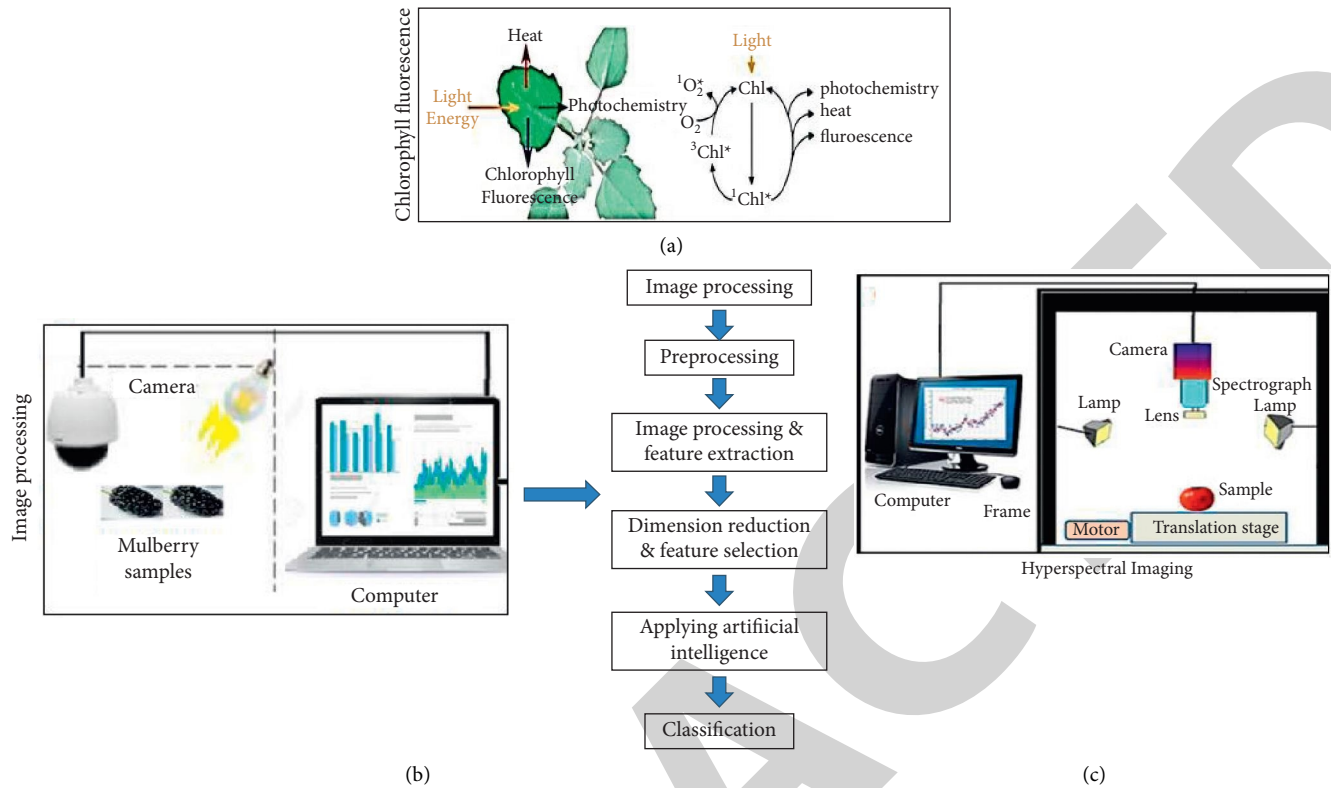


FIGURE 3: (a) Working mechanism of CF [122], (b) imaging system for mulberry classification [123], and (c) main component of HSI system [1].

different storage temperatures. Dilute alkali-soluble pectin (23.52–91.78 g/kg), water-soluble pectin (17.33–117.44 g/kg), chelator soluble pectin (22.91–135.52 g/kg), and total soluble pectin (63.77–344.75 g/kg) were analyzed in the study. The best prediction outcomes were recorded from dilute alkali-soluble pectin and total soluble pectins in the Dashi cultivar stored at room temperature, delivering satisfactory residual predictive deviation results of 2.31 and 1.93, respectively [145].

8. Conclusion and Future Prospects

Investigations on mulberry fruit's health-promoting bioactive compounds proved the disease-fighting attributes in treating various chronic dysfunctions. Clinical trials proved its positive impact against cardiovascular problems, HIV, diabetes, different type of cancers, and obesity and that it can prevent body cell damage, strengthen the nervous system, and can also alleviate many other chronic ailments. Despite these efforts, future work can be focused on the detection of new phytochemicals with more effective and green extraction methods, such as ultrasonication, supercritical fluid extraction, cold plasma method, low polarity water, pulsed electric field, and their integration with other non-thermal methods. In addition, due to the rich phytochemicals and anti-oxidants, the fruit is more worthy for dieticians and health care industries in future research domains. However, some of the polyphenols mechanism

in the human body is still not clear yet, need to be thoroughly studied in future works. Furthermore, the stability of polyphenols is also a challenge and further work can be addressed by proposing novel methods in enhancing their stability.

Likewise, fast, green, label-free, and non-destructive methods are also required for accurate and non-interruptive assessment of mulberry fruit attributes. Numerous imaging methods (such as soft X-ray imaging, laser backscattering imaging, multispectral imaging, resonance imaging, thermal imaging, microwave imaging, and others) and spectroscopic approaches such as surface enhance Raman spectroscopy, near-infrared spectroscopy, Fourier transforms infrared spectroscopy, and others may be the smart choice in upcoming projects.

Data Availability

The data set supporting the conclusions of this article is included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

All the authors equally contributed to this article.

Acknowledgments

This research was funded by National Natural Science Foundation of China (Grant no. 31771900).

References

- [1] A. Hussain, H. Pu, and D.-W. Sun, "Innovative nondestructive imaging techniques for ripening and maturity of fruits - a review of recent applications," *Trends in Food Science & Technology*, vol. 72, pp. 144–152, 2018.
- [2] G. D'urso, J. J. Mes, P. Montoro, R. D. Hall, and R. C. De Vos, "Identification of bioactive phytochemicals in mulberries," *Metabolites*, vol. 10, no. 1, p. 7, 2020.
- [3] E. Gómez-Mejía, C. L. Roriz, S. A. Heleno et al., "Valorisation of black mulberry and grape seeds: chemical characterization and bioactive potential," *Food Chemistry*, vol. 337, Article ID 127998, 2021.
- [4] I. Kim and J. Lee, "Variations in anthocyanin profiles and antioxidant activity of 12 genotypes of mulberry (*Morus* spp.) fruits and their changes during processing," *Antioxidants*, vol. 9, no. 3, 242 pages, 2020.
- [5] S. Shreelakshmi, M. S. Nazareth, S. S. Kumar, P. Giridhar, K. H. Prashanth, and N. P. Shetty, "Physicochemical composition and characterization of bioactive compounds of mulberry (*Morus indica* L.) fruit during ontogeny," *Plant Foods for Human Nutrition*, vol. 76, pp. 1–7, 2021.
- [6] S. Parida, K. Rayaguru, and J. Panigrahi, "Mulberry cultivation and its phytochemical benefits: a review," *Journal of Natural Remedies*, vol. 21, no. 5, pp. 33–48, 2020.
- [7] V. K. Ramappa, D. Srivastava, P. Singh et al., "Mulberries: a promising fruit for phytochemicals, nutraceuticals, and biological activities," *International Journal of Fruit Science*, vol. 20, pp. 1–26, 2020.
- [8] P. Wen, T.-G. Hu, R. J. Linhardt, S.-T. Liao, H. Wu, and Y.-X. Zou, "Mulberry: a review of bioactive compounds and advanced processing technology," *Trends in Food Science & Technology*, vol. 83, pp. 138–158, 2019.
- [9] M. Arslan, H. E. Tahir, M. Zareef et al., "Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques," *Trends in Food Science & Technology*, vol. 107, pp. 80–113, 2021.
- [10] M. Arunkumar, A. Rajendran, S. Gunasri, M. Kowsalya, and C. Krithika, "Non-destructive fruit maturity detection methodology-A review," *Materials Today Proceedings*, 2021.
- [11] A. Hussain, H. Pu, and D.-W. Sun, "Measurements of lycopene contents in fruit: a review of recent developments in conventional and novel techniques," *Critical Reviews in Food Science and Nutrition*, vol. 59, no. 5, pp. 758–769, 2019.
- [12] S. Nayab, K. Razzaq, S. Ullah et al., "Genotypes and harvest maturity influence the nutritional fruit quality of mulberry," *Scientia Horticulturae*, vol. 266, Article ID 109311, 2020.
- [13] F. Koyuncu, M. Çetinbas, and İ. Erdal, "Nutritional constituents of wild-grown black mulberry (*Morus nigra* L.)," *Journal of Applied Botany and Food Quality*, vol. 87, 2014.
- [14] S. M. Paunović, P. Mašković, and M. Milinković, "Determination of primary metabolites, vitamins and minerals in black mulberry (*Morus nigra*) berries depending on altitude," *Vegetable Oils in Food Technology: Composition, Properties Uses*, vol. 62, no. 3, pp. 355–360, 2020.
- [15] B. Ma, Y. Luo, L. Jia et al., "Genome-wide identification and expression analyses of cytochrome P450 genes in mulberry (*Morus notabilis*)," *Journal of Integrative Plant Biology*, vol. 56, no. 9, pp. 887–901, 2014.
- [16] S. Ercisli, M. Tosun, B. Duralija, S. Voća, M. Sengul, and M. Turan, "Phytochemical content of some black (*Morus nigra* L.) and purple (*Morus rubra* L.) mulberry genotypes," *Food Technology and Biotechnology*, vol. 48, no. 1, 2010.
- [17] Y. Jiang and W.-J. Nie, "Chemical properties in fruits of mulberry species from the Xinjiang province of China," *Food Chemistry*, vol. 174, pp. 460–466, 2015.
- [18] E. M. Sánchez-Salcedo, P. Mena, C. García-Viguera, J. J. Martínez, and F. Hernández, "Phytochemical evaluation of white (*Morus alba* L.) and black (*Morus nigra* L.) mulberry fruits, a starting point for the assessment of their beneficial properties," *Journal of Functional Foods*, vol. 12, pp. 399–408, 2015.
- [19] E. M. Sánchez-Salcedo, E. Sendra, Á. A. Carbonell-Barrachina, J. J. Martínez, and F. Hernández, "Fatty acids composition of Spanish black (*Morus nigra* L.) and white (*Morus alba* L.) mulberries," *Food Chemistry*, vol. 190, pp. 566–571, 2016.
- [20] Q. Yuan and L. Zhao, "The mulberry (*Morus alba* L.) fruit-A review of characteristic components and health benefits," *Journal of Agricultural and Food Chemistry*, vol. 65, no. 48, 10383 pages, 2017.
- [21] M. R. Bozhüyük, M. Pehlivan, K. Tuncay, and B. Dođru, "Organic acid composition of selected mulberry genotypes from Aras valley," *Atatürk Üniversitesi Ziraat Fakültesi Dergisi*, vol. 46, no. 2, pp. 69–74, 2015.
- [22] M. Gundogdu, F. Muradoglu, R. I. G. Sensoy, and H. Yilmaz, "Determination of fruit chemical properties of *Morus nigra* L., *Morus alba* L. and *Morus rubra* L. by HPLC," *Scientia Horticulturae*, vol. 132, pp. 37–41, 2011.
- [23] E. Gómez-Mejía, C. L. Roriz, S. A. Heleno et al., "Valorisation of black mulberry and grape seeds: chemical characterization and bioactive potential," *Food Chemistry*, vol. 337, Article ID 127998, 2020.
- [24] S. C. Lourenço, M. Moldão-Martins, and V. D. Alves, "Antioxidants of natural plant origins: from sources to food industry applications," *Molecules*, vol. 24, no. 22, p. 4132, 2019.
- [25] M. Wang, L.-X. Gao, J. Wang et al., "Diels-Alder adducts with PTP1B inhibition from *Morus notabilis*," *Phytochemistry*, vol. 109, pp. 140–146, 2016.
- [26] T. Mahmood, F. Anwar, M. Abbas, M. C. Boyce, and N. Saari, "Compositional variation in sugars and organic acids at different maturity stages in selected small fruits from Pakistan," *International Journal of Molecular Sciences*, vol. 13, no. 2, pp. 1380–1392, 2012.
- [27] Y.-G. Liu, J.-L. Yan, Y.-Q. Ji, W.-J. Nie, and Y. Jiang, "Black mulberry ethanol extract attenuates atherosclerosis-related inflammatory factors and downregulates PPAR γ and CD36 genes in experimental atherosclerotic rats," *Food & Function*, vol. 11, no. 4, pp. 2997–3005, 2020.
- [28] S. Sarkhel, "Nutrition importance and health benefits of mulberry leaf extract: a review," *Journal of Pharmacognosy and Phytochemistry*, vol. 9, no. 5, pp. 689–695, 2020.
- [29] Q. Jin, J. Yang, L. Ma, J. Cai, and J. Li, "Comparison of polyphenol profile and inhibitory activities against oxidation and α -glucosidase in mulberry (*GenusMorus*) cultivars from China," *Journal of Food Science*, vol. 80, no. 11, Article ID C2440, 2015.
- [30] K. Thakur, Y.-Y. Zhu, J.-Y. Feng et al., "Morin as an imminent functional food ingredient: an update on its enhanced efficacy in the treatment and prevention of metabolic

- syndromes," *Food & Function*, vol. 11, no. 10, pp. 8424–8443, 2020.
- [31] M. Tomas, G. Toydemir, D. Boyacioglu, R. Hall, J. Beekwilder, and E. Capanoglu, "The effects of juice processing on black mulberry antioxidants," *Food Chemistry*, vol. 186, pp. 277–284, 2015.
- [32] N. Kishore and D. Kumar, "Pharmacological Properties of Mulberry (*Morus Alba*)," *Assessment of Medicinal Plants for Human Health: Phytochemistry, Disease Management, Novel Applications*, vol. 37, 2020.
- [33] W. Kalt, A. Cassidy, L. R. Howard et al., "Recent research on the health benefits of blueberries and their anthocyanins," *Advances in nutrition (Bethesda, Md.)*, vol. 11, no. 2, pp. 224–236, 2020.
- [34] Q. Du, J. Zheng, and Y. Xu, "Composition of anthocyanins in mulberry and their antioxidant activity," *Journal of Food Composition and Analysis*, vol. 21, no. 5, pp. 390–395, 2008.
- [35] C. Chen, L.-J. You, A. M. Abbasi, X. Fu, R. H. Liu, and C. Li, "Characterization of polysaccharide fractions in mulberry fruit and assessment of their antioxidant and hypoglycemic activities in vitro," *Food & Function*, vol. 7, no. 1, pp. 530–539, 2016.
- [36] C. Chen, B. Zhang, X. Fu, and R. H. Liu, "A novel polysaccharide isolated from mulberry fruits (*Morus alba* L.) and its selenide derivative: structural characterization and biological activities," *Food & Function*, vol. 7, no. 6, pp. 2886–2897, 2016.
- [37] H.-P. Huang, Y.-W. Shih, Y.-C. Chang, C.-N. Hung, and C.-J. Wang, "Chemoinhibitory effect of mulberry anthocyanins on melanoma metastasis involved in the Ras/PI3K pathway," *Journal of Agricultural and Food Chemistry*, vol. 56, no. 19, pp. 9286–9293, 2008.
- [38] F. Sheng, Y. Wang, X. Zhao, N. Tian, H. Hu, and P. Li, "Separation and identification of anthocyanin extracted from mulberry fruit and the pigment binding properties toward human serum albumin," *Journal of Agricultural and Food Chemistry*, vol. 62, no. 28, pp. 6813–6819, 2014.
- [39] C. Qin, Y. Li, W. Niu, Y. Ding, R. Zhang, and X. Shang, "Analysis and characterisation of anthocyanins in mulberry fruit," *Czech Journal of Food Sciences*, vol. 28, no. 2, pp. 117–126, 2010.
- [40] L. Butkhup, W. Samappito, and S. Samappito, "Phenolic composition and antioxidant activity of white mulberry (*Morus alba* L.) fruits," *International Journal of Food Science and Technology*, vol. 48, no. 5, pp. 934–940, 2013.
- [41] M. I. De Rosas, L. Deis, L. Martínez, M. Durán, E. Malovini, and J. B. Cavagnaro, "Anthocyanins in nutrition: biochemistry and health benefits," in *Psychiatry and Neuroscience Update*, pp. 143–152, Springer, 2019.
- [42] A. Mandal, "Nutraceutical and medicinal property of mulberry fruits: a review on its pharmacological potential," MDPI, Basel, Switzerland, 2020.
- [43] G. B. Escher, M. B. Marques, M. A. V. do Carmo et al., "*Clitoria ternatea* L. petal bioactive compounds display antioxidant, antihemolytic and antihypertensive effects, inhibit α -amylase and α -glucosidase activities and reduce human LDL cholesterol and DNA induced oxidation," *Food Research International*, vol. 128, Article ID 108763, 2020.
- [44] Y. Jiang, M. Dai, W.-J. Nie, X.-R. Yang, and X.-C. Zeng, "Effects of the ethanol extract of black mulberry (*Morus nigra* L.) fruit on experimental atherosclerosis in rats," *Journal of Ethnopharmacology*, vol. 200, pp. 228–235, 2017.
- [45] C.-S. Wu, T.-J. Chung, Y.-J. Lee, J.-D. Hsu, and H.-J. Lee, "Mulberry supplementation reduces lipid deposition and protects hamster retina from oxLDL damage," *Journal of Functional Foods*, vol. 71, Article ID 104007, 2020.
- [46] M. L. Chávez-González, L. Sepúlveda, D. K. Verma et al., "Conventional and emerging extraction processes of flavonoids," *Processes*, vol. 8, no. 4, p. 434, 2020.
- [47] W.-T. Ju, O.-C. Kwon, H.-B. Kim, G.-B. Sung, H.-W. Kim, and Y.-S. Kim, "Qualitative and quantitative analysis of flavonoids from 12 species of Korean mulberry leaves," *Journal of Food Science & Technology*, vol. 55, no. 5, pp. 1789–1796, 2018.
- [48] T. Mahmood, F. Anwar, M. Abbas, and N. Saari, "Effect of maturity on phenolics (phenolic acids and flavonoids) profile of strawberry cultivars and mulberry species from Pakistan," *International Journal of Molecular Sciences*, vol. 13, no. 4, pp. 4591–4607, 2012.
- [49] M. K. Gecer, M. Akin, M. Gundogdu, S. P. Eyduran, S. Ercisli, and E. Eyduran, "Organic acids, sugars, phenolic compounds, and some horticultural characteristics of black and white mulberry accessions from Eastern Anatolia," *Canadian Journal of Plant Science*, vol. 96, no. 1, pp. 27–33, 2016.
- [50] I. Giavasis, "Bioactive fungal polysaccharides as potential functional ingredients in food and nutraceuticals," *Current Opinion in Biotechnology*, vol. 26, pp. 162–173, 2014.
- [51] J.-H. Xie, M.-L. Jin, G. A. Morris et al., "Advances on bioactive polysaccharides from medicinal plants," *Critical Reviews in Food Science and Nutrition*, vol. 56, no. sup1, pp. 60–84, 2016.
- [52] X. Xu, P. Xu, C. Ma, J. Tang, and X. Zhang, "Gut microbiota, host health, and polysaccharides," *Biotechnology Advances*, vol. 31, no. 2, pp. 318–337, 2013.
- [53] W. Wei, W. Zhou, N. Zang, and L. Jiang, "Structural analysis of a polysaccharide from *fructus mori albae*," *Carbohydrate Polymers*, vol. 70, no. 3, pp. 341–344, 2007.
- [54] C.-J. Liu and J.-Y. Lin, "Protective effects of strawberry and mulberry fruit polysaccharides on inflammation and apoptosis in murine primary splenocytes," *Journal of Food and Drug Analysis*, vol. 22, no. 2, pp. 210–219, 2014.
- [55] R. Tian, "Separation, purification and composition analysis of polysaccharides from *Mori fructus*," *West China Journal of Pharmaceutical Sciences*, vol. 29, pp. 401–404, 2014.
- [56] C. Chen, L.-J. You, A. M. Abbasi, X. Fu, and R. H. Liu, "Optimization for ultrasound extraction of polysaccharides from mulberry fruits with antioxidant and hyperglycemic activity in vitro," *Carbohydrate Polymers*, vol. 130, pp. 122–132, 2015.
- [57] Y. Chen, W. Zhang, T. Zhao et al., "Adsorption properties of macroporous adsorbent resins for separation of anthocyanins from mulberry," *Food Chemistry*, vol. 194, pp. 712–722, 2016.
- [58] J. S. Lee, A. Synytsya, H. B. Kim et al., "Purification, characterization and immunomodulating activity of a pectic polysaccharide isolated from Korean mulberry fruit *Oddi* (*Morus alba* L.)," *International Immunopharmacology*, vol. 17, no. 3, pp. 858–866, 2013.
- [59] S. Zhao, C. H. Park, J. Yang et al., "Molecular characterization of anthocyanin and betulinic acid biosynthesis in red and white mulberry fruits using high-throughput sequencing," *Food Chemistry*, vol. 279, pp. 364–372, 2019.
- [60] T. Bao, Y. Li, J. Xie, Z. Jia, and W. Chen, "Systematic evaluation of polyphenols composition and antioxidant activity of mulberry cultivars subjected to gastrointestinal digestion and gut microbiota fermentation," *Journal of Functional Foods*, vol. 58, pp. 338–349, 2019.

- [61] A. Castañeda-Ovando, M. d. L. Pacheco-Hernández, M. E. Páez-Hernández, J. A. Rodríguez, and C. A. Galán-Vidal, "Chemical studies of anthocyanins: a review," *Food Chemistry*, vol. 113, no. 4, pp. 859–871, 2009.
- [62] J. Bernatoniene and D. Kopustinskiene, "The role of catechins in cellular responses to oxidative stress," *Molecules*, vol. 23, no. 4, p. 965, 2018.
- [63] Y. P. Hwang, J. H. Choi, H. J. Yun et al., "Anthocyanins from purple sweet potato attenuate dimethylnitrosamine-induced liver injury in rats by inducing Nrf2-mediated antioxidant enzymes and reducing COX-2 and iNOS expression," *Food and Chemical Toxicology*, vol. 49, no. 1, pp. 93–99, 2011.
- [64] K.-C. Cheng, C.-J. Wang, Y.-C. Chang et al., "Mulberry fruits extracts induce apoptosis and autophagy of liver cancer cell and prevent hepatocarcinogenesis in vivo," *Journal of Food and Drug Analysis*, vol. 28, no. 1, pp. 84–93, 2020.
- [65] R. A. Kadam, N. D. Dhumal, and V. B. Khyade, "The Mulberry, *Morus alba* (L.): the medicinal herbal source for human health," *International Journal of Current Microbiology and Applied Sciences*, vol. 8, no. 4, pp. 2941–2964, 2019.
- [66] I. Khalifa, D. Xia, K. Dutta, J. Peng, Y. Jia, and C. Li, "Mulberry anthocyanins exert anti-AGEs effects by selectively trapping glyoxal and structural-dependently blocking the lysyl residues of β -lactoglobulins," *Bioorganic Chemistry*, vol. 96, Article ID 103615, 2020.
- [67] C. Ren, Y. Zhang, W. Cui et al., "A polysaccharide extract of mulberry leaf ameliorates hepatic glucose metabolism and insulin signaling in rats with type 2 diabetes induced by high fat-diet and streptozotocin," *International Journal of Biological Macromolecules*, vol. 72, pp. 951–959, 2015.
- [68] H. Zhang, Z. Ma, X. Luo, and X. Li, "Effects of mulberry fruit (*Morus alba* L.) consumption on health outcomes: a mini-review," *Antioxidants*, vol. 7, no. 5, 69 pages, 2018.
- [69] M. Y. Naeem, "Medicinal potentials and health benefits of black mulberry," *Eurasian Journal of Food Science Technology*, vol. 4, no. 1, pp. 1–5, 2020.
- [70] X. Yang, L. Yang, and H. Zheng, "Hypolipidemic and antioxidant effects of mulberry (*Morus alba* L.) fruit in hyperlipidaemia rats," *Food and Chemical Toxicology: An International Journal Published for the British Industrial Biological Research Association*, vol. 48, no. 8-9, pp. 2374–2379, 2010.
- [71] C.-C. Chen, L.-K. Liu, J.-D. Hsu, H.-P. Huang, M.-Y. Yang, and C.-J. Wang, "Mulberry extract inhibits the development of atherosclerosis in cholesterol-fed rabbits," *Food Chemistry*, vol. 91, no. 4, pp. 601–607, 2005.
- [72] A. Anchalee Sirikancharon, A. Akkarach Bumrungpert, W. Wiroje Kaewruang, T. Tipanee Senawong, and P. Patcharanee Pavadhgul, "The effect of mulberry fruits consumption on lipid profiles in hypercholesterolemic subjects: a randomized controlled trial," *Journal of Pharmacy and Nutrition Sciences*, vol. 6, no. 1, pp. 7–14, 2016.
- [73] Y. Wang, L. Xiang, C. Wang, C. Tang, and X. He, "Antidiabetic and antioxidant effects and phytochemicals of mulberry fruit (*Morus alba* L.) polyphenol enhanced extract," *PLoS One*, vol. 8, no. 7, 2013.
- [74] Y. Jiao, X. Wang, X. Jiang, F. Kong, S. Wang, and C. Yan, "Antidiabetic effects of *Morus alba* fruit polysaccharides on high-fat diet- and streptozotocin-induced type 2 diabetes in rats," *Journal of Ethnopharmacology*, vol. 199, pp. 119–127, 2017.
- [75] F. Yan, G. Dai, and X. Zheng, "Mulberry anthocyanin extract ameliorates insulin resistance by regulating PI3K/AKT pathway in HepG2 cells and db/db mice," *Journal of Nutritional Biochemistry*, vol. 36, pp. 68–80, 2016.
- [76] C. Guo, R. Li, N. Zheng, L. Xu, T. Liang, and Q. He, "Antidiabetic effect of ramulus mori polysaccharides, isolated from *Morus alba* L., on STZ-diabetic mice through blocking inflammatory response and attenuating oxidative stress," *International Immunopharmacology*, vol. 16, no. 1, pp. 93–99, 2013.
- [77] C.-H. Peng, L.-K. Liu, C.-M. Chuang, C.-C. Chyau, C.-N. Huang, and C.-J. Wang, "Mulberry water extracts possess an anti-obesity effect and ability to inhibit hepatic lipogenesis and promote lipolysis," *Journal of Agricultural and Food Chemistry*, vol. 59, no. 6, pp. 2663–2671, 2011.
- [78] H. H. Lim, S. O. Lee, S. Y. Kim, S. J. Yang, and Y. Lim, "Anti-inflammatory and antiobesity effects of mulberry leaf and fruit extract on high fat diet-induced obesity," *Experimental Biology and Medicine*, vol. 238, no. 10, pp. 1160–1169, 2013.
- [79] H.-P. Huang, Y.-C. Chang, C.-H. Wu, C.-N. Hung, and C.-J. Wang, "Anthocyanin-rich Mulberry extract inhibit the gastric cancer cell growth in vitro and xenograft mice by inducing signals of p38/p53 and c-jun," *Food Chemistry*, vol. 129, no. 4, pp. 1703–1709, 2011.
- [80] Y. Li, Z. Yang, S. Jia, and K. Yuan, "Protective effect and mechanism of action of mulberry marc anthocyanins on carbon tetrachloride-induced liver fibrosis in rats," *Journal of Functional Foods*, vol. 24, pp. 595–601, 2016.
- [81] J.-J. Chang, M.-J. Hsu, H.-P. Huang, D.-J. Chung, Y.-C. Chang, and C.-J. Wang, "Mulberry anthocyanins inhibit oleic acid induced lipid accumulation by reduction of lipogenesis and promotion of hepatic lipid clearance," *Journal of Agricultural and Food Chemistry*, vol. 61, no. 25, pp. 6069–6076, 2013.
- [82] T. H. Kang, J. Y. Hur, H. B. Kim, J. H. Ryu, and S. Y. Kim, "Neuroprotective effects of the cyanidin-3-O- β -D-glucopyranoside isolated from mulberry fruit against cerebral ischemia," *Neuroscience Letters*, vol. 391, no. 3, pp. 122–126, 2006.
- [83] D.-Q. Jiang, Y. Guo, D.-H. Xu, Y.-S. Huang, K. Yuan, and Z.-Q. Lv, "Antioxidant and anti-fatigue effects of anthocyanins of mulberry juice purification (MJP) and mulberry marc purification (MMP) from different varieties mulberry fruit in China," *Food and Chemical Toxicology*, vol. 59, pp. 1–7, 2013.
- [84] Y. Qiu, S. Zhang, H. Yu et al., "Identification and characterization of two novel geminiviruses associated with paper mulberry (*Broussonetia papyrifera*) leaf curl disease," *Plant Disease*, 2020.
- [85] P. Sharma, "Mulberry as a life savior-a review," *Journal of Pharmacognosy and Phytochemistry*, vol. 9, no. 2, pp. 2445–2451, 2020.
- [86] J. Wattanathorn, S. Muchimapura, W. Thukhammee et al., "Mulberry fruits protects against age-related cognitive decline," *American Journal of Applied Sciences*, vol. 9, no. 9, p. 1503, 2012.
- [87] N. P. Minh and D. T. A. Dao, "Investigation of *Saccharomyces cerevisiae* in fermented mulberry juice," *International Journal of Scientific & Technology Research*, vol. 2, no. 11, pp. 329–338, 2013.
- [88] M. E. Abd El-Hack, M. T. El-Saadony, M. E. Shafi et al., "Antimicrobial and antioxidant properties of chitosan and its derivatives and their applications: a review," *International Journal of Biological Macromolecules*, 2020.
- [89] M. Özgen, S. Serçe, and C. Kaya, "Phytochemical and antioxidant properties of anthocyanin-rich *Morus nigra* and

- Morus rubra fruits,” *Scientia Horticulturae*, vol. 119, no. 3, pp. 275–279, 2009.
- [90] Á. Calín-Sánchez, J. J. Martínez-Nicolás, S. Munera-Picazo, Á. A. Carbonell-Barrachina, P. Legua, and F. Hernández, “Bioactive compounds and sensory quality of black and white mulberries grown in Spain,” *Plant Foods for Human Nutrition*, vol. 68, no. 4, pp. 370–377, 2013.
- [91] M. Isabelle, B. L. Lee, C. N. Ong, X. Liu, and D. Huang, “Peroxyl radical scavenging capacity, polyphenolics, and lipophilic antioxidant profiles of mulberry fruits cultivated in southern China,” *Journal of Agricultural and Food Chemistry*, vol. 56, no. 20, pp. 9410–9416, 2008.
- [92] W. Zhang, F. Han, J. He, and C. Duan, “HPLC-DAD-ESI-MS/MS analysis and antioxidant activities of non-anthocyanin phenolics in mulberry (*Morus alba*L.),” *Journal of Food Science*, vol. 73, no. 6, pp. 512–518, 2008.
- [93] M. M. Natić, D. Č Dabić, A. Papetti et al., “Analysis and characterisation of phytochemicals in mulberry (*Morus alba* L.) fruits grown in Vojvodina, North Serbia,” *Food Chemistry*, vol. 171, pp. 128–136, 2015.
- [94] A. García-Sánchez, A. G. Miranda-Díaz, and E. G. Cardona-Muñoz, “The role of oxidative stress in physiopathology and pharmacological treatment with pro-and antioxidant properties in chronic diseases,” *Oxidative Medicine Cellular Longevity*, vol. 2020, 2020.
- [95] Ö. Kalkışım, “Determination of the pomological and morphological properties of white mulberry types growing in transition region between mild and continental climates,” *Journal of Food Agriculture and Environment*, vol. 11, no. 1, 2013.
- [96] G. Chandrasegaran, C. Elanchezhian, K. Ghosh, and S. Sethupathy, “Berberine chloride ameliorates oxidative stress, inflammation and apoptosis in the pancreas of Streptozotocin induced diabetic rats,” *Biomedicine & Pharmacotherapy*, vol. 95, pp. 175–185, 2017.
- [97] C. Guo, J. Yang, J. Wei, Y. Li, J. Xu, and Y. Jiang, “Antioxidant activities of peel, pulp and seed fractions of common fruits as determined by FRAP assay,” *Nutrition Research*, vol. 23, no. 12, pp. 1719–1726, 2003.
- [98] M. M. Awais and M. Akhtar, “Evaluation of some sugarcane (*saccharum officinarum* L.) extracts for immunostimulatory and growth promoting effects in industrial broiler chickens,” *Pakistan Veterinary Journal*, vol. 32, no. 3, 2012.
- [99] S. E. Bharani, M. Asad, S. S. Dhamanigi, and G. K. Chandrakala, “Immunomodulatory activity of methanolic extract of *Morus alba* Linn. (mulberry) leaves,” *Pakistan Journal of Pharmaceutical Sciences*, vol. 23, no. 1, pp. 63–8, 2010.
- [100] J. Ni and L. Zhang, “Cancer cachexia: definition, staging, and emerging treatments,” *Cancer Management and Research*, vol. 12, pp. 5597–5605, 2020.
- [101] E. M. Choi and J. K. Hwang, “Effects of *Morus alba* leaf extract on the production of nitric oxide, prostaglandin E2 and cytokines in RAW264.7 macrophages,” *Fitoterapia*, vol. 76, no. 7-8, pp. 608–613, 2005.
- [102] A. Trivellini, M. Lucchesini, R. Maggini et al., “Lamiaceae phenols as multifaceted compounds: bioactivity, industrial prospects and role of “positive-stress,”” *Industrial Crops and Products*, vol. 83, pp. 241–254, 2016.
- [103] J. C. Jeong, S. W. Jang, T. H. Kim, C. H. Kwon, and Y. K. Kim, “Mulberry fruit (*Morus fructus*) extracts induce human glioma cell death in vitro through ROS-dependent mitochondrial pathway and inhibits glioma tumor growth in vivo,” *Nutrition and Cancer*, vol. 62, no. 3, pp. 402–412, 2010.
- [104] D. Muhammad, N. Chand, S. Khan, A. Sultan, and M. Mushtaq, “Hepatoprotective role of milk thistle (*Silybum marianum*) in meat type chicken fed aflatoxin B 1 contaminated feed,” *Pakistan Veterinary Journal*, vol. 32, no. 3, 2012.
- [105] F. Muhammad, M. S. Zafar, T. Khaliq, I. Javed, and M. K. Saleemi, “Nephroprotective effects of *Morus alba* linn against isoniazid-induced toxicity in albino rabbits,” *Pakistan Veterinary Journal*, vol. 34, no. 4, 2014.
- [106] S. Aydin, O. Yilmaz, and Z. Gokçe, “Protective effect of *Morus nigra* L.(mulberry) fruit extract on the liver fatty acid profile of Wistar rats,” *Pakistan Journal of Zoology*, vol. 47, no. 1, pp. 255–261, 2015.
- [107] A. J. Day and G. Williamson, “Biomarkers for exposure to dietary flavonoids: a review of the current evidence for identification of quercetin glycosides in plasma,” *British Journal of Nutrition*, vol. 86, no. 1, pp. 105–110, 2001.
- [108] L. K. Liu, H. J. Lee, Y. W. Shih, C. C. Chyau, and C. J. Wang, “Mulberry anthocyanin extracts inhibit LDL oxidation and macrophage-derived foam cell formation induced by oxidative LDL,” *Journal of Food Science*, vol. 73, no. 6, 2008.
- [109] V. R. Muddapu, S. A. P. Dharshini, V. S. Chakravarthy, and M. M. Gromiha, “Neurodegenerative diseases—is metabolic deficiency the root cause?” *Frontiers in Neuroscience*, vol. 14, 2020.
- [110] C. Kaewmool, S. Udomruk, T. Phitak, P. Pothacharoen, and P. Kongtawelert, “Cyanidin-3-O-Glucoside protects PC12 cells against neuronal apoptosis mediated by LPS-stimulated BV2 microglial activation,” *Neurotoxicity Research*, vol. 37, no. 1, pp. 111–125, 2020.
- [111] B. K. Oh, K.-S. Oh, K.-I. Kwon, S. Y. Ryu, Y. S. Kim, and B. H. Lee, “Melanin-concentrating hormone-1 receptor binding activity of pheophorbides isolated from *Morus alba* leaves,” *Phytotherapy Research*, vol. 24, no. 6, pp. 919–923, 2010.
- [112] Y. You, X. Yuan, H. J. Lee, W. Huang, W. Jin, and J. Zhan, “Mulberry and mulberry wine extract increase the number of mitochondria during brown adipogenesis,” *Food & Function*, vol. 6, no. 2, pp. 401–408, 2015.
- [113] T. Wu, X. Qi, Y. Liu et al., “Dietary supplementation with purified mulberry (*Morus australis* Poir) anthocyanins suppresses body weight gain in high-fat diet fed C57BL/6 mice,” *Food Chemistry*, vol. 141, no. 1, pp. 482–487, 2013.
- [114] M. L. Amodio, F. Ceglie, M. M. A. Chaudhry, F. Piazzolla, and G. Colelli, “Potential of NIR spectroscopy for predicting internal quality and discriminating among strawberry fruits from different production systems,” *Postharvest Biology and Technology*, vol. 125, pp. 112–121, 2017.
- [115] Y. Lu, Y. Huang, and R. Lu, “Innovative hyperspectral imaging-based techniques for quality evaluation of fruits and vegetables: a review,” *Applied Sciences*, vol. 7, no. 2, p. 189, 2017.
- [116] M. Moudache, C. Nerín, M. Colon, and F. Zaidi, “Antioxidant effect of an innovative active plastic film containing olive leaves extract on fresh pork meat and its evaluation by Raman spectroscopy,” *Food Chemistry*, vol. 229, pp. 98–103, 2017.
- [117] D. Urbonaviciene and P. Viskelis, “The cis-lycopene isomers composition in supercritical CO2 extracted tomato by-products,” *Lebensmittel-Wissenschaft und -Technologie- Food Science and Technology*, vol. 85, pp. 517–523, 2017.
- [118] K. Wang, D.-W. Sun, Q. Wei, and H. Pu, “Quantification and visualization of α -tocopherol in oil-in-water emulsion based delivery systems by Raman microspectroscopy,”

Retraction

Retracted: Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] J. Zhou and M. Yang, "Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3975853, 8 pages, 2022.

Research Article

Bone Region Segmentation in Medical Images Based on Improved Watershed Algorithm

Jun Zhou ^{1,2} and Mei Yang ^{3,4}

¹School of Artificial Intelligence, Chongqing Business Vocational College, Chongqing 401331, China

²Chongqing Research Institute of Business Economics, Chongqing, China

³Baliuag University, Baliuag, Bulacan, Philippines

⁴School of General Education, Chongqing Business Vocational College, Chongqing 401331, China

Correspondence should be addressed to Mei Yang; 1198768040@qq.com

Received 3 December 2021; Revised 6 February 2022; Accepted 24 February 2022; Published 22 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Jun Zhou and Mei Yang. 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.

Watershed algorithm is widely used in image segmentation, but it has oversegmentation in image segmentation. Therefore, an image segmentation algorithm based on K -means and improved watershed algorithm is proposed. Firstly, Gaussian filter is used to denoise human skeleton image. K -means clustering algorithm is used to segment the denoised image and the connected component with the largest area is extracted as the initial human skeleton region. The initial bone region was morphologically opened and then morphologically closed to eliminate the noise. Morphologically open operation is used to disconnect other human tissues that adhere to the human bone region and eliminate the background noise with small area, while closed operation smoothes the edge of the human bone region and fills the fracture in the contour line. Secondly, the watershed segmentation algorithm is implemented on the image after morphological operation. The similarity degree of two blocks is defined according to the mean difference of gray level of adjacent blocks and the mean value of standard deviation of gray level of pixels in the edge of the block 4-neighborhood. The adaptive threshold T is generated by Otsu method for histogram of gradient amplitude image. If the similarity degree is greater than T , the image blocks will be merged; otherwise, the image blocks will not be merged. The proposed image segmentation algorithm is used to extract and segment the human bone region from 100 medical images containing human bone. The number of blocks segmented by watershed algorithm is 2775 to 3357, but the number of blocks segmented by the proposed algorithm is 221 to 559. The experimental results show that the proposed algorithm effectively solves the oversegmentation problem of watershed algorithm and effectively segments the image target.

1. Introduction

A series of classic image segmentation algorithms—such as watershed segmentation algorithm, active contour growth methods, threshold-based segmentation methods, regional growth methods, clustering algorithms, and artificial neural network segmentation methods—have been applied to image segmentation over recent years [1]. The watershed transformation algorithm combining morphology algorithm and regional growth method features good extraction of enclosure edges and high precision of target segmentation. Image segmentation refers to a process whereby an image is segmented in accordance with certain rules into a

number of parts or subsets; it works as a key step from image processing to image analyzing [2]. To well perform image segmentation and processing, watershed transformation has been extensively used. Reference [3] segmented by using watershed transformation the gray and white target objects from medical nuclear magnetic resonance (MR) images. This algorithm introduces geographical definitions of such as mountain ridge lines and catchment basin into the recognition of image targets, where the gray pixel is considered as the altitude. This algorithm starts with a lower threshold which, however, enables effective target segmentation and achieves the most optimal segmentation progressively as the threshold increases, while

keeping adjacent targets from merging. In this way, it cracks the problem where adjacent targets cannot be segmented correctly by global thresholding because of their closeness. Watershed transformation is sensitive to noises and potential for excessive image segmentation, though featuring rapidness, simplicity, and directness and a good performance in segmenting images with lower contrast and in simplifying subsequent processing [4].

Fuzzy clustering segmentation algorithm is highly concerned in the research of image segmentation. It adopts the degree of membership to represent the category to which a pixel belongs and the fuzzy *C-means* algorithm to segment MR images. This algorithm, however, increases the complexity of computation and notably, requires further improvement on its antinoise performance [5–7]. Compared with the fuzzy *C-means* algorithm, the *K-means* algorithm features a lower complexity of computation and a lower degree of overlapping of its segmentation targets [8, 9].

We come up with a new image segmentation algorithm that integrates *K-means* clustering algorithm and improved watershed algorithm. First of all, we used *K-means* clustering algorithm to carry out initial clustering and, thus, to extract the interested targets. Then, a 4-neighboring zone similarity-based improved watershed algorithm was proposed to segment target zones on the images initially clustered by the *K-means*. Finally, the human bone zones were extracted from 100 medical bone images. The experimental results demonstrate that our proposed algorithm effectively cracks the excessive segmentation problem of watershed algorithm and also enables the segmentation of image targets. Compared with the watershed segmentation algorithm, the proposed algorithm can solve the problem of image oversegmentation and is easy to be disturbed by noise.

2. Methods and Materials

2.1. Gaussian Filter Denoising. It is necessary to denoise the images before the target segmentation on them; the purpose of Gaussian filter denoising was making the target boundary clearer in the image and removing noise from the image. Gaussian filter is a linear filter, which can effectively suppress noise and smooth images. The working principle is similar to the mean filter, which takes the mean value of pixels in the filter window as the output. However, the coefficients of the window template are different from those of the mean filter. The template coefficients of the mean filter are all the same as 1, while the template coefficients of the Gaussian filter decrease with the increase of distance from the center of the template. Therefore, compared with the mean filter, Gaussian filter has less blur on the image. Gaussian filter was used to carry out the denoising. Its template coefficient decreased as its distance to the center of the template expands, and the images it processed featured a lower degree of fuzziness [10, 11]. Therefore, the Gaussian filter is capable of controlling noise effectively and flattening images when in comparison with other mean filters. The 2-dimensional Gaussian filter used for image denoising is shown as

$$h(x, y) = e^{-\frac{x^2 + y^2}{2\sigma^2}}, \quad (1)$$

where (x, y) represents the coordinates of pixels in an image and σ the standard deviation. The larger the standard deviation σ is, the more blurred image will be after noise elimination by Gaussian filtering. Therefore, in order to make the subsequent target contour clearer, the value of σ is set to a smaller range from 0 to 1.

2.2. K-Means Clustering Algorithm. Clustering is to divide a data set into different classes or clusters according to a specific standard (e.g., distance) so that the similarity of data objects in the same cluster is as large as possible, and the difference of data objects not in the same cluster is as large as possible. After clustering, the data of the same class should be gathered together as much as possible, and the data of different classes should be separated as much as possible.

K-means is the most commonly used clustering algorithm based on Euclidean distance, which holds that the closer the distance between two objects, the greater the similarity. *K-means* has the following advantages: easy to understand, the clustering effect is good, although it is local optimal, but often local optimal is enough [12]. When dealing with large data sets, the algorithm can ensure good scalability. The effect is good when the cluster approximates Gaussian distribution. The algorithm has low complexity. The *K-means* clustering algorithm is an unsupervised clustering algorithm and capable of initializing the segmentation. There are usually a host of zones with similar greyscale, which could lead to a great many local minimums. Given this account, the *K-means* clustering algorithm was first used to perform initial clustering in order to decrease the excessive segmentation of the original watershed algorithm. Also, this algorithm features rapidness, simplicity, and effectiveness in categorizing [13]. Let us set (x, y) as the pixel coordinates of digital images and $f(x, y)$ as the greyscale function. Meanwhile, $Q_j^{(i)}$ represents the j th zone after the i th clustering and $\mu_j^{(i+1)}$ the mean of the j th category (the clustering center) after the $i+1$ th clustering.

After the *K-means* segmented the image clustering into K categories, the connected components with the largest area S_i ($i=1, 2, \dots, k$) were extracted to be the human bone zones given the fact that the bone zones are often a piece of connected zone, which is relatively completed and largest in area in a medical image [14]. Let S_i be a pixel subset of an image. If there are connection paths between all the pixels of the S_i , then the pixels are considered to be connected with each other. To extract the objects of interest to the maximum extent, the 8-neighboring pattern was applied. For any pixel p in S_i , the set of pixels that connects to this pixel in S_i is called the connected component of S_i . The area of the connected component S_i is the count of the number of pixels in S_i . argmax_i represents that if the maximum area zone value in the brackets is corresponding to S_i , then the corresponding zone S_i is the human bone zone.

$$S_i = \text{argmax}_i (S_1, \dots, S_i, \dots, S_k). \quad (2)$$

2.3. Denoising by Morphological Operations.

Mathematical morphology developed from set theory, which was originally used to analyze the geometric structure of metal materials and geological samples. It uses structural elements with certain morphology to measure and extract the corresponding shape features in the image so as to achieve the purpose of image analysis and recognition [15]. At present, mathematical morphology has become a new method in the field of digital image processing and pattern recognition, and image processing using mathematical morphology has some unique characteristics.

- (1) It reflects the logical relationship between pixels in an image, not the simple numerical relationship
- (2) It is a nonlinear image processing method and has irreversibility
- (3) It can be implemented in parallel
- (4) It can be used to describe and define various set parameters and features of images

Mathematical morphology firstly takes binary images as the research object, which is called binary morphology. Later, it extends binary morphology to gray image processing, which is called gray morphology. Basic operations include Dilation, Erosion, Opening, and Closing, from which various morphological processing algorithms can be derived and combined.

The bone zone connection components extracted may be mingled with background images, so the bone zone connected component images S_i extracted were converted into binary images BW . These images were denoised by first morphological open operations and then morphological close operations, where \oplus denotes the morphological dilation and \ominus the morphological erosion.

Morphological close operation:

$$A \bullet B = (A \oplus B) \ominus B. \quad (3)$$

Morphological open operation:

$$A \circ B = (A \ominus B) \oplus B. \quad (4)$$

Binary dilation could fill the small holes (compared with the structural elements) in the image and the small concave parts at the edge of the image so as to filter the outside of the image; binary erosion, in contrast, could remove the smaller structural constituents of the image so as to filter the inner of the image and shrink the image [16]. The process of erosion first and dilation then is an open operation, which enables smoothing the contour of the image and cutting off the disconnections and eliminate the tiny convex objects [17]. The process of dilation first and erosion then is a close operation, which also smoothens the contour but, oppositely, removes the narrow disconnections and the long, thin gaps to clear up the small holes and patch the fractures in the contour line [18].

2.4. *Watershed Segmentation Algorithm.* Watershed segmentation based on mathematical morphology is widely used in the field of image segmentation. Judging from the

description and definition process of watershed segmentation, each local minimum in the image to be segmented corresponds to a separate region in the segmentation result, and the region contour of the object to be segmented will be obtained at the end of segmentation [19]. The number of regions is determined by the number of local minima. Due to noise and local irregularity in the image to be segmented, the number of local minima will be larger than the target object with practical significance in the image, resulting in a large number of false contours, which interferes with the recognition of the target contour of interest in the image. This phenomenon is called "oversegmentation." When watershed transform is directly applied to gradient images, the influence of noise and some local irregularities of gradient images often leads to oversegmentation. A common method to solve this problem is to preprocess the image to be segmented, bring prior information into the segmentation process, and limit the number of regions allowed to be segmented. The main idea of the watershed segmentation based on markers is to use the forced minimum calibration algorithm to determine the minimum region in the gradient image first and to compulsorily take the extracted markers as the minimum value of the gradient image to modify the original gradient image. Based on the modified gradient image, watershed segmentation is applied to complete image segmentation. The watershed segmentation algorithm based on markers adopts internal markers and external markers [20]. A marker is a connecting component. The internal marker is related to a target of interest and the external marker is related to the background. The selection of tags includes preprocessing and defining a set of criteria for selecting tags. Marker selection criteria can make gray value, connectivity, size, shape, texture, and other features. After having the internal marks, only these internal marks are used as the minimum value regions for segmentation, and the watershed of the segmentation result is used as the external mark. Then, other segmentation techniques, such as thresholding, are used for each segmented region to separate the background from the target. In this paper, a watershed segmentation algorithm based on markers is implemented by combining morphological preprocessing.

3. Improved Watershed Segmentation Algorithm

Watershed transformation is a typical segmentation method that is based on zones. Its conception originated from geography, and a watershed is geographically defined as the line that separates the zones in which rainwater accumulates (i.e., the catchment basins). Another conception is that the catchment basin is being immersed in the water and there is a hole at its local minimum point (the lowest part of the catchment basin). From this lowest point, the water slowly flows into the catchment basin. When the water originally in different catchment basins is about to converge together, dams are established to stop their convergency. When the immersing completes, each basin is inundated by the water and completely enclosed by the dams. These dams are termed as the watershed line, or the watershed for short [21].

The specific procedure of the watershed segmentation algorithm is as follows:

Let M_1, M_2, \dots, M_k be the local minimum points of image $f(x, y)$, respectively, and B_1, B_2, \dots, B_k be the sets of all points in the catchment basin corresponding to each local minimum point. B_k^n represents the set of all points in the catchment basin B_k at the n th phase, and B^n represents the set of all B_k^n . f_{\min} and f_{\max} represent the maximum and the minimum of the image greyscale, respectively. In the light of the description above, we obtain

$$B^n = \bigcup_{k=1}^k (B_k^n). \quad (5)$$

And the set of all catchment basins is

$$B^{f_{\max}+1} = \bigcup_{k=1}^k (B_k). \quad (6)$$

$T(n)$ represents the set of the points immersed in the water in the image; that is, $T(n) = \{(i, j) | f(i, j) < n\}$, and then, there is

$$B_k^n = B_k \cap T(n). \quad (7)$$

In this way, a binary image corresponding to the original image could be structured. If the point of the original image has $f(i, j) \in T(n)$, then this point is set as 0, and otherwise, as 1. By this way of counting the "0," the number of points below the water level at the n th phase could be acquired [22]. The specific steps are as follows.

First, the set of initial catchment basins was set as $B^{f_{\min}+1} = T(f_{\min}+1)$, and then, the algorithm begins the recursive step. The water level rose from $n=f_{\min}+1$ to $n=f_{\max}+1$ by integer value, and the sets of points immersed in each phase were recorded as $T(n)$. Let there be Q connected components B_q^n at the n th phase $T(n)$, where $q \leq Q$. Let there be an established B^{n-1} at the $n-1$ th phase. For each q , the connected components B_q^n and B^{n-1} were compared. If the two components meet different relationships, then different operations are carried out.

- (1) $B_q^n \cap B^{n-1} = \Phi$, then a new connected component appears; that is, a new catchment basin comes into being. B_q^n and B^{n-1} are merged.
- (2) $B_q^n \cap B^{n-1}$ contains a connected component of $B_q^n \cap B^{n-1}$. This indicates the presence of this catchment basin in B^{n-1} . B_q^n and its corresponding part in B^{n-1} are merged.
- (3) $B_q^n \cap B^{n-1}$ contains two or more connected components of B^{n-1} .

A watershed is required to establish in B_q^n , and its structure may adopt the morphological dilation operation in order to stop the combination of catchment basins [23]. If the images are segmented by the original watershed algorithm, excessive segmentation will be engendered, usually. Here, we come up with an improved watershed algorithm, which mainly aims to subsequently process the segmentation results and to merge the excessively segmented pieces in

accordance with certain rules. The specific steps are as follows:

- (1) Let the image be segmented by watershed segmentation into N pieces, and each piece be S_i and as large as n_i (where $0 < i < N$), and the greyscale at $(x_i, y_i) \in S_i$, (x_i, y_i) be $I(x_i, y_i)$. The mean greyscale of each pieces M_i is counted by

$$M_i = \frac{1}{n_i} \sum_{x_i, y_i} I(x_i, y_i). \quad (8)$$

- (2) Let S_i and S_j be neighboring pieces; then, the greyscale mean difference of the two pieces M_{ij} is

$$M_{ij} = \sqrt{(M_i - M_j)^2}. \quad (9)$$

- (3) Let S_i and S_j be neighboring pieces, $(x_i, y_i) \in S_i$, $(x_j, y_j) \in S_j$; and let (x_i, y_i) and (x_j, y_j) be two neighboring points on the edge of the segmented piece, and both meet the 4-neighboring zone relationship. N_{ij} represents the number of pixel points on the edge of the segmented piece, which meets the 4-neighboring zone relationship. The mean greyscale standard deviation of these pixel points is defined as

$$P_{ij} = \frac{1}{N_{ij}} \sum_{(x_i, y_i)(x_j, y_j)} \sqrt{(I(x_i, y_i) - I(x_j, y_j))^2}. \quad (10)$$

- (4) The similarity of two neighboring segmented pieces is defined as

$$C_{ij} = \frac{\sqrt{(M_{ij} + P_{ij})^2}}{2}. \quad (11)$$

- (5) Neighboring pieces are merged.

The histogram based on gradient amplitude images generates a self-adaptive threshold T . The gradient amplitude of the image is first computed. The gradient of the image function $f(x, y)$ is a vector with size and direction at point (x, y) . Let G_x and G_y represent the gradients in the x -direction and y -direction, respectively. The vector of this gradient can be expressed as the following formula:

$$\Delta f(x, y) = [G_x, G_y]^T = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]^T. \quad (12)$$

Vector amplitude is

$$\text{mag}(\Delta f) = g(x, y) = \sqrt{\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}}. \quad (13)$$

Direction angle is

$$\phi(x, y) = \arctan \left| \frac{\partial f / \partial y}{\partial f / \partial x} \right|. \quad (14)$$

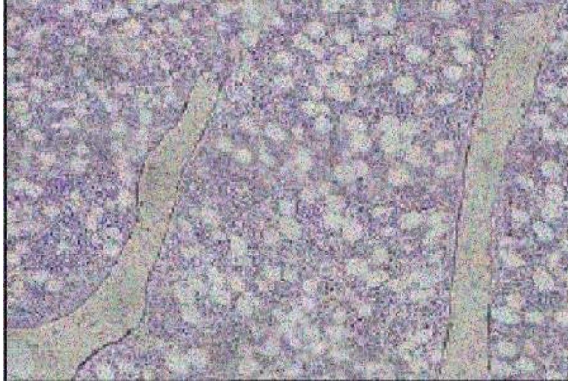


FIGURE 1: Medical image of human bones interfered by noises.

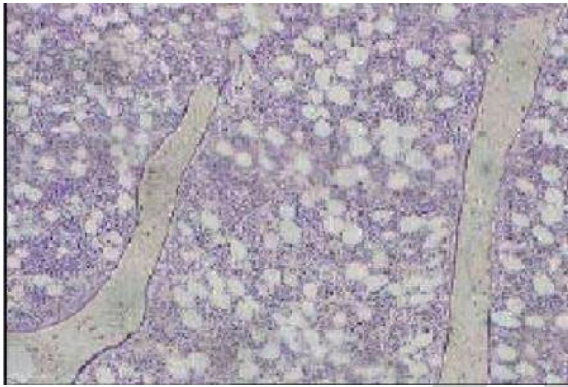


FIGURE 2: Denoising by Gaussian filter.

For digital images, difference is often used to approximate derivatives:

$$\begin{aligned} G_x &= f(x, y) - f(x - 1, y), \\ G_y &= f(x, y) - f(x, y - 1). \end{aligned} \quad (15)$$

After the gradient amplitude of the image is computed, the *Otsu* method [24] is employed for global thresholding, and the globally optimal threshold T is self-adaptively generated. Let us assume that the histogram distribution of an image is expressed as

$$p_r(r_q) = \frac{n_q}{n}, \quad q = 0, 1, 2, \dots, L - 1, \quad (16)$$

where n represents the sum of pixels in the image, n_q the number of pixels of greyscale q , and L all the possible greyscales in the image. Let the initially selected threshold be T , C_1 be the pixel at the greyscale $\{0, 1, \dots, k\}$, and C_2 be a group of pixels at the greyscale $\{k + 1, \dots, L - 1\}$. The *Otsu* methods selects the threshold T such that interclass variance between C_1 and C_2 maximizes. The ratio of interclass variance to overall image greyscale variance is a measurement that can divide the image greyscale into two categories [25], and its value range is $[0, 1]$. At last, the threshold T generated was used to threshold the neighboring segments.

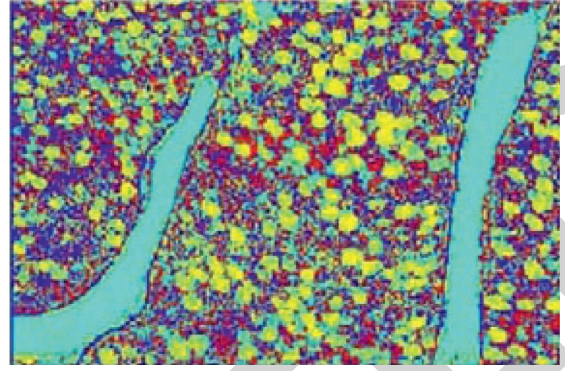
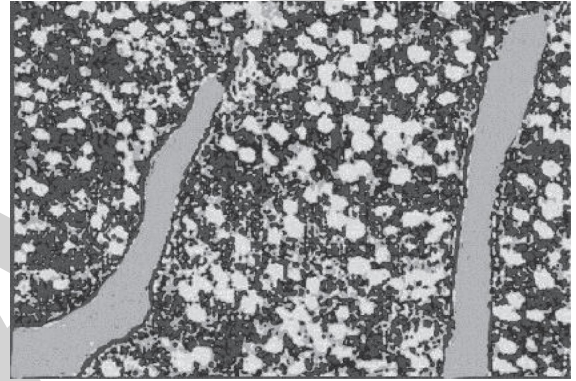
FIGURE 3: K-means clustering effect ($K=4$).

FIGURE 4: K-means clustering segmentation greyscale image.

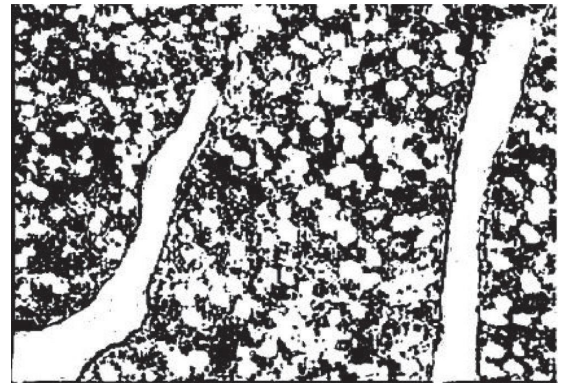


FIGURE 5: Otsu thresholding binary image.

$$\begin{cases} \text{if } C_{ij} > T, & S_i \text{ and } S_j \text{ do not merge,} \\ \text{if } C_{ij} < T, & S_i \text{ and } S_j \text{ merge.} \end{cases} \quad (17)$$

4. Experimental Results and Analysis

The algorithm proposed was compared with the original watershed algorithm in terms of the segmentation performance, where 100 2D medical images of human bones were selected as the experimental subjects. The bone images interfered with salt-and-pepper noise served as the experimental subjected (shown as Figure 1) and denoised by



FIGURE 6: Denoising by morphological operations.

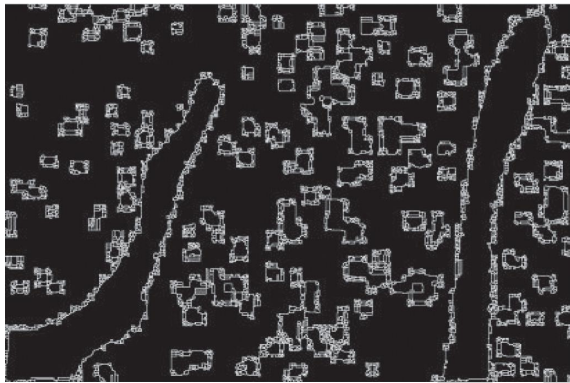


FIGURE 7: Target boundary extraction by watershed transformation.

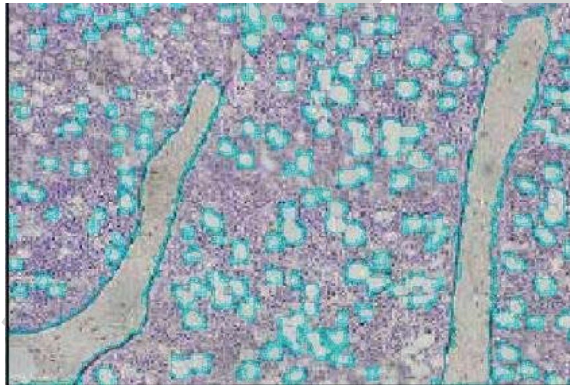


FIGURE 8: Segmentation of bone images.

Gaussian filter (shown as Figure 2), the standard deviation σ of Gaussian filter is 0.8. The *K-means* was applied to carry out initial clustering, with the parameter $K=4$, since the medical images of human bones are regularly made up of bones, soft tissues, fats, and background. The effect is shown in Figure 3. Computer hardware environment: the CPU model is Intel Core I7-4510, the main frequency of the computer is 2 GHz, and the memory is 16G.

The algorithm proposed was applied to segment the images of human bones. First, Figure 3 was converted into a greyscale image, shown in Figure 4. The greyscale image was

then converted by *Otsu* thresholding into a binary image, shown as Figure 5. Next, Figure 6 was processed by morphological close operations to remove the noise after by morphological open operations, where a 7×7 structure unit was selected as the morphological operation structural unit. By the morphological open operations, the human bones were cut off from the connection with other human tissues, and the background noises of a smaller area were also removed. Meanwhile, the morphological close operations flattened the edges of the bone zones, bridged the narrower disconnections, eliminated the smaller hollows, and patched the fractures in the contour line [12]. The results are shown in Figure 7. The watershed transformation in Figure 7 was segmented. Corresponding results were shown in Figure 8. The connectivity of watershed transformation was set to 4.

Since excessive segmentation is the most critical problem of the original watershed algorithm, the number of segments was used as the index for comparison. 100 human bone images were selected to carry out the target segmentation experiment; the numbers of segments made by the original watershed algorithm, the improved watershed algorithm, and the proposed algorithm were compared. The results are shown in Figure 9.

5. Conclusion

We come up with an image segmentation algorithm by combining *K-means* clustering and an improved watershed algorithm in order to crack the excessive segmentation and the noise sensitivity that result from the original watershed algorithm. First of all, *K-means* initial clustering was carried out upon the removal of gauss noise from images, followed by threshold division using the *Otsu* method. Then, morphological open and close operations were employed to eliminate human body noises and thereby to extract human bone zones. On the basis of watershed transformation, next, the similarity was computed for segment integration and precise segmentation of human bone zones. After the application of our proposed algorithm to medical segmentation of bone zones, the experimental results demonstrate the good viability of this algorithm in solving the excessive segmentation of the original watershed algorithm and in segmenting the bone zones from other human tissues. Our research findings are of practical significance to medical image segmentation, providing a reliable means to both computer auxiliary therapy and clinical diagnoses. In this article, however, we experimented only on the medical images of human bones. Therefore, further in-depth explorations are required for the ways of applying our algorithm to a wider range of medical images of other tissues, organs, and so forth. Since the target to be segmented in the image is constantly changing, but the number of targets' classification needs to be known when using the *K-means* algorithm for initial segmentation, the proposed image segmentation algorithm is still not ideal in the number of image blocks generated after image segmentation. All these aspects should be studied in depth. At the same time, the algorithm proposed in the article may be useful for brain

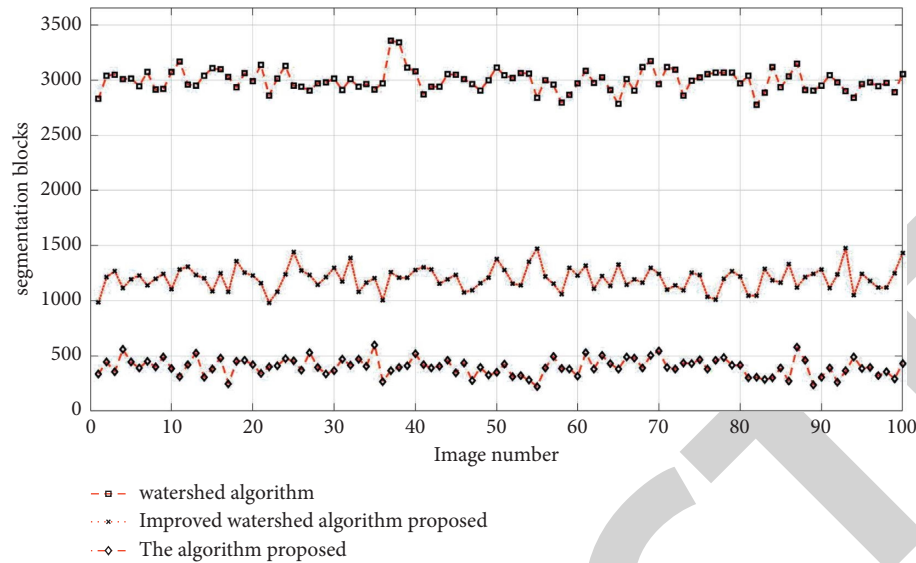


FIGURE 9: Comparison of the number of image segmentation blocks.

imaging detection of mental illness, and further research is needed.

Data Availability

The data included in this paper are available without any restriction.

Disclosure

The same title was presented in the “2021 International Conference on Intelligent Computing Automation and Applications” (ICAA) [25].

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The authors would like to thank the financial support by “Noise Elimination and Identification Model of Acoustic Emission Signal Based on Deep Learning” Project of Science and Technology Research Program of Chongqing Education Commission of China (Grant no. KJZD-K201904401), “Research on Key Technologies of Iris Recognition Based on Wavelet Packet Analysis” Project of Science and Technology Research Program of Chongqing Education Commission of China (Grant no. KJZD-K202004401), “Research on Key Technology of Precision Heavy-Duty Joint Reducer for Special Robot in Nuclear Field” Project of Science and Technology Research Program of Chongqing Education Commission of China (Grant no. KJZD-k202104401), Chongqing Education Science Planning Project (Grant no. 2020-GX-414), Artificial Intelligence Stage achievement of Application Collaborative Innovation Center of Chongqing Business Vocational College, Stage Achievement of Artificial intelligence Trainer Master

Studio of Chongqing Vocational College of Business, and Stage Achievements of Chongqing Vocational College of Commerce VR Interactive Device Space digital Display Application Technology Promotion Center.

References

- [1] Z. Wang, P. Chen, and J. Pan, “Target detection based on complex deep learning,” *Journal of Chongqing University of Technology (Natural Sciences)*, vol. 32, pp. 172–173, 2018.
- [2] X. M. Zhang, P. Wang, and Y. P. Bai, “Improved PSO-SVM based SAR image segmentation recognition,” *Journal of Chongqing University of Technology (Natural Sciences)*, vol. 32, pp. 166–173, 2018.
- [3] V. Grau, R. Kikinis, M. Alcaniz, and S. K. Warfield, “Cortical gray matter segmentation using an improved watershed transform,” in *Proceedings of the 28th Annual International Conference of Engineering in Medicine and Biology Society*, pp. 618–621, New York, NY, USA, 2016.
- [4] X. H. Yuan, D. Xu, and L. Xia, “Target image segmentation based on morphological filter and watershed,” *Data Acquisition and Processing*, vol. 22, pp. 455–489, 2015.
- [5] D. Zhang, D. M. Monro, and S. Rakshit, “Eyelash removal method for human iris recognition,” in *Proceedings of the IEEE International Conference Image Processing*, pp. 285–288, Quebec, Canada, September 2015.
- [6] Q. P. Wu and C. Wu, “Rapid robust kernel spatial fuzzy clustering segmentation,” *China Journal of Image and Graphics*, vol. 23, pp. 1838–1851, 2018.
- [7] C. Y. Jiang, J. X. Liu, H. X. Zhong, H. Y. Li, and D. J. Li, “Intracranial hemorrhage CT image segmentation based improved FCM fuzzy clustering,” *China Medical Equipment*, vol. 33, pp. 16–20, 2018.
- [8] L. Ren, H. C. Lai, Q. Z. Chen, and X. Wang, “Cotton segmentation algorithm based on improved K -means clustering and HSV model,” *Computer Engineering and Design*, vol. 34, pp. 22–24, 2015.
- [9] Y. Zhang, Q. Xie, and C. Zhang, “Key algorithms for segmentation of copperplate printing image based on deep

Retraction

Retracted: Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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. Khan, V. Saravanan, C. N. Gnanaprakasam, T. J. Lakshmi, N. Deb, and N. A. Othman, "Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9985933, 8 pages, 2022.

Research Article

Privacy Protection of Healthcare Data over Social Networks Using Machine Learning Algorithms

Shakir Khan ¹, V. Saravanan ², Gnanaprakasam C. N ³, T. Jaya Lakshmi ⁴,
Nabamita Deb ⁵, and Nashwan Adnan Othman ⁶

¹College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

²Department of Computer Science, College of Engineering and Technology, Dambi Dollo University, Dambi Dollo, Oromia Region, Ethiopia

³Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai 600119, Tamilnadu, India

⁴Department of Computer Science and Engineering, SRM University, Amaravati, AP, India

⁵Department of Information Technology, Gauhati University, Gawahati, Assam 781014, India

⁶Department of Computer Science, College of Science, Knowledge University, Erbil 44001, Iraq

Correspondence should be addressed to V. Saravanan; saravanan@dadu.edu.et

Received 31 January 2022; Revised 12 February 2022; Accepted 17 February 2022; Published 24 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Shakir Khan 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 rapid development of mobile medical care, medical institutions also have the hidden danger of privacy leakage while sharing personal medical data. Based on the k-anonymity and l-diversity supervised models, it is proposed to use the classified personalized entropy l-diversity privacy protection model to protect user privacy in a fine-grained manner. By distinguishing solid and weak sensitive attribute values, the constraints on sensitive attributes are improved, and the sensitive information is reduced for the leakage probability of vital information to achieve the safety of medical data sharing. This research offers a customized information entropy l-diversity model and performs experiments to tackle the issues that the information entropy l-diversity model does not discriminate between strong and weak sensitive features. Data analysis and experimental results show that this method can minimize execution time while improving data accuracy and service quality, which is more effective than existing solutions. The limits of solid and weak on sensitive qualities are enhanced, sensitive data are reduced, and the chance of crucial data leakage is lowered, all of which contribute to the security of healthcare data exchange. This research offers a customized information entropy l-diversity model and performs experiments to tackle the issues that the information entropy l-diversity model does not discriminate between strong and weak sensitive features. The scope of this research is that this paper enhances data accuracy while minimizing the algorithm's execution time.

1. Introduction

With the rapid development of mobile medical technology, the gradual expansion of medical data sharing, and the continuous updating of data mining technology and deep learning technology, the sharing of medical data between different hospitals has become more convenient. The mining and sharing data and information also create enormous economic and social values. However, a problem cannot be ignored in data release and sharing. To calculate the

anonymous cost, this article presents a technique based on the extension level. The different aggregate levels of the attribute must be built first. That is, the privacy of medical patients is leaked. If medical institutions do not fully consider data privacy issues when sharing data, illegal users (attackers) can use data released by other institutions to speculate in series or even use data vulnerabilities released by the same hospital at different times to obtain medical patients' privacy sensitivity and information, thereby causing an unpredictable risk of leakage to patient privacy. In the

past, when medical institutions shared or released medical data for privacy protection, they would choose to remove some personal identification information, such as name, address, and phone number. However, attackers can still obtain some insensitive user information through other means. For example, this information is used to correspond with the user's disease diagnosis data to get the patient's privacy about the disease. This attack is also called a link attack [1].

Table 1 is a medical datasheet. The hospital did not explicitly give the patient's name when released. However, assuming that the attacker obtains the voting table of the voters in the user's jurisdiction, as shown in Table 2 on the network, the attacker can link the common attributes of the two tables, such as zip code (430056), to infer the patient's status, name (Kevin), and the disease "overweight." If an evil attacker sells this information to a weight loss center, it will directly leak the patient's (Kevin) private information.

This research paper is organized as follows: Section 1 describes the introduction. The privacy protection based on the principle of anonymity is described in Section 2. Section 3 describes the information entropy I-diversity model. The personalized information entropy is described in Section 4. Section 5 describes experimental results and analysis. The conclusion is described in Section 6.

2. Privacy Protection Based on the Principle of Anonymity

Privacy protection based on the principle of anonymity mainly processes the relevant attributes in the data table through technical means, such as data generalization and data suppression [2, 3], before data are released or shared and do not release or restrict the release of specific data. In this way, personal identification information loses association with sensitive attributes and achieves the purpose of privacy protection [4, 5]. Researchers are also proposing various security protocols for keeping the information confidential and secure which is shared among the users and servers in wireless network using several IoT devices [6].

2.1. k-Anonymous Thought. Among the many privacy protection methods based on the anonymity principle, k-anonymity has become a necessary technical means for privacy protection in data release because it protects private information while ensuring data availability [7, 8]. The core idea is to publish low-precision data through generalization and concealment techniques. Each record has at least the same quasi identifier attribute value as the other $k - 1$ data in the data table, reducing privacy caused by link attacks leakage [9].

Definition 1. k-anonymity: given a data table $U \{B_1, B_2, \dots, B_n\}$, the quasi-identifier of U is RI, if and only if each value sequence in $U[RI]$ in $U[RI]$ appears at least k times. It is said that table U satisfies k-anonymity on QI. $T[QI]$ represents the projection of the tuple of table U on RI.

TABLE 1: The sheet of medical data.

Age	Sex	Zip	Disease
22	F	103658	Short breath
24	M	158083	Influenza
25	M	158086	Fever
27	F	350186	Insomnia
29	M	213045	Influenza
33	F	120654	Hepatitis
35	M	430056	Obesity
36	M	131548	Emphysema

TABLE 2: The list of voter poll.

Name	Age	Sex	Zip	Party
Tim	25	M	132635	Member
Linda	28	F	151346	N/a
Kevin	35	M	430056	Member
Mary	37	F	350186	Member

2.2. k-Anonymity. The essence of k-anonymity requires that every record in the dataset has the same projection on the quasi identifier with at least $k - 1$ records. Therefore, the probability that the form where an individual is located is determined that does not exceed $1/k$. Generalization [10] is a technical way to achieve anonymized privacy protection. Its essence is to replace specific values with generalized values or intervals and increase attackers' difficulty in obtaining individual private information by reducing data accuracy. Table 3 is an anonymous medical data table satisfying $k = 2$ after generalization.

2.3. k-Anonymity Disadvantages. Although the k-anonymity algorithm improves the security of published information, it loses part of the data availability due to the need to generalize and conceal specific attributes of the data table. At the same time, the k-anonymity algorithm has the disadvantage of inaccurate query results during the calculation process, especially in the scenario where users are scarce. In addition, it will generate a larger anonymous area, thereby increasing the communication overhead.

3. Information Entropy I-Diversity Model

The generalized data table satisfies k-anonymity, which ensures that a particular user is in the set of k individuals of the same category so that individual users with the same quasi identifier are indistinguishable, thereby achieving a certain degree of anonymity protection. However, we suppose that k tuples in the same equivalence class have the same value on the sensitive attributes. Because it preserves private information while preserving availability of data, k-anonymity has now become a necessary technical technique for privacy protection in data release. The main concept is to use generalization and concealing methods to release low-precision data. Every record in the datasheet seems to have at least the very same quasi-identifier parameter value as the other $k - 1$ data, limiting privacy risks caused by link threats. The advantages of this algorithm are

TABLE 3: The meeting 2- anonymous data sheet.

Age	Sex	Zip	Disease
(20, 30]	M	1211**	Pneumonia
(20, 30]	M	1211**	Influenza
[30, 40]	F	1315**	Diabetes
[30, 40]	F	1315**	Diabetes
(40, 50]	*	1526**	Heart disease
(40, 50]	*	1526**	Hypertension

* signifies the approximate value.

that the k-anonymity technique increases the privacy of released information; it reduces the data accessibility by requiring the generalizations and concealment of certain data table properties. In that case, the individual user records will be attacked by homogeneity and cause attribute leakage, such as the second equivalence class in Table 3.

To solve the privacy leakage problem caused by homogeneity attacks, literature [11] proposes the l -diversity model, which requires each equivalence class to contain at least l well-presented sensitivities. Attribute value, taking into account the constraints on sensitive attributes. If each equivalence class in the data table has l different sensitive attribute values, then the data table is said to satisfy the l -diversity rule. Literature [7] also gives an information entropy l -diversity rule.

Definition 2. Entropy l -diversity: given a data table $U\{B_1, B_2, \dots, B_n\}$, the quasi-identifier attribute is $RI\{B_i, B_{i+1}, \dots, B_j\}$, the sensitive attribute is TA, and $T = \{T_i, T_{i+1}, \dots, T_j\}$ is the sensitive attribute value. Table U satisfies k-anonymity, and its equivalence class set is $F = \{F_1, F_2, \dots, F_n\}$, if and only if for each equivalence class $F_i = 1, 2, \dots, n \subseteq F$, all satisfy. In formula (1), it is said that the data table U satisfies the information entropy l -diversity.

$$P(E_i, s) \lg P(E_i, s) \geq \lg(1), \quad (1)$$

$$\sum_{t \in T} Q(F_i, t) \lg Q(F_i, t) \geq \lg(l). \quad (2)$$

Among them, $Q(F_i, t)$ is the frequency of the sensitive attribute value s in the equivalence class F_i ; $\sum_{t \in T} Q(F_i, t) \lg Q(F_i, t)$ is the information entropy of the equivalence class F_i , also known as the information entropy diversity sex, denoted as entropy (F_i). Information entropy reflects the distribution of attributes. The larger the information entropy, the more even the distribution of sensitive attribute values in the equivalence, and the more difficult it is to derive specific individuals. By type

(1) It can be seen that if the equivalence class satisfies the information entropy l -diversity, then the information entropy of the equivalence class is at least $\lg(l)$. Table 4 is an equivalent class in the anonymous data table.

Table 4 shows that the information entropy l -diversity calculation results of an equivalence class are as follows:

$$\frac{4}{5} \lg \frac{4}{5} + \frac{1}{5} \lg \frac{1}{5} \approx 0.217 \approx \lg 1.165. \quad (3)$$

From the results, for the equivalence class, the diversity of information entropy is $\lg 1.65$, and the value of parameter l

TABLE 4: The meeting 5- anonymous equivalence class.

Age	Sex	Zip	Disease
(20, 30]	M	1236**	Influenza
(20, 30]	M	1236**	Influenza
(20, 30]	M	1236**	Influenza
(20, 30]	M	1236**	Diabetes
(20, 30]	M	1236**	Influenza

** signifies the approximate value.

cannot exceed 1.65. Only one attribute can be taken, considering the definition of l -diversity. There is at least one different sensitive attribute value in the price category. For the published datasheet, this conclusion is obviously of little significance.

And 4 of the sensitive attributes in the equivalence class are “flu.” For many patients, this is not a sensitive attribute. We suppose that the equivalence class contains four sensitive details: “tuberculosis” sensitive points. Assuming that the attacker knows that someone is in the equivalence class, the attacker is confident to speculate that the person has the characteristics of “emphysema” disease tendency, which is unacceptable for the patient [12]. Medical information contains many nonsensitive attribute values such as “flu” or “fever,” and the disclosure of these attribute values will not infringe on individual privacy. Therefore, the information entropy l -diversity model does not distinguish between sensitive attribute values and cannot reflect the risk of privacy leakage in this case. This paper proposes a personalized information entropy l -diversity model to protect users’ medical data privacy.

4. Personalized Information Entropy l -Diversity Model

4.1. Personalized Information Entropy. Diversity model definition has given the deficiencies of the information entropy l -diversity model; on the one hand, it is necessary to increase the information entropy value of the equivalence class. On the other hand, it is essential to distinguish sensitive attribute values and reduce information leakage with solid and sharp attributes.

Therefore, the sensitive attribute value can be divided into a robust and sensitive value SV (sensitive value) and a weak sensitive value DV (do not care value), modifying the information entropy l -diversity rule to obtain a new personalized information entropy l -diversity rule.

Definition 3. Personalized information entropy l -diversity: given a data table $U\{B_1, B_2, \dots, B_n\}$, $RI\{B_i, B_{i+1}, \dots, B_j\}$ is the quasi identifier of U , and TA is the sensitive attribute, $T = \{T_i, T_{i+1}, \dots, T_j\}$ is the set of sensitive attribute values, SV represents the strong sensitivity value, DV is the weak sensitivity value, $|SV|$ is the number of strong sensitivity values, $|DV|$ is the weak sensitivity value Number. Table U satisfies k-anonymity, and its equivalence class set is $F = \{F_1, F_2, \dots, F_n\}$, if and only if for each equivalence class $F_i = 1, 2, \dots, n \subseteq F$, all satisfy. In formula (2), it is said that the data table U satisfies the individualized information entropy l -diversity.

$$\frac{|DV| + |SV|}{|SV|} 10^{\sum_{SV \in S} Q(F_i, SV) \lg Q(F_i, SV)}. \quad (4)$$

Among them, $Q(F_i, SV)$ is the frequency of strong and sensitive attribute values appearing in the equivalence class. Rate $|DV| + |SV|/|SV| 10^{\sum_{SV \in S} Q(F_i, SV) \lg Q(F_i, SV)}$ is the letter of personalized equivalence class diversity of information entropy.

It can be seen from formula (2) that it is necessary to calculate the frequency $Q(F_i, SV)$ of the susceptible attribute value in the equivalence class instead of calculating the weakly sensitive attribute value that will reduce the value of $Q(F_i, SV) \lg Q(F_i, SV)$ frequency of occurrence. Formula (2) is used to calculate the information entropy of the equivalence class in Table 4. There are $DV = \{\text{flu}\}$, $SV = \{\text{emphysema}\}$, $|DV| = 1$, $|SV| = 1$, and SV appears in the equivalence class. The probability is $1/5$, and then the improved diversity of information entropy is

$$1 = 1 + 1 \times 10^{-15} \lg 15 \approx 2.2828, \quad (5)$$

$$m = \frac{1+1}{1} * 10^{-\frac{1}{5}} \lg \frac{1}{5} \approx 2.2828.$$

According to the calculation result, the value of l does not exceed 2.2828. Then, l is 2, and the equivalence class satisfies 2-diversity. Personalized information entropy l -diversity is compared with information entropy l -diversity, it improves the information entropy of the equivalence class and reduces the correspondence between the private information and the identity information derived from the link in the equivalence class.

4.2. Information Loss Measurement. The anonymous privacy protection model based on k -anonymity and its improvement will inevitably produce information loss while protecting private information, which will affect data accuracy [13]. This is the anonymization cost. The anonymity cost is generated when the original data are generalized and suppressed in preprocessing operations. The anonymity cost measurement is an indicator to measure the information loss after the data is anonymized, and it can also judge the optimization degree of the anonymized dataset. The smaller the information loss, the greater the data accuracy, and higher is the data availability, and vice versa. Therefore, in the process of anonymization operation, the cost of anonymity should be reduced as much as possible.

This article adopts the method based on the generalization level to measure the anonymity cost and to use this method to measure the anonymity cost. It is necessary to construct the domain generalization level of the attribute. The amount of information in each layer in the domain generalization hierarchy is different. Generally, for the same quality, data at a higher level of generalization have less information than data at a lower level. It is calculated as follows:

$$Qre(ST) = 1 - \frac{\sum_{i=1}^N \sum_{j=1}^{N_A} h / |DGH_{B_i}|}{N * N_A}. \quad (6)$$

Among them, Qre represents the accuracy of the data, which is the original data table; ST is the generalized data table; N means the number of attributes in the data table; N_A is the number of records in the dataset; $|DGH_{B_i}|$ is the generalized hierarchical structure of the attribute B_i Height; h represents the height of the attribute B_i in the generalization hierarchy.

Definition 4. Domain generalization hierarchy: let A be the attribute of data table U ; there is a function f_h : $h = 0, 1, \dots, n-1$ such that $B_0 \xrightarrow{h_0} B_1 \xrightarrow{h_1} \dots \xrightarrow{h_{n-1}} B_n$, and $B = B_0$, $|B_n| = 1$; then, the generalization domain layer of attribute A on f_h : $h = 0, 1, \dots, n-1$ can be expressed as $\cup_{h=0}^n h$, denoted as $|DGH_B|$.

$\{A_0, A_1, A_2, A_3\}$ shows the bottom-up generalization process of Zip attributes, and each layer represents a generalization domain of the fact. As the DG keeps going up, the generalization degree of the quality gets higher and higher until it finally reaches the inhibited state. The generalization process is described as follows:

$$\begin{aligned} A_3 &= \{*****\} \\ &\uparrow \\ A_2 &= \{021**\} \\ &\uparrow Z \\ A_1 &= \{0213^*, 0214^*\} \\ &\uparrow \\ A_0 &= \{02138, 02139, 02141, 02142\} \end{aligned}$$

5. Experimental Results and Analysis

This experiment uses the incognito algorithm proposed in [14] to complete the anonymous operation process. The basic idea of the Incognito algorithm is to use global recoding technology to perform generalization operations on the original dataset in a bottom-up breadth-first manner, and at the same time, perform necessary pruning and iterative functions on the generalization graph to make the original dataset gradually optimize, to achieve anonymity effect. The main premise behind the incognito algorithm is to use worldwide recoding techniques to accomplish bottom-up breadth-first generalization operations on the entire dataset, while also performing required pruning and repetitive features on the generalization graph to progressively enhance the actual information set and achieve secrecy. With an adjustable sequence reduction limit, the incognito algorithm constructs a set of all potential k -anonymous full-domain extensions of T . The approach checks monosubsets of the quasi-identifier first, and iterates, testing k -anonymity with regard to considerably larger subsets, in a similar fashion of the subset characteristic. The program in this paper mainly considers the algorithm execution time and the information loss of the data table.

5.1. Experimental Data and Experimental Environment.

The dataset used in the experiment is the Adult database in UCI [15], which is the most commonly used data source for k-anonymity research. The database has 32,206 pieces of data with a size of 5.5 MB, and the dataset contains a total of 15 attributes. We select eight attributes as the attribute set of quasi identifiers and select the Disease attribute as the sensitive attribute. Table 5 describes the structure of the experimental dataset.

The experiment uses MySQL 5.5 to store data; the algorithm is implemented in Java; the experiment running environment is a 3.3 GHz Intel® Corei5 processor with 4 GB RAM.

We select Disease as the experimental sensitive attribute. The Disease attribute contains ten values, randomly generating disease types and using sensitivity weights to measure sensitivity [16]. The larger the value, the higher the sensitivity, as shown in Table 6. In the experiment, diseases with a sensitivity weight lower than 0.5 are set as weakly sensitive attributes.

5.2. Time Complexity Analysis. The solution in this paper first needs to calculate the personalized information entropy of the solid and sensitive value |SV| and the weak sensitive

value |DV| diversity $|DV| + |SV|/|SV|10^{\sum_{SV \in S} Q(F_i, SV) \lg Q(F_i, SV)}$, [1] so the computational cost is linear. The time complexity is $P(n)$. Secondly, the information loss metric needs to be calculated: $Qre(ST) = 1 - \sum_{i=1}^N \sum_{j=1}^{N_A} h(|DGH_{Bi}|/N * N_A)$, and there are two cumulative multiplication operations in the calculation process, so the calculation overhead at this stage is $P(n^2)$. Finally, the attributes need to be processed at the domain generalization level. The computational cost of each generalization is linear, so the computational complexity is $P(n)$.

5.3. Execution Time Analysis. It can be seen that Figure 1 with Table 7 shows that as the number of RIs increases, the execution time of the three models will increase. This is because as the RI value increases, the equivalence class also rises.

More quasi-identifier attributes must be recorded, which requires more generalization times. This process also requires the algorithm to execute more cycles to increase the execution time.

At the same time, it can be seen that as the QI value increases, the execution time of the proposed scheme is shorter than that of the information entropy l -diversity model. This is because it takes longer for the information entropy l -diversity model to judge the strong and weakly sensitive attributes in the equivalence class.

Figure 2 with Table 8 describes the change of data accuracy with the value of k or l in the three anonymous models when QI records increase from 0 to 1,000. The abscissa is the number of documents, and the ordinate is the time to execute different algorithms. It can be seen that the solution in this paper improves the data accuracy while reducing the execution time of the algorithm.

TABLE 5: The structure of Adult dataset.

Attributes	Difference value	Weights
Age	74	4
Gender	2	2
Race	5	2
Educate	16	4
Employer	7	3
Country of citizenship	41	3
Profession	14	2
Disease	10	—

TABLE 6: The weight of disease.

Numbering	Disease	Weights
1	Influenza	0.11
2	Obesity	0.12
3	Fever	0.13
4	Angrily kick	0.31
5	Insomnia	0.41
6	Chest pain	0.42
7	Hepatitis	0.51
8	Myocarditis	0.91
9	Tuberculosis	0.92
10	Depression	0.93

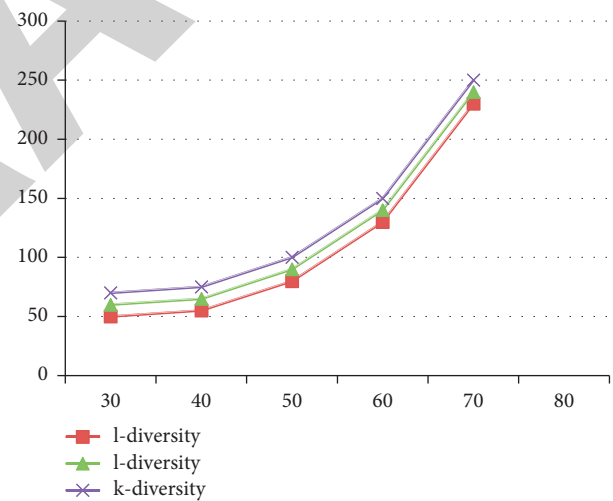


FIGURE 1: Comparison of execution time with the number of quasi identifiers.

TABLE 7: Comparison of execution time with the number of quasi identifiers.

S. No.	l -diversity	l -diversity	k -diversity
30	50	60	70
40	55	65	75
50	80	90	100
60	130	140	150
70	230	240	250

5.4. Precision Analysis. Figure 3 with Table 9 describes the change of data accuracy with the value of k or l in the three anonymous models when QI records increase from 1,000 to 3,500. The abscissa is the number of documents, and the

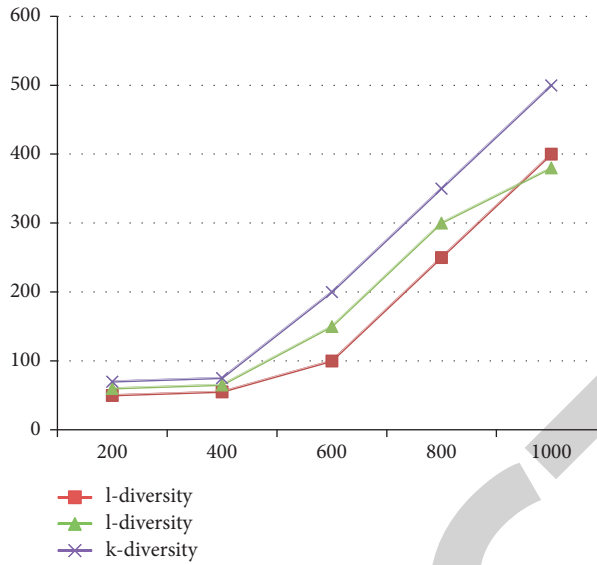


FIGURE 2: Comparison of execution time with the number of records.

TABLE 8: Comparison of execution time with the number of records.

S. No.	<i>l</i> -diversity	<i>l</i> -diversity	<i>k</i> -diversity
200	50	60	70
400	55	65	75
600	100	150	200
800	250	300	350
1000	400	380	500

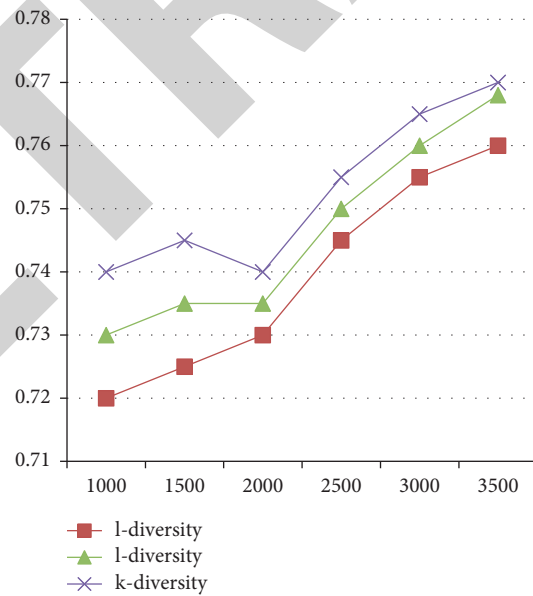
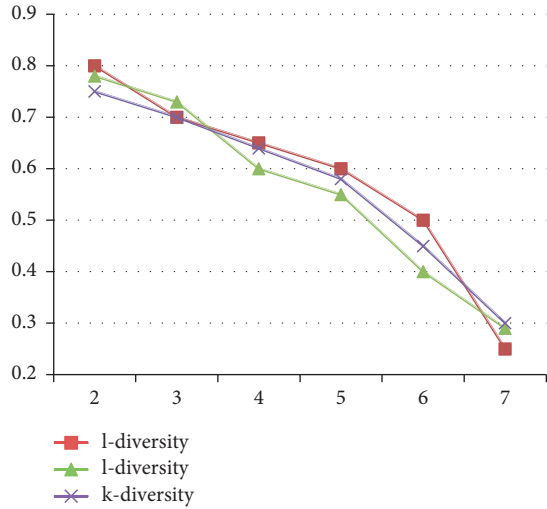


FIGURE 3: Comparison of data accuracy with the number of records.

TABLE 9: Comparison of data accuracy with the number of records.

S. No.	<i>l</i> -diversity	<i>l</i> -diversity	<i>k</i> -diversity
1000	0.72	0.73	0.74
1500	0.725	0.735	0.745
2000	0.73	0.735	0.74
2500	0.745	0.75	0.755
3000	0.755	0.76	0.765
3500	0.76	0.768	0.77

FIGURE 4: Comparison of data accuracy with l , k value.TABLE 10: Comparison of data accuracy with l , k value.

S. No.	l -diversity	l -diversity	k -diversity
2	0.8	0.78	0.75
3	0.7	0.73	0.7
4	0.65	0.6	0.64
5	0.6	0.55	0.58
6	0.5	0.4	0.45
7	0.25	0.29	0.3

ordinate is the data accuracy of the unknown dataset. It can be found that as the number of records increases, the accuracy of this solution is higher than other solutions.

5.5. Information Loss Analysis. Figure 4 with Table 10 describes the change of data accuracy with the value of k or l in the three anonymous models when the number of QIs is 0–8. The abscissa is the value of k and l , and the ordinate is the data accuracy of the anonymous dataset.

Figure 3 describes as the values of k and l increase, the accuracy of the data shows a downward trend. As the importance of k and l increases, the number of tuples that need to be generalized in the equivalence class increases. The higher the generalization level, the greater is the information loss, and the data accuracy will decrease. Under the same circumstances, the information loss of personalized information entropy l -diversity is higher than that of information entropy l -diversity. This is because customized information entropy l -diversity has more substantial constraints on anonymity than information entropy l -diversity. Higher-level generalization needs to be aligned with the identifier, so the information loss is relatively significant.

6. Conclusion

Aiming at the problem that the information entropy l -diversity model does not distinguish between strong and weak sensitive attributes, this paper proposes a personalized information entropy l -diversity model and conducts

experiments. The privacy of the proposed model is better than other models. Because it secures personal data while preserving data availability, k -anonymity, based on the incognito principle, has become an essential technical technique for privacy protection in data release. The basic concept is to use generalizations and concealing methods to publish data with low accuracy. Each record in the datasheet has at least the same quasi-identifier attribute value as the other $k - 1$ data in the database, decreasing security leaks caused by connection threats. The drawback of the proposed algorithm is that erroneous query results throughout the calculating process, especially when users are scarce. It will also create a broader anonymous area, which will increase the communication overhead. The experimental results show that the performance of this solution in terms of execution time and data accuracy is better than the information entropy l -diversity model and k -anonymity model, and it has better privacy. It can be used in mobile medical systems to protect medical users. Private data will not be leaked. The scope of this research is that the data analysis and trial findings reveal that this strategy is more effective than previous alternatives in terms of reducing execution time while boosting data accuracy and service quality.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] B. K. Tripathy, A. Maity, B. Ranajit, and D. Chowdhuri, "A fast p -sensitive l -diversity Anonymisation algorithm," in *Proceedings of the 2011 IEEE Recent Advances in Intelligent Computational Systems*, pp. 741–744, Trivandrum, India, September 2011.
- [2] H. Zhu, S. Tian, and K. Lu, "Privacy-preserving data publication with features of independent l -diversity," *The Computer Journal*, vol. 58, no. 4, pp. 549–571, 2015.
- [3] J. Han Jianmin, T. Cen Tingting, and J. Yu Juan, "An l -MDAV microaggregation algorithm for sensitive attribute l -diversity," in *Proceedings of the 2008 Twenty Seventh Chinese Control Conference*, pp. 713–718, Kunming, China, July 2008.
- [4] M. A. Enam, S. Sakib, and M. S. Rahman, "An algorithm for l -diversity clustering of a point-set," in *Proceedings of the 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 1–6, Cox'sBazar, Bangladesh, February 2019.
- [5] M. Testard, J. C. Nivelet, T. Matos, and G. Levannier, "Tight approximation of bit error probability for L -diversity non-coherent M -ary FSK frequency hopping system with binary convolutional code and soft Viterbi decoder: diversity, bit interleaver size and Reed-Solomon outer code effects analysis on receiver performance for $M=8$," *MILCOM 97 MILCOM 97 Proceedings*, vol. 1, pp. 313–317, 1997.
- [6] J. Jianmin Han, H. Huiqun Yu, and J. Yu, "An improved l -diversity model for numerical sensitive attributes," in *Proceedings of the 2008 Third International Conference on*

Retraction

Retracted: Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System

Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 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] Q. Li, X. Chai, C. Zhang, X. Wang, and W. Ma, "Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8936103, 10 pages, 2022.

Research Article

Prediction Model of Ischemic Stroke Recurrence Using PSO-LSTM in Mobile Medical Monitoring System

Qingjiang Li ¹, Xuejiao Chai ², Chunqing Zhang ³, Xinjia Wang ⁴, and Wenhui Ma ³

¹School of Medical Technology, Qiqihar Medical University, Qiqihar, Heilongjiang 161000, China

²School of Public Health, Qiqihar Medical University, Qiqihar, Heilongjiang 161000, China

³Clinical Teaching Center, Qiqihar Medical University, Qiqihar, Heilongjiang 161000, China

⁴School of General Practice and Continuing, Qiqihar Medical University, Qiqihar, Heilongjiang 161000, China

Correspondence should be addressed to Wenhui Ma; mawenhui@qmu.edu.cn

Received 9 December 2021; Revised 13 January 2022; Accepted 25 January 2022; Published 24 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Qingjiang Li 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.

Aiming at the problems of low prediction accuracy and low sensitivity of traditional ischemic stroke recurrence prediction methods, which limits its application range, by introducing an adaptive particle swarm optimization (PSO) algorithm into the Long and Short-Term Memory (LSTM) model, a prediction model of ischemic stroke recurrence using deep learning in mobile medical monitoring system is proposed. First, based on the clustering idea, the particles are divided into local optimal particles and ordinary particles according to the characteristic information and distribution of different particles. By updating the particles with different strategies, the diversity of the population is improved and the problem of local optimal solution is eliminated. Then, by introducing the adaptive PSO algorithm into the LSTM, the PSO-LSTM prediction model is constructed. The optimal super parameters of the model are determined quickly and accurately, and the model is trained combined with the patient's clinical data. Finally, by using SMOTE method to process the original data, the imbalance of positive and negative sample data is eliminated. Under the same conditions, the proposed PSO-LSTM prediction model is compared with two traditional LSTM models. The results show that the prediction accuracy of PSO-LSTM model is 92.0%, which is better than two comparison models. The effective prediction of ischemic stroke recurrence is realized.

1. Introduction

Researches show that the death caused by cerebrovascular diseases ranks first and second among the causes of death in China. The incidence rate of ischemic stroke accounts for more than 75% of cerebrovascular diseases [1]. In recent years, advanced technologies such as deep learning methods and mobile medical monitoring have been developed and widely applied in medicine. In response to the high incidence rate, high mortality rate, and high recurrence rate of ischemic stroke, the coping strategies gradually changed from treatment-based to prediction-based [2, 3]. Early recurrence prediction and effective response measures for cerebrovascular diseases are the key factors for the prevention and treatment of stroke recurrence, which is also the focus of research in the medical industry [4–6].

At present, more than 94% of ischemic stroke is caused by specific and controllable factors such as lifestyle, hypertension, and aging [7, 8]. Therefore, the incidence rate and mortality rate of ischemic stroke patients can be significantly decreased according to the main causes and current physiological state of ischemic stroke [9, 10]. In [11], elderly patients, who had early transient ischemic attack and had a poor prognosis, were divided into recurrence group and nonrecurrence group. The criterion for judging the poor prognosis was whether there was recurrence one month after the onset. According to logistic regression analysis, the clinical characteristics of the two groups were compared, and a guiding method that can predict the recurrence of ischemic stroke was proposed. However, this method only analyzes the characteristics after the early onset of the disease, which has some limitations. Reference [12] took patients with

depression and cerebrovascular diseases as the research object and used the diagnosis of depression as the index date to track and record the recurrence of cerebrovascular diseases. The results showed that the risk of cerebrovascular diseases recurrence mainly depends on age and physical health but had little relationship with depression and psychotropic drugs. This method mainly analyzed the effect of depression on the recurrence of cerebrovascular diseases but did not give the specific factors affecting the recurrence of cerebrovascular diseases. The results of [13] showed that asymptomatic cerebral infarction is an independent predictive factor of clinical cerebrovascular events and can be used as a prediction index of cerebrovascular events recurrence to predict ischemic stroke. However, this method only proves the predictive effect of asymptomatic cerebral infarction on cerebrovascular diseases and cannot be applied to most patients. Reference [14] obtained relevant feature information by learning the clinical characteristics of stroke patients using the bidirectional LSTM method and enhanced the utilization efficiency of index information according to the attention mechanism. The authors proposed a risk prediction model of ischemic stroke based on the bidirectional LSTM (BiLSTM) network and attention mechanism. However, this method does not classify the clinical characteristics of patients in detail, and the prediction accuracy is not optimal. Reference [15] obtained the patient's characteristic information by learning the electronic medical record of patients with cardiovascular and cerebrovascular diseases using the Recurrent Neural Network (RNN), proposed a cardiovascular and cerebrovascular disease risk prediction model based on electronic medical record data mining, and improved the prediction accuracy of the model by fusing various types of clinical data. However, this method does not consider the difference of pathogenesis in different patients, and there are some defects. Reference [16] designed the embolization mechanism of Shuxuetong injection in prevention of acute ischemic stroke recurrence and carried out relevant experiments. The experimental results provided an effective basis for the effectiveness and safety of Shuxuetong injection in reducing stroke recurrence in patients with ischemic stroke. However, this method only studied and verified the preventive effect of Shuxuetong injection and did not put forward a substantive prediction basis. Reference [17] scored patients with acute ischemic stroke based on Cox regression analysis, divided the risk of ischemic stroke recurrence under different scores by establishing a scoring mechanism, and realized the prediction of ischemic stroke recurrence by calibrating and distinguishing the scores. However, the score of this method has not been fully verified when applied to external data sets and has poor prediction accuracy.

Based on the above analysis, aiming at the problems of small application scope and low accuracy of most existing ischemic stroke recurrence prediction methods, an ischemic stroke recurrence prediction model using deep learning in mobile medical monitoring system is proposed. The adaptive PSO algorithm can improve the population diversity and solve the problem of falling into the local optimal solution. It can effectively predict the recurrence of ischemic stroke by

introducing PSO into LSTM and combining the patient's clinical data.

2. Model Establishment

2.1. Problem Description. The recurrence prediction of ischemic stroke can be mathematically described as the feature mapping relationship between the prediction index of the current time or a future time t and the corresponding label.

For the historical data of h patients, set the prediction index data set as $A = \{a_1, a_2, a_3, \dots, a_h\}$ and the label set corresponding to the data set as $B = \{b_1, b_2, b_3, \dots, b_h\}$. On the basis of constructing the prediction model of ischemic stroke recurrence, the feature mapping relationship between the prediction index data set A and the corresponding label set B is obtained by training the model. According to the obtained feature mapping relationship between the prediction data and the corresponding label, for the newly emerging patients outside the label set, based on the patient's prediction data, the optimal risk prediction result $B' = \{b'_1, b'_2, b'_3, \dots, b'_h\}$ is obtained through the ischemic stroke recurrence prediction model.

2.2. LSTM Neural Network Prediction Model of Adaptive PSO

2.2.1. PSO Algorithm of Adaptive Learning Strategy. The establishment of the prediction model is based on the adaptive PSO algorithm, and the traditional adaptive PSO algorithm is improved by introducing LSTM to construct the ischemic stroke recurrence prediction model. It can use the PSO algorithm of adaptive learning strategy to match the data features of ischemic stroke patients with the topology of LSTM neural network, so as to achieve higher prediction performance.

In order to improve the diversity of the initial population as much as possible, the clustering idea is the basic idea in the process of model construction. First, based on the feature information of different particles in the particle swarm and their respective distribution, the whole particle swarm is divided into multiple subgroups with different features. Then, different learning strategies are adopted for different subgroups to improve the diversity of the whole population [18].

The process of dividing particle swarm is based on simplified PSO algorithm and simplified PSO algorithm with extreme value disturbance [19]. The algorithms can automatically obtain the cluster center of the sample data set and can realize high-performance clustering and fast search for any shape of data. The basic basis is that the cluster center contains two basic features surrounded by points with low local density, and the information of the two features is far away from the points with high local density. The basic principle is as follows:

If there is a population E in the search space of W dimension, and the population is composed of e particles, the population E can be expressed as $E = \{l_k\}_{k=1}^e$, where l_k represents the k -th particle in the population E , which contains W dimensions and can be expressed as $l_k = \{l_{k1}, l_{k2}, l_{k3}, \dots, l_{kW}\}$. Two variables are given from the

w -th dimension of the k -th particle: local density and distance. The local density of the particle is defined as ρ_{kw} , and its expression is as formula (1):

$$\rho_{kw} = \sum_{k \neq j} \exp \left[-\frac{D_{kj}^2}{D_T} \right], \quad (1)$$

where D_{kj} represents the Euclidean distance between the k -th particle and the j -th particle. D_T represents the truncation distance.

The distance between the k -th particle and the j -th particle with higher local density is defined as σ_{kw} , and its expression is as follows:

$$\sigma_{kw} = \min_{j: \rho_{jw} > \rho_{kw}} \{D_{kj}\}. \quad (2)$$

For the sample data with the largest local density ρ_{kw} , the value of σ_{kw} is $\max\{D_{kj}\}$.

As can be seen from formula (2), if the density of l_{kw} is the maximum local density, σ_{kw} will be much greater than the distance σ of its nearest particle. Therefore, the center of each divided subgroup is generally some particles with very large σ , and the local density ρ of these particles is also relatively large. When selecting the cluster centers for different subgroups, particles with relatively large distance σ and local density ρ can be selected. For the particles other than the cluster center, according to the l_{jw} of the particle, the particle can be divided into subgroups where the sample with local density greater than l_{jw} and closest to l_{jw} is located.

After the overall division of particle swarm, according to the division results of subgroups, the particles in each subgroup are reclassified into two categories: locally optimal particles and ordinary particles, and they are iteratively updated by different update methods to achieve the purpose of increasing population diversity.

Ordinary particles expand the local search ability under the guidance of the optimal particle and update iteratively based on

$$l_{kw} = \alpha \cdot l_{kw} + \beta_1 R_{1w} (L_{POkw} - l_{kw}) + \beta_2 R_{2w} (I_{POkw} - l_{kw}). \quad (3)$$

In formula (3), α represents the inertia weight coefficient. β_1 and β_2 represent learning factors. R_{1w} and R_{2w} represent random numbers which obey uniform distribution in the interval $[0,1]$. L_{POkw} represents the optimal position information of the w -th dimension of the k -th particle. I_{POkw} represents the optimal location information in the i -th subgroup.

For locally optimal particles, in order to strengthen the information interaction between different subgroups, they are generally updated by collecting the information of different subgroups. The update process is shown in formula (4).

$$l_{kw} = \alpha \cdot l_{kw} + \beta_1 R_{1w} (L_{POkw} - l_{kw}) + \beta_2 R_{2w} \left(\frac{1}{I} \sum_{i=1}^I I_{POkw} - l_{kw} \right). \quad (4)$$

In formula (4), I represents the total number of subgroups.

From the above analysis, it can be seen that, in a certain subgroup, the main function of local optimal particles is to guide the search direction of the whole subgroup. They can not only guide other ordinary particles to learn, but also undertake the task of information exchange between different subgroups. If the local optimal particle updates with the same update strategy as other ordinary particles, the local optimal particle will lose the ability to interact with other subgroups, and its search direction is likely to deviate from the optimal search direction. At this time, the subgroup will fall into the trap of local optimal solution. Therefore, the local optimal particle needs to break through the constraints of the subgroup and interact with other subgroups in the process of updating, so as to ensure the correctness of the search direction by obtaining effective information from other subgroups, as shown in (4). In this way, information sharing among different subgroups can improve population diversity.

2.2.2. PSO-LSTM Prediction Model. The clinical data of patients with ischemic stroke can be regarded as a time series. There are many factors inducing their disease recurrence, and all the factors are very complex, uncertain, nonlinear, and unstable [20]. In order to accurately predict the factors inducing disease recurrence to the greatest extent, a prediction model for ischemic stroke recurrence is constructed based on LSTM according to the common features of time series.

The network structure of LSTM is mainly affected by some parameters in the model [21]. In order to make LSTM more suitable for the prediction of ischemic stroke recurrence, a new LSTM ischemic stroke recurrence prediction model based on PSO algorithm is constructed by fusing and optimizing the adaptive PSO algorithm and LSTM.

It can be seen from the previous analysis that the adaptive PSO algorithm has many unique advantages, such as simple algorithm design, fast calculation speed, and convergence speed, and makes up for the defect that it is easy for the ordinary PSO algorithm to fall into the trap of local optimal solution. The adaptive PSO algorithm greatly improves the performance in finding the optimal solution [22]. PSO algorithm combined with LSTM can quickly and accurately determine the optimal super parameters and finally realize the effective prediction of ischemic stroke recurrence.

The basic structure of LSTM ischemic stroke recurrence prediction model based on PSO algorithm is shown in Figure 1 below.

The model construction process mainly includes the following steps:

- (1) The adaptive PSO algorithm is used to optimize some parameters of LSTM, including time window parameters, the number of hidden layer units, and batch processing parameters. The position information of different particles in particle swarm optimization is initialized based on the value range of super parameters.

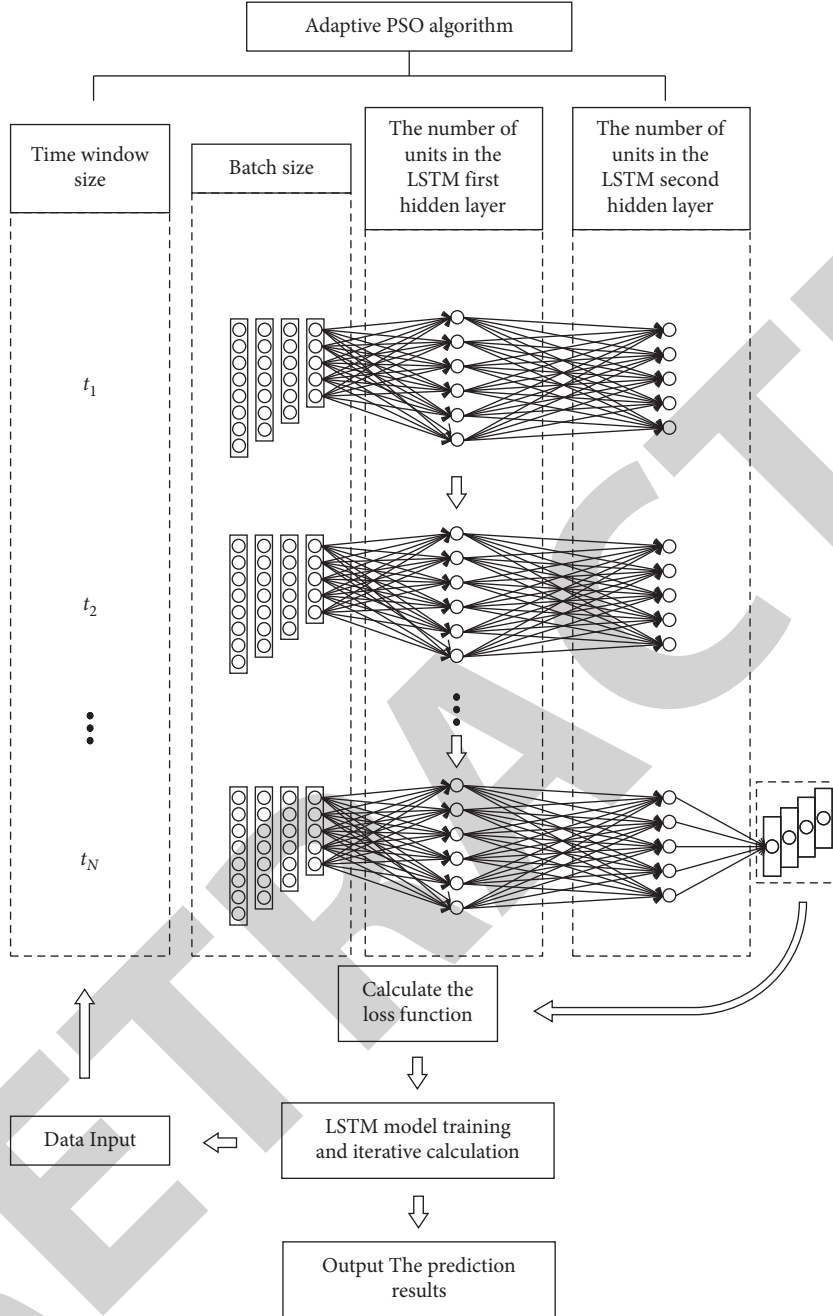


FIGURE 1: The structure of the LSTM ischemic stroke recurrence prediction model based on adaptive PSO.

- (2) Based on local density of the particle ρ_{kw} and the distance σ_{kw} between this particle and the particle with higher local density calculated by formulas (1) and (2), finally, the adaptive partition of particle swarm is realized, and several subgroups are obtained.
- (3) Based on the super parameter value corresponding to the particle position information, the LSTM is constructed. The training data are used to learn and train the constructed model. Then use the validation data to predict and verify the trained LSTM.

- (4) Construct fitness values for particles. The definition of fitness function is shown in formula (5), which is the average absolute percentage error of the model on the validation data set.

$$F = \frac{1}{S_v} \sum_{i=1}^{S_v} \frac{|n'_i - n_i|}{n_i} \quad (5)$$

In formula (5), S_v represents the total amount of data in the validation data set. n'_i represents the predicted value of the i -th validation data. n_i represents the true value of the i -th validation data.

- (5) The fitness values of particles in all different subgroups are calculated, and based on their fitness values, these particles are divided into ordinary particles, subgroup optimal particles, and global optimal particles. On this basis, for different classes of particles, their position information is calculated and updated with the formulas shown in equations (3) and (4).
- (6) Determine whether to terminate the calculation. If the termination conditions are met, it means that the global optimal solution of the optimization objective has been obtained. If the termination conditions are not met, the subpopulation should be regrouped according to the updated particle position information, and steps 2–5 should be repeated until the termination conditions are met. The global optimal solution of the super parameter is obtained.
- (7) The LSTM is constructed according to the global optimal solution of the obtained hyperparameters, and the model is trained and predicted based on the clinical data of patients with ischemic stroke.

3. Prediction of Ischemic Stroke Recurrence

3.1. Multifactor Determination. Because the factors affecting the recurrence of ischemic stroke are very complex, it is necessary to screen the relevant indicators. In the medical field, logistic regression analysis is usually widely used to analyze and study the causal relationship between independent variables and dependent variables [23]. Next, logistic regression analysis is used to analyze and determine the multiple factors affecting the recurrence of ischemic stroke. The dependence between independent variables and dependent variables can be characterized by regression coefficients, and the calculation of regression coefficients can be obtained by calculating category probability [24]. The calculation methods of category probability and regression coefficient are shown in formulas (6) and (7), respectively.

$$P(b = c|a) = \frac{\exp[\eta_c + \sum_{m=1}^M \lambda_{cm} a_m]}{1 + \sum_{c=1}^{C-1} (\eta_c + \lambda_{cM} a_M)} \quad (6)$$

$$\ln \left[\frac{P(b = c|a)}{P(b = C|a)} \right] = \eta_c + \sum_{m=1}^M \lambda_{cm} a_m \quad (7)$$

In formulas (6) and (7), a represents the independent variable. c represents the dependent variable. M represents the number of independent variables.

When $M = 1$, there is only one independent variable. At this time, the model carries out single factor analysis to analyze the impact of a single independent variable on the dependent variable. When $M > 1$, the number of independent variables is more than one; the model carries out multifactor analysis to analyze the comprehensive impact on the dependent variables when multiple independent variables change at the same time. Finally, the model comprehensively analyzes the risk factors of ischemic stroke

recurrence by changing the number of independent variables.

3.2. Determination of Input and Output Variables.

According to the problem description, the main problem to be solved in constructing the prediction model of ischemic stroke recurrence is the feature mapping relationship between the prediction index data set A and the corresponding label set B , that is, the feature mapping relationship between input variables and output variables. In order to predict the recurrence of ischemic stroke, firstly, it should be clear whether the input variables and output variables meet the requirements of the model. Given two data sets $D = \{(d_{i(t-1)}, d_{it})\}_{i=1}^h$ and $G = \{g_i \in \{0, 1\}\}$, where i represents the i -th patient, $d_{i(t-1)}$ represents the predictive index data of the i -th patient at the time $t - 1$ and contains a group of multiple incentives affecting the recurrence of ischemic stroke. G represents the diagnostic label of each sample.

The following is an analysis of the input variable (the prediction data set A) and the output variable (the corresponding label set B).

- (1) The input variable (the prediction data set A)

First, the LSTM ischemic stroke recurrence prediction model based on PSO algorithm is calculated through the determined d . For the i -th patient, based on the data index $d_{i(t-1)}$ of the current time $t - 1$, the change d_t' of the continuous value in the data index d_t of the future time t is fitted. The fitting process is shown as follows.

$$d_{it}' = \mu [\text{PSO} - \text{LSTM}(d_{i(t-1)}, d_{it})]. \quad (8)$$

On this basis, the data indexes of time $t - 1$ and time t are combined to form the prediction index data input variable A , as follows.

$$A = \text{concat}[d_{i(t-1)}, d_{it}']. \quad (9)$$

- (2) The output variable (the corresponding label set B)

For the i -th patient, the real label set B of the patient is obtained by converting the obtained sample diagnostic label G into a one-dimensional array, as follows.

$$B = \{b_1, b_2, b_3, \dots, b_h\}. \quad (10)$$

In formula (10), $b_i \in g_i, i = 1, 2, \dots, h$.

After obtaining the prediction index data set A and the corresponding label set B , the proposed prediction model searches the feature mapping relationship between A and B and finally obtains the prediction results of ischemic stroke recurrence risk, as shown in formula (11).

$$B' = \text{soft max}[\text{LSTM}(A, B)]. \quad (11)$$

3.3. Data Acquisition and Preprocessing. Before using LSTM ischemic stroke recurrence prediction model based on PSO

algorithm to predict the recurrence risk, it is necessary to collect patient data and preprocess these data.

Data collection is mainly carried out through the big data management platform for stroke patients. The platform is mainly based on data access and import tools. It takes the medical institutions, sanitary places, healthcare institutions, physical examination centers, and institutions of various hospitals scattered all over the country as the collection objects and collects the source data of different stroke patients. Finally, a uniquely researchable and structured patient case information database about stroke patients was formed [25]. Data collection for patients mainly includes the following aspects: Personal information, past medical history, family history, laboratory data, inpatient diagnosis and treatment data, periodic follow-up data and physical examination data, etc. The data import tool of the platform can provide a compatible heterogeneous data acquisition interface for different types of stroke heterogeneous data sources and can import data from a variety of relational databases from different patients and institutions. In terms of data acquisition strategy, the data import tool of the platform can realize the access and import of full, batch, and real-time data. For offline data, the platform can also import log data files such as HDFS, FTP, and text files. In addition, it can also import streaming data such as Flume and Kafka in real time.

The purpose of data preprocessing is to clean, interpolate missing values, eliminate abnormal data, and standardize data format and other operations for the source data of stroke patients with complex, extensive, and diverse data forms and types. Data preprocessing unifies the data format of stroke patients to improve the overall data quality to a certain extent [26]. On this basis, the influencing factors of ischemic stroke are assigned as the independent variables of modeling. Finally, these data are more suitable for the requirements of model construction.

3.4. Model Training. In order to find the best model parameter, in the process of training the model, the calculation error is calculated by calculating the loss function at each step, and on this basis, the optimizer is used for reverse adjustment and update.

In the binary classification loss problem, the cross-entropy loss function is widely used to calculate the loss. It can reflect the effect of model training by calculating the error between the predicted value and the real label. The cross-entropy loss function is shown in formula (12).

$$F_{Lce} = - \sum_{i=1}^h b_i \log(b'_i). \quad (12)$$

In formula (12), h represents the total number of sample data.

The difference between the probability of ischemic stroke recurrence b'_i and the real label b_i was calculated by maximum likelihood operation.

In the process of reverse adjustment and update using the optimizer, in order to reduce the loss of model training and avoid falling into the trap of local optimal solution,

Adam optimizer is used for reverse calculation to adjust the weight parameters of the network. On this basis, the adaptive learning rate is designed by calculating the first-order moment estimation and second-order moment estimation of the gradient.

The gradient ∇t at time t is calculated based on the loss function F_{Lt} of the target. The calculation process is shown in formula (13).

$$\nabla t \leftarrow \Delta_\gamma F_{Lt}(y_{t-1}). \quad (13)$$

In formula (13), γ represents the update parameter corrected by moment estimation.

According to the gradient ∇t at time t calculated by formula (13), the first-order and second-order moment estimates m_{1t} and m_{2t} are calculated. The calculation process is shown in formulas (14) and (15), respectively.

$$m_{1t} \leftarrow \xi_1 m_{1t-1} + (1 - \xi_1) \nabla t, \quad (14)$$

$$m_{2t} \leftarrow \xi_2 m_{2t-1} + (1 - \xi_2) \nabla t^2. \quad (15)$$

In formulas (14) and (15), ξ_1 and ξ_2 represent the attenuation index of moment estimation, and their values are $\xi_1 = 0.900$ and $\xi_2 = 0.999$.

4. Experiments and Analysis

4.1. Parameter Setting. The parameters of the model are set before the experiment. The proposed PSO-LSTM model in this paper mainly includes four parts: the input layer, the first LSTM layer, the second LSTM layer, and the output layer. The loss function adopts the cross-entropy loss function, the optimizer adopts the Adam algorithm optimizer, and the construction of the network model is completed based on the Keras framework. The super parameters in LSTM model mainly include time window size, batch size, training times, and the number of neurons in hidden layer.

In order to minimize the error and influence of human factors on the model as much as possible, four super parameters in the LSTM model are set based on the clinical data of actual ischemic stroke patients, which are as follows: the value range of time window size is set to $[1, 20]$, the value range of batch size is set to $[1, 60]$, and the value range of the number of neurons in the hidden layer is set to $[10, 30]$. The training times are mainly determined by the loss of the model. Since the loss function of the model will gradually converge after 500 iterations, the training time of the model is 500. In addition, other relevant parameters are set as follows: the total number of particles in the particle swarm is set to 50, the maximum number of iterations is set to 300, the inertia weight of velocity is set to $\alpha = 0.85$, and the sum of acceleration coefficient is set to 1.8.

4.2. Evaluation Index. In order to effectively measure the accuracy of LSTM ischemic stroke recurrence prediction model based on PSO algorithm, the following five evaluation indexes are used to evaluate the experimental results.

(1) Accuracy: the calculation method is

$$E_a = \frac{S_{TT} + S_{FF}}{S_{TT} + S_{TF} + S_{FT} + S_{FF}}. \quad (16)$$

(2) Sensitivity: the calculation method is

$$E_{se} = \frac{S_{TT}}{S_{FT} + S_{FF}}. \quad (17)$$

(3) Specificity: the calculation method is

$$E_{sp} = \frac{S_{FF}}{S_{FT} + S_{FF}}. \quad (18)$$

(4) Positive prediction rate: it is calculated as follows:

$$E_p = \frac{S_{TT}}{S_{TT} + S_{FT}}. \quad (19)$$

(5) Negative prediction rate: it is calculated as follows:

$$E_n = \frac{S_{FF}}{S_{TF} + S_{FF}}. \quad (20)$$

(6) F1_score: the calculation method is shown in the following equation:

$$E_{F1} = \frac{2E_p \cdot E_{se}}{E_p + E_{se}}. \quad (21)$$

In formulas (15)–(21), S_{TT} represents the number of patients who are actually with recurrent ischemic stroke and are correctly predicted as the patients with recurrent ischemic stroke. S_{TF} indicates the number of patients who are actually with recurrent ischemic stroke but are incorrectly predicted as patients without recurrent ischemic stroke. S_{FT} indicates the number of patients who are not with recurrent ischemic stroke but are incorrectly predicted as patients with recurrent ischemic stroke. S_{FF} indicates the number of patients who are not with recurrent ischemic stroke and are correctly predicted as patients without recurrent ischemic stroke.

4.3. Normalization. During the experiment, the clinical data of all ischemic stroke patients, i.e., the input sample values, were normalized to compress the values within the value range [0, 1] [27]. Next, it takes 2000 total cholesterol concentration sample data as an example for normalization. The numerical changes of sample data before and after normalization are shown in Figure 2.

It can be seen from Figure 2 that, after normalization, the total cholesterol concentration of the sample data is compressed from [0, 18] to [0, 1] on the basis of maintaining the basic characteristics of the original data.

4.4. Experimental Results. In order to eliminate the influence caused by the imbalance of positive and negative sample data in the process of data collection, the collected sample data

are balanced by Synthetic Minority Oversampling Technique (SMOTE). The data processed by SMOTE method not only solves the problem of data imbalance between positive and negative samples but also expands the diversity of data samples to a certain extent. The calculation results of different evaluation indexes before and after SMOTE method are shown in Table 1 below.

It can be seen from Table 1 that, compared with the results of ischemic stroke recurrence prediction using the original data, the accuracy, sensitivity, specificity, positive prediction rate, negative prediction rate, and F1_score of recurrence prediction using the data processed by SMOTE method have improved. The prediction accuracy reached 92%, with a relative increase of 14%. In addition, it can be seen that, before using SMOTE method to process the data, the specificity of the prediction results is 81%, the sensitivity is 69%, and the difference between them is 12%, which shows that the negative sample has more feature information than the positive sample in the data used in the learning process of the model, and there are more omissions in predicting the recurrence of ischemic stroke patients. The positive prediction rate is 62% and the negative prediction rate is 86%, with a difference of 24%, which shows that the model has a high probability of misjudgment in the prediction process. When the SMOTE method is used to process the data, the specificity of the prediction results is 90%, and the sensitivity is 91%. They are basically the same. The positive prediction rate is 85%, and the negative prediction rate is 83%. They are also basically the same. The test results of positive and negative samples have been significantly improved and achieved a relatively balanced effect.

Next, the prediction results of PSO-LSTM ischemic stroke recurrence prediction model proposed in this paper are compared with the prediction methods proposed in [14, 15]. The prediction results are shown in Table 2 below.

It can be seen from Table 2 that the accuracy of PSO-LSTM ischemic stroke recurrence prediction model proposed in this paper is 92.0%, which is improved compared with the other two prediction models. The sensitivity and specificity of the other two prediction models are lower than that of PSO-LSTM model. This is because the model proposed in [14] has relatively poor learning ability for time-series feature data set. Reference [15] does not introduce attention mechanism, while PSO-LSTM model assigns corresponding attention weight to each time-series feature, so the sensitivity and specificity of the model have been improved. In addition, it can be seen that the positive prediction rate and negative prediction rate of the prediction results of PSO-LSTM model are relatively high. This is because the adaptive PSO optimization algorithm is not introduced in [14, 15], and the introduction of PSO can reduce the prediction omission in the prediction process. In addition, SMOTE method is used to process the sample data, which eliminates the imbalance of positive and negative sample data and makes F1_score increase.

In order to better illustrate the consistency between the prediction results of PSO-LSTM ischemic stroke recurrence prediction model proposed in this paper and the actual results, the receiver operating characteristic (ROC) curve of

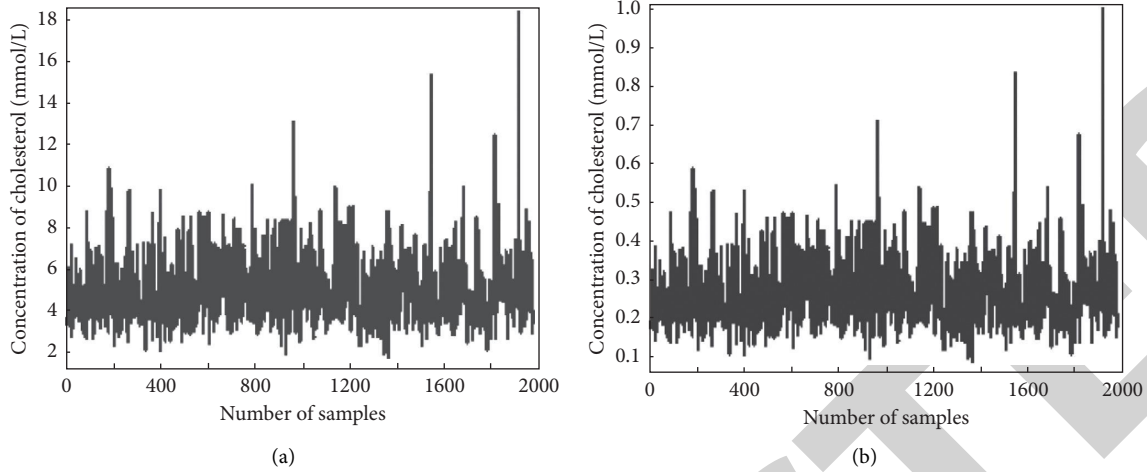


FIGURE 2: Sample data before and after normalization. (a) Initial source data. (b) Normalized data.

TABLE 1: Calculation results of different evaluation indexes before and after processing by SMOTE method.

Evaluation index	Before SMOTE processing (%)	After SMOTE processing (%)
Accuracy (E_a)	78.2	92.0
Sensitivity (E_{se})	69.3	91.2
Specificity (E_{sp})	81.1	90.5
Positive prediction rate (E_p)	62.6	85.3
Negative prediction rate (E_n)	86.4	83.7
F1_score (E_{F1})	65.8	88.5

TABLE 2: Evaluation indexes for prediction results of different methods.

Evaluation index	Model		
	PSO-LSTM (%)	Reference [14]	Reference [15]
Accuracy (E_a)	92.0	86.2	85.1
Sensitivity (E_{se})	91.2	84.4	83.5
Specificity (E_{sp})	90.5	82.2	80.1
Positive prediction rate (E_p)	85.3	82.6	83.1
Negative prediction rate (E_n)	83.7	80.7	80.2
F1_score (E_{F1})	88.5	82.5	81.9

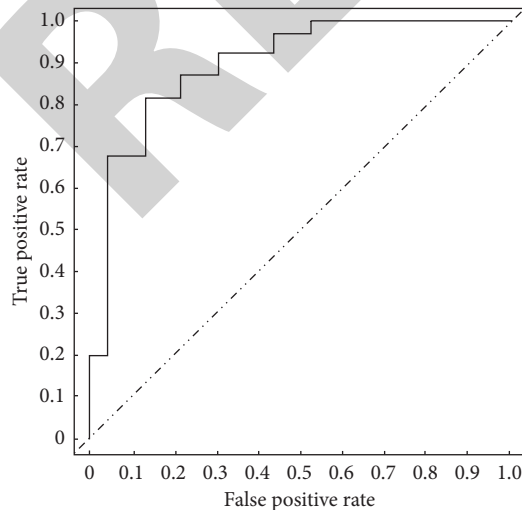


FIGURE 3: ROC curve of PSO-LSTM prediction model.

the model is analyzed below. The ROC curve of PSO-LSTM prediction model is shown in Figure 3.

As can be seen from Figure 3, the ROC curve of PSO-LSTM prediction model is relatively far from the 45° classifier baseline with discrimination of 0, and the lower area of ROC curve reaches 0.89, which shows that PSO-LSTM prediction model has strong discrimination and good performance. By introducing the adaptive PSO algorithm into LSTM, the rapid determination of the optimal super parameters is realized based on the historical characteristics of patient clinical data. By using SMOTE method to process the original data, the effective prediction of ischemic stroke recurrence is realized and the prediction accuracy is improved.

5. Conclusion

According to the historical clinical data of ischemic stroke patients, reasonable prediction of stroke recurrence can

effectively reduce the mortality of patients. Therefore, an ischemic stroke recurrence prediction model using deep learning in mobile medical monitoring system is proposed by introducing adaptive PSO algorithm into LSTM. In order to solve the clustering and searching problem of the existing prediction models for the clinical data of stroke patients, the proposed model introduces adaptive learning strategy based on PSO algorithm. By dividing the types and update methods of particles, it avoids the possibility of falling into local optimization on the basis of improving the diversity of the population and improves the clustering performance and searching speed of the model. Experiments based on the data from big data management platform for stroke patients show that the accuracy of PSO-LSTM ischemic stroke recurrence prediction model proposed in this paper is 92% and F1_score is 88%, which are better than the prediction performance of the other two models. In addition, the sensitivity and specificity, positive prediction rate, and negative prediction rate of PSO-LSTM prediction model are improved compared with the other two models, and the lower area of ROC curve reaches 0.89, which has better performance. Future work will further study the prediction effect of the proposed PSO-LSTM prediction model on patients with ischemic stroke and other types of diseases and study the performances of different diseases on the prediction model.

Data Availability

The data included in this paper are available without any restriction.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the science and technology research project of Heilongjiang Provincial Department of Education (no. 2018-KYYWF-0104) and application research of mobile medical monitoring system for stroke disease prevention and rehabilitation (subject no. 2018-KYYWF-0104).

References

- [1] H. Kong and J. Chen, "Medical monitoring and management system of mobile thyroid surgery based on internet of things and cloud computing," *Wireless Communications and Mobile Computing*, vol. 2021, no. 2, 10 pages, Article ID 7065910, 2021.
- [2] W. An, Y. Zhang, Z. Tang, and Z. Yang, "Evaluation of carotid plaque by time-intensity curve quantitative analysis of contrast-enhanced ultrasound combined with Essen stroke risk score in predicting the risk of recurrence of ischemic stroke," *China Journal of Modern Medicine*, vol. 30, no. 20, pp. 17–21, 2020.
- [3] J. Liu, H. Zhao, C. Liu, and Q. Jia, "Privacy data security policy of medical cloud platform based on lightweight Algorithm model," *Scientific Programming*, vol. 2021, no. 4, 9 pages, Article ID 5543714, 2021.
- [4] Y. Yang, K. Yan, Y. Li, Y. Qiang, J. Min, and Z. Cairong, "Value of CHA2DS2-VASc score in predicting stroke recurrence in first-ever ischemic stroke survivors without atrial fibrillation," *Journal of Southern Medical University*, vol. 40, no. 6, pp. 786–792, 2020.
- [5] X. M. Chen, L. Wang, J. Y. Jiang, Y. Gao, and J. Li, "Association of total cholesterol/high density lipoprotein cholesterol ratio and stroke recurrence," *Chinese Journal of Practical Nervous Diseases*, vol. 23, no. 9, pp. 771–775, 2020.
- [6] X. M. Chen, L. Wang, J. Y. Jiang et al., "Association of neuroimaging markers of cerebral small vessel disease with short-term outcomes in patients with minor cerebrovascular events," *BMC Neurology*, vol. 21, no. 1, pp. 44–52, 2021.
- [7] S. Chabrier, P. Lasjaunias, and M. Tardieu, "Specificities of arterial ischemic stroke in childhood," *Archives de Pediatrie*, vol. 8, no. 3, pp. 299–307, 2021.
- [8] Y. Gao and Y. Xie, "Review on risk factors for recurrence of ischemic stroke," *World Science and Technology-Modernization of Traditional Chinese Medicine*, vol. 22, no. 2, pp. 452–458, 2020.
- [9] Y. Yu, X. Zhang, and R. Fu, "Study on influencing factors and predictive models of stroke recurrence in cerebral ischemic stroke patients treated by aspirin," *Chinese Journal of New Drugs and Clinical Remedies*, vol. 38, no. 11, pp. 681–686, 2019.
- [10] N. Altaf, N. Kandiyil, A. Hosseini, R. Mehta, S. MacSweeney, and D. Auer, "Risk factors associated with cerebrovascular recurrence in symptomatic carotid disease: a comparative study of carotid plaque morphology, microemboli assessment and the European carotid surgery trial risk model," *Journal of American Heart Association*, vol. 3, no. 3, 129 pages, Article ID e000173, 2014.
- [11] X. Bi, Y. Zhang, and X. Wang, "Analysis of influencing factors of early poor prognosis in elderly patients with transient ischemic attack," *Chinese Journal of Behavioral Medicine and Brain Science*, vol. 30, no. 1, pp. 52–57, 2021.
- [12] W. Cai, C. Mueller, H. Shetty, G. Perera, and R. Stewart, "Predictors of cerebrovascular event reoccurrence in patients with depression: a retrospective cohort study," *BMJ Open*, vol. 10, no. 1, 51 pages, Article ID e031927, 2020.
- [13] Y.-W. Wang and G.-M. Zhang, "New silent cerebral infarction in patients with acute non-cerebral amyloid angiopathy intracerebral hemorrhage as a predictor of recurrent cerebrovascular events," *Medical Science Monitor*, vol. 25, no. 2, pp. 418–426, 2019.
- [14] Y. Luo, Y. Shao, and D. Chen, "Risk prediction of annual stroke of ischemic stroke based on BiLSTM-Attention model," *Journal of Donghua University*, vol. 47, no. 4, pp. 62–68, 2021.
- [15] Y. An, N. Huang, R. Yang, and X. Chen, "Deep learning-based model for risk prediction of cardiovascular diseases," *Chinese Journal of Medical Physics*, vol. 36, no. 9, pp. 1103–1112, 2019.
- [16] H.-Q. Gu, X.-W. Xie, J. Jing et al., "Shuxuetong for Prevention of recurrence in Acute Cerebrovascular events with Embolism (SPACE) trial: rationale and design," *STROKE AND VASCULAR NEUROLOGY*, vol. 5, no. 3, pp. 311–314, 2020.
- [17] D. Strambo, A. Zachariadis, D. Lambrou et al., "A score to predict one-year risk of recurrence after acute ischemic stroke," *International Journal of Stroke*, vol. 16, no. 5, pp. 602–612, 2021.
- [18] A. Wei, Y. Zhang, P. I. Yani, Z. Y. Yang, and Z. H. Tang, "Contrast-enhanced ultrasound quantitative analysis of carotid plaque for predicting recurrence of ischemic stroke,"

Retraction

Retracted: Computational Intelligence Approaches in Developing Cyberattack Detection System

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. S. Alzahrani and F. W. Alsaade, "Computational Intelligence Approaches in Developing Cyberattack Detection System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4705325, 16 pages, 2022.

Research Article

Computational Intelligence Approaches in Developing Cyberattack Detection System

Mohammed Saeed Alzahrani  and Fawaz Waselallah Alsaade 

College of Computer Science and Information Technology, King Faisal University, P.O. Box 4000, Al-Ahsa, Saudi Arabia

Correspondence should be addressed to Fawaz Waselallah Alsaade; falsaade@kfu.edu.sa

Received 4 January 2022; Revised 7 February 2022; Accepted 21 February 2022; Published 18 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Mohammed Saeed Alzahrani and Fawaz Waselallah Alsaade. 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 Internet plays a fundamental part in relentless correspondence, so its applicability can decrease the impact of intrusions. Intrusions are defined as movements that unfavorably influence the focus of a computer. Intrusions may sacrifice the reputability, integrity, privacy, and accessibility of the assets attacked. A computer security system will be traded off when an intrusion happens. The novelty of the proposed intelligent cybersecurity system is its ability to protect Internet of Things (IoT) devices and any networks from incoming attacks. In this research, various machine learning and deep learning algorithms, namely, the quantum support vector machine (QSVM), k-nearest neighbor (KNN), linear discriminant and quadratic discriminant long short-term memory (LSTM), and autoencoder algorithms, were applied to detect attacks from signature databases. The correlation method was used to select important network features by finding the features with a high-percentage relationship between the dataset features and classes. As a result, nine features were selected. A one-hot encoding method was applied to convert the categorical features into numerical features. The validation of the system was verified by employing the benchmark KDD Cup database. Statistical analysis methods were applied to evaluate the results of the proposed study. Binary and multiple classifications were conducted to classify the normal and attack packets. Experimental results demonstrated that KNN and LSTM algorithms achieved better classification performance for developing intrusion detection systems; the accuracy of KNN and LSTM algorithms for binary classification was 98.55% and 97.28%, whereas the KNN and LSTM attained a high accuracy for multiple classification (98.28% and 97.07%). Finally, the KNN and LSTM algorithms are fitting-based intrusion detection systems.

1. Introduction

The Internet of Things (IoT) could be defined as interlinked systems that focus on standardized mechanisms that communicate large amounts of data [1] between Internet-connected machines. Artificial intelligence (AI), or the quality of being smart, is being introduced to gadgets, devices, houses, businesses, and maybe even communities as a result of the current innovations in IoT. IoT is considered one of the most rapidly evolving disciplines of present technology advancement, contributing significantly to a variety of domains ranging from agriculture to self-driving cars. Because it interacts with each and every form of linked system in everyday life, IoT is known as the use of Internet through everything that can help people in their daily lives.

Fundamental firewalls are static defense systems that act as channels. They are not fit for perceiving an attack. They generally obstruct all traffic with the exception of the packets coordinating a few guidelines; for example, packets are bound to a specific port or originate from the secure Internet Protocol (IP) addresses. These rules are constructed physically by the system overseer as indicated by the network security approach. This implies that the productivity of a firewall relies on how talented the administrator is [2].

The quantity of smart interconnected devices is expected to reach 1 billion in 2025 [2]. IoT is made up of numerous layers, which includes a specific layer called the network layer. The architecture of the network layer depends on the Internet, which is based on different communication layers, and is primarily capable of sending network packets among

servers. Furthermore, the network layer is a complicated and vulnerable component of the IoT structure that contributes to a variety of security problems.

Nonetheless, a number of security mechanisms exist to solve security concerns [3]. To enable a set of connected devices to function successfully and address security issues, these mechanisms must be installed in the IoT ecosystem and/or endpoints. However, many security devices require a significant amount of computing power and storage space [4]. To address these limitations, several techniques, including lightweight cryptography and authentication processes, can be used [5]. The vast number of sensors, nodes, servers, or machines associated and interlinked through the IoT architecture is indeed a major source of security concern, as a security incident in either a single node or sensor might cause the entire system to collapse. Cyberattacks, distributed denial-of-service (DDoS) hacks, ransomware, distant monitoring, packet-forwarding attacks, and privacy breaches are by far the most prevalent security vulnerabilities that IoT systems confront. A firewall is generally the first point of security against intrusions in IoT devices, although this is not an efficient option due to the wide range and complication of IoT infrastructures. Intrusion detection systems (IDSs) have risen in importance as a result of its reliability. In 1980, Spafford and James [6] offered a description of an IDS for the very first time. IDSs are designed to detect intrusions in a certain network domain. An intrusion through an IoT context can become a host that attempts to access neighboring nodes without taking permission. An IDS has three major components: a client, a screening test, and a reaction module. The client is entirely accountable for managing data from the tracking actions data stream. The intrusion prevention mechanism detects evidence of intrusion and delivers alarms. Then, the reaction module can be activated using the results that the analysis engine provides. IDSs have improved in reliability and efficiency over time, but hackers have created more diverse attack tactics to circumvent these tracking systems. Furthermore, typical IDSs are incapable of dealing well with IoT's numerous network elements, such as interconnected layers [7].

Researchers have been urged to use decentralized IDSs in addition to different machine learning techniques, including artificial neural networks (ANNs), deep learning, and optimization algorithms, because of recent advances in intelligent machines. Typical ANNs are limited in their ability to cope with the complications of IDS systems. Enhanced technology by addressing such limitations is necessary for IDSs to achieve their potential. The major objective of this paper is to apply blockchains to a multi-agent system and to evaluate its performance using a benchmarking dataset [8, 9].

The main contribution of this study is to apply various machine learning and deep learning algorithms to detect intrusion intelligently. A smart IDS can help to protect the IoT environment from any updated attacks. The system has the ability to detect and prevent cyberattacks in IoT networks. In this study, we investigate various machine learning and deep learning to detect attacks with binary and multiple

classes to determine the performance of each model. The network dataset has many network features that obstruct the IDS system from quick detection, enabling the selection of significant features that can help the system save time with a high detection rate. In this study, we use the correlation method to find features that have a significant relationship with classes. Finally, different AI algorithms are investigated to improve the performance and efficiency of IDS systems.

2. Related Work

Although the techniques of the Internet of Things are essential for enhancing real-world intelligent systems, such as applications used in smart cities, home automation, and smart factories, and their massive scale and omnipresence have presented unique security concerns [10, 11]. Additionally, because IoT systems are typically used in an uncontrolled environment, an intruder with malevolent aims may gain access to these systems [12, 13]. Snooping can sometimes be employed to get confidential details from such a transmission medium, since IoT components are generally interconnected across wireless networks [14, 15]. Due to their limited power and computing capabilities, IoT-connected devices may not have installed advanced security measures to address the upper edge of such security concerns. Specific attack interfaces emerge on something like a constant basis as a result of the IoT's complexity and interrelated settings [16, 17].

As a result, particularly in contrast to typical computing systems, IoT networks are much more exposed. To mitigate threats faced by IoT-connected devices, appropriate diagnostic and preventive strategies must be developed. Furthermore, a line of defense in distributed systems must be established to defend IoT systems from cyberattacks. IDSs are used to solve this problem [18, 19]. Machine learning-based IDSs that provide security for IoT networks or exploited IoT systems have been reported in many studies. IDSs that are implemented in cloud-based IoT networks [20], sensor networks [21, 22], cyber-physical applications [20], and wireless mobile networks [23, 24] have all been covered by the literature. Classical IDS approaches, on the other hand, are much less efficient or effective for the provision of security networks due to their unique attributes, such as limited power, pervasiveness, diversity, constrained bandwidth utilization, and global connectivity, as noted above. Deep learning and machine learning-based approaches have recently found traction for detecting cyber threats, particularly those affecting IoT networks. This is due to the fact that machine learning- and deep learning-based approaches may detect both benign and malignant abnormalities in an IoT network.

To discover the characteristics of patterns, IoT servers and network flow can be monitored and examined. Any divergence from all learned norms can be leveraged to spot abnormal activity and unusual behavior. Moreover, technologies based on machine learning and deep learning have been used to predict unknown or zero-day cyberattacks. As a result, machine learning- and deep learning-based techniques provide reliable security measures for IoT devices and

systems. Several studies have investigated various strategies for developing IDSs for IoT applications, but the majority of the abovementioned surveys did not include the adoption of machine learning or deep learning approaches, such as detection methods in IoT networks and associated compact components. The focus of several studies [25–30] was on investigating IoT security challenges broadly and their categorization in different layers related to applications, networks, cryptography, and access restrictions. An inclusive study that provides a comprehensive evaluation of machine learning and deep learning algorithms that can be adopted in IDS applications in IoT network settings is still needed, as is a key emphasis of this work.

The researchers in [31] focused on the problems with IoT security somewhere within the network layers. A study published in [32] investigated IDS technologies for IoT networks. A preliminary examination of machine learning's applicability in the domain of IoT confidentiality and protection was addressed in [33]. Furthermore, they highlighted bandwidth limitations, processing power limitations, and a lack of suitable space as obstacles in applying any machine learning-based security mechanisms for IoT interconnected systems. Other studies [34, 35] explored the possibility of using machine learning and data mining algorithms to identify malicious attacks and intrusions in IoT networks by incorporating these algorithms in IDSs and recognizing abnormalities or using network data classification. The authors in [20] pointed out differences among IDSs that operate on cellular broadband and wireless communication networks, particularly IoT networks. Due to basic architectural differences, applying machine learning approaches to IoT IDSs involves special attention to the details of cyberattacks, supporting protocols (including both telecommunications and networking), and the application layer. A further study reported in [21] explored how IDSs can be implemented in mobile ad hoc networks. Three major kinds of IDS layouts can be used in mobile ad hoc networks (MANETs). A layered architecture is the first layout that is organized with several hierarchical layers. For deployment in a decentralized and collaborative setting, the second architectural is also flattened. The third layout can be a combination of the first two employed in mobile agents. An additional study [22] explored a number of intrusion detection techniques for mobile ad hoc-based IDS architectures. These IDS techniques, as per the authors, can indeed be divided into several classified methods based on the basic principles employed to identify an intrusion. Rules, metrics, optimizations, signatures, contexts, popularity scores, or pathways can always be utilized as principles in IoT systems. Anomaly discovery, exploitation, signature-based algorithms, and evolutionary algorithms were eventually included in the list of hybrid technologies.

Other classification criteria have been proposed as well [22]. For example, these include real-time/offline, attack type, and effectiveness of detection (scalability, reliability, timeliness, etc.). Other authors provided further classification criteria, such as legitimacy, intrusion patterns, and identification efficacy (scalability, reliability, timeliness, etc.) [22]. In a different study, the author discussed a

categorization of IDS for wireless sensor networks (WSN) depending on the IDS agent's configuration model [29]. The configuration model might be decentralized, centralized, or mixed, with the last model being recommended as the ideal fit for WSNs. A similar survey presented in [30] categorized WSNs relying on IDS by utilizing IDS detecting class criteria. Outlier detection, abuse detection, and recognition based on configuration were among the categories discovered. A further facet of the virtualized IoT ecosystem was examined and described in [15] in which the authors of this study evaluated and categorized several cloud-based IDSs that influence the confidentiality, authenticity, and reliability of cloud computing that depend on IoT networks. Hypervisor-based IDS, host-based IDS (HIDS), network-based IDS (NIDS), and scalable IDS were all discussed as well. The authors in [29] introduced a research study on IoT-based IDS and specifically focused on IDS design. They looked at current IoT standards, protocols, and solutions, as well as IoT privacy concerns and detecting categories, before proposing an IoT IDS design. In [36], the authors presented a new multiphase anomaly identification technique based on Boruta Firefly aided partitioning density-based spatial clustering of applications with noise (BFA-PDBSCAN). Furthermore, they assumed that their suggested approach provided better experimental results in matching the specified methods of density-based spatial clustering of applications with noise (DBSCAN) and hierarchical density-based spatial clustering of applications with noise (HDBSCAN).

The researchers in [37] presented an integrated data processing approach for outlier identification and classification that incorporates grey wolf optimization (GWO) and convolutional neural network (CNN) algorithms. The researchers stated that their method outperformed existing state-of-the-art IDSs in terms of effectiveness and detection accuracy. A sophisticated autoencoder-based anomaly detector system was utilized to analyze and diagnose IoT botnet intrusions [38]. The approach involved obtaining statistical properties from behavior snapshots of typical IoT edge device data patterns and developing a deep learning-based autoencoder just on extracted features from the used dataset. Furthermore, the reconstruction of errors for traffic measurements was matched to a threshold to determine whether they are normal or abnormal. The authors assessed the suggested identification approach using the BASHLITE and Mirai botnets dataset created with the help of industrial IoT systems.

3. Materials and Methods

Figure 1 displays the formwork of the proposed system for detecting intrusion from a real dataset.

3.1. Dataset. The KDD Cup dataset was employed to investigate our proposed system. The NSL-KDD is an updated version of the KDD Cup dataset proposed by McHugh [39]. Furthermore, each record consists of 41 features, and these features can be described as either normal or attacks. The

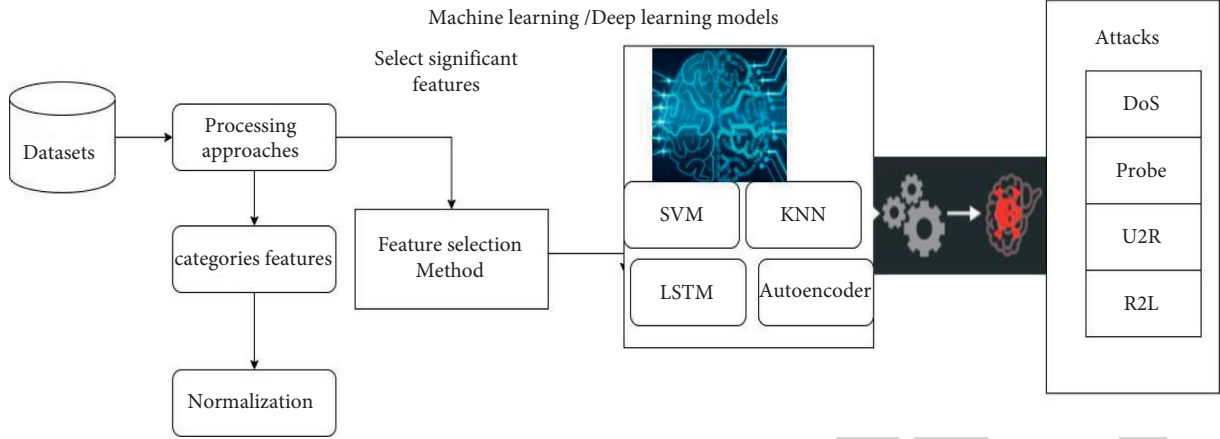


FIGURE 1: Proposed system.

KDD Cup and NSL-KDD datasets contain three major intrusions, namely, denial-of-service (DOS), probe, root to local (R2L), and user to root (U2R). Table 1 demonstrates the feature names for the KDD Cup dataset.

Furthermore, the attack types of the KDD Cup datasets are clustered into four different attack classes: (1) DoS, which includes attacks that cause the slowing or shutting down of a machine by sending more traffic information to the server than the system is able to handle. DoS attacks affect legitimate network traffic or access to services; (2) R2L includes attacks that provide illegal local access to a machine by sending remote deceiving packets to the system; (3) U2R includes attacks that provide root access, and in this case, the hacker finds out the system vulnerability and starts using the system as a normal user; and (4) probe includes attacks that can avoid security control systems by gathering information about the network. The attack categories of the KDD Cup are reported in Table 2.

3.2. Preprocessing. The processing method was applied to select significant features from the dataset.

3.2.1. One-Hot Encoding. One-hot encoding was proposed to convert categorical features, namely, protocol type, service, and flag, into numerical features. One-hot encoding is used to assign each string to a new binary value [0, 1]. Table 3 shows the categorical features of both datasets.

3.2.2. Normalization Method. After transforming the categorical features, the data were processed using min-max normalization methods for normalizing the data to avoid overlap in the training process that can occur when handling the largest dataset. In the normalization method used to scale the dataset in the same range, we put the scaling range of data between 0 and 1.

$$z_n = \frac{x - y}{x - y} (x_i - y_i) + y_i, \quad (1)$$

TABLE 1: Feature names of the KDD Cup dataset.

S. No	Feature names
1	Duration
2	Protocol type
3	Service
4	Src-byte
5	Dst-rate
6	Flag
7	Land
8	Wrong_fragment
9	Urgent
10	Hot
11	Num_failed_login
12	Logged_in
13	Num_compromised
14	Root_shell
15	Su_atte- + mpted
16	Num_root
17	Num_file_creation
18	Num_shells
19	Num_acces_shells
20	Num_outbound_cmds
21	Is_hot_Login
22	Ist_guest_Login
23	Count
24	Serror_rate
25	Rerror-rate
26	Same-Srv-rate
27	Diff-Srv-rate
28	Srv_Count
29	Srv_serror_rate
30	Srv_rerror_rate
31	Srv_Diff_host_rate
32	Dst_host_count
33	Dst_host_srv_count
34	Dst_host_same_srv_count
35	Dst_host_diff_srv_count
36	Dst_host_same_src_port_rate
37	Dst_host_srv_diff_host_rate
38	Dst_host_serror_rate
39	Dst_host_srv_serror_rate
40	Dst_host_rerror_rate
41	Dst_host_srv_rerror_rate

TABLE 2: All types of attacks in the KDD Cup.

Attacks in datasets	Type of attacks in KDD Cup
DoS Probe	Back, Land, Neptune, Pod, Smurf, Teardrop, Mailbomb, Processtable, Udpstor m, Apache2, Worm Satan, IPSweep, Nmap, Portsweep, Mscan, Sa int
R2L	Guess_password, Ftp_write, Imap, Phf, Multihop, Warezmaster, Xlock, Xsnoop, Smpgue, ss, Smpgetattack, Httptunnel, Sendmail, Named
U2R	Buffer_overflow, Loadmodule Rootkit, Perl, Sqlattack, Xterm, Ps
Attacks in datasets	Type of attacks in NSL-KDD
DoS Probe	Back, Land, Neptune, Pod, Smurf, Teardrop, Mailbomb, Processtable, Udpstor m, Apache2, Worm Satan, IPSweep, Nmap, Portsweep, Mscan, Saint
R2L	Guess_password, Ftp_write, Imap, Phf, Multihop, WarezmasterXlock, Xsnoop, Smpgue, ss, Smpgetattack, Httptunnel, Sendmail, Named
U2R	Buffer_overflow, Loadmodule Rootkit, Perl, Sqlattack, Xterm, Ps

TABLE 3: Categorical features.

S. No	Feature name
2	Service
3	Flag
6	Protocol type

where y and x are the minimum and maximum data, respectively. The maximum range is represented by y_i , whereas the minimum range is indicated by x_i [0].

3.2.3. *Feature Selection Method.* Correlation analysis was used to find correlations between the features and classes. It is also used to find significant patterns between features of datasets for intrusion detection.

$$R = \frac{n\sum(x \times y) - (\sum x)(\sum y)}{[n\sum(x^2) - \sum(x^2)] \times [n\sum(y^2) - \sum(y^2)]} \times 100\%, \quad (2)$$

where R is Pearson's correlation coefficient approach, x is input training, and y is target (classes). We considered the threshold value to be 0.50, the features with a greater-than-0.50 relationship with classes were selected, and everything else was excluded. Table 4 shows the selected features among 41 features of the KDD Cup datasets. According to the results of the correlation analysis method, the same_srv_rate had a high correlation among 75% all features; therefore, we considered these features as significant.

3.3. *Classification Algorithms.* In this section, the classification algorithm is presented.

3.3.1. *Support Vector Machine (SVM).* The support vector machine (SVM) is a prevalent supervised nonlinear technique that can be applied to distribute data sequentially and nonsequentially for classification tasks. SVM is used for text classification, image processing, and anomaly analysis. Furthermore, it has the ability to deliver good accuracy for high-dimensional vector space data and symbolizes data training features in space maps. The data features of the

TABLE 4: Selected features.

No.	Feature name	Correlation (ranking%)
23	count	0.576257
30	srv_serror_rate	0.648135
24	serror_rate	0.650527
38	dst_host_serror_rate	0.651740
39	dst_host_srv_serror_rate	0.654855
12	logged_in	0.690053
36	dst_host_same_srv_rate	0.693525
33	dst_host_srv_count	0.722356
26	same_srv_rate	0.751746

several classes are distinguished based on a maximum margin in the hyperplane. The decision boundary that can be achieved by the SVM technique is represented by the extreme margin space for determining the distance between the training samples of two or more classes. The equation for the SVM classifier is given as follows:

$$K(X, X') = \exp\left(-\frac{\|X - X'\|^2}{2\sigma^2}\right), \quad (3)$$

where X, X' is the feature vector for the training of the evaluated dataset, $\|X - X'\|^2$ denotes the squared Euclidean difference among two feature inputs, and σ is a free parameter.

3.3.2. *KNN Algorithm.* When the KNN algorithm is adopted for the classification task, it performs the classification of various feature values by computing the distance between each pair. An integer number not more than 20 usually specifies the k parameter in this algorithm. While working on the KNN algorithm, the decided neighbors can be represented by various objects that have been accurately identified and categorized. This technique only identifies the class of the sample and can be based on the class of the neighboring one or various samples in the decision making regarding categorization. KNN is utilized to determine the k values, which are near a set of values through the training dataset, and the majority of these k values fall to a confirmed class; furthermore, the input sample is classified. The equation that was applied for the KNN algorithm is written as follows:

$$\sqrt{(x_1 - x_2) + (y_1 - y_2)}. \quad (4)$$

The k value is utilized to find and calculate the nearest points in the feature vectors. As such, the value must be distinctive.

3.3.3. Long Short-Term Memory (LSTM). Hochreiter and Schmidhuber [40] proposed the long short-term memory (LSTM) approach for learning long-term information interdependence. An LSTM's flow is similar to that of the recurrent neural network (RNN) method. The difference in how the cells are operated between the LSTM and RNN approaches is that there are four gates in each LSTM unit, specifically the input, candidate, forget, and output gates. The forget gate determines whether data should be saved or destroyed. The cells are refreshed by the input gate, while the output gate always determines the hidden state in the LSTM. The LSTM also has an incorporated memory block and gate mechanism that allows it to resolve vanishing gradient point problems and disintegration gradient complications through the RNN learning process [41, 42]. The structure of the LSTM technique is expressed in Figure 2.

The computing equations that are associated with the LSTM structure in Figure 1 are as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot X_t + W_f \cdot h_{t-1} + b_f), \\ i_t &= \sigma(W_i \cdot X_t + W_i \cdot h_{t-1} + b_i), \\ S_t &= \tanh(W_c \cdot X_t + W_c \cdot h_{t-1} + b_c), \\ C_t &= i_t * S_t + f_t * S_{t-1}, \\ o_t &= \sigma(W_o \cdot X_t + W_o \cdot h_{t-1} + V_o \cdot C_t + b_o), \\ h_t &= o_t + \tanh(C_t). \end{aligned} \quad (5)$$

The mathematical symbolization in the above equations can be interpreted and expressed as follows:

X_t is the vector of the input data that progress to the memory cell at time t . W_i , W_f , W_c , W_o , and V_o are the weight matrices. b_i , b_f , b_c , and b_o represent bias vectors. h_t is the specified value of the memory cell at time t . S_t and C_t are the defined values of the candidate state of the memory cell and the state of the memory cell at time t , individually.

σ and \tanh are the activation functions in the LSTM network.

i_t , f_t , and o_t are acquired values for the input gate, the forget gate, and the output gate at time t . These gates have values in the range of 0 to 1 over the nonlinear sigmoid activation function.

3.3.4. Deep Autoencoder Algorithm. Encoders and decoders are two primary components of an autoencoder technique. An encoder component reduces the dimensionality of input data into the lowest dimensional exemplification form, while the decoder reproduces input data depending on the lowest data representation, which is made by the encoder component. Autoencoders, on the other hand, automatically encode all data of the input layer and forward these data into

hidden layers before finally decoding the data into the production layer (output layer) in the network [43–47]. Considering the efficiency of autoencoders in discovering different sorts of attacks, the recognition accuracy of an autoencoder-based deep learning model for IDSs might be highly dependent on the nature of the autoencoder model's design and hyperparameter configurations. As a result, finding ideal settings of autoencoders that can lead to better detection accuracy is crucial. Earlier mainstream studies described individually obtaining the right model by running several tests with specific datasets. Human procedure testing takes a long time in intrusion detection tasks, and they must be performed whenever data are updated [48–52]. The deep autoencoder (DAE) model for IDSs achieved through two processes can handle the IoT network security problem. These processes are training and testing [53, 55]. The system utilizes a training dataset to generate a classifier obtained by the selected DAE. In the testing process, an IDS uses the autoencoder model to recognize the class of each sample in the testing dataset to evaluate the overall performance of the system when it can be applied to an online environment. Figure 3 illustrates the suggested DAE structure for intrusion detection that consists of three different layers: the input, hidden, and output layers.

3.4. Performance Measures. The performance measures were used to test the outcomes of the proposed model. Accuracy, false positive, precision, true positive, and time were used. The equations for performance measures are as follows:

(a) Accuracy

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%, \quad (6)$$

(b) Precision

$$\text{precision} = \frac{TP}{TP + FP} \times 100\%. \quad (7)$$

(c) F-score

$$F - \text{score} = \frac{2 * \text{precision} * \text{sensitivity}}{\text{precision} + \text{sensitivity}} \times 100\%, \quad (8)$$

where TN represents true negative, TP represents true positive, FP represents false positive, and FN represents false negative.

4. Experimental Results

This section describes the experimental analysis of the proposed model developed during the research phase. Two experiments were conducted to improve the IDS. The experiment was conducted and evaluated by utilizing the KDD Cup dataset. Python programming language was used to implement all machine learning and deep learning algorithms to design the model. The Jupyter platform was used to run all code. In this study, two experiments were prepared to classify and identify intrusions from the IoT platform.

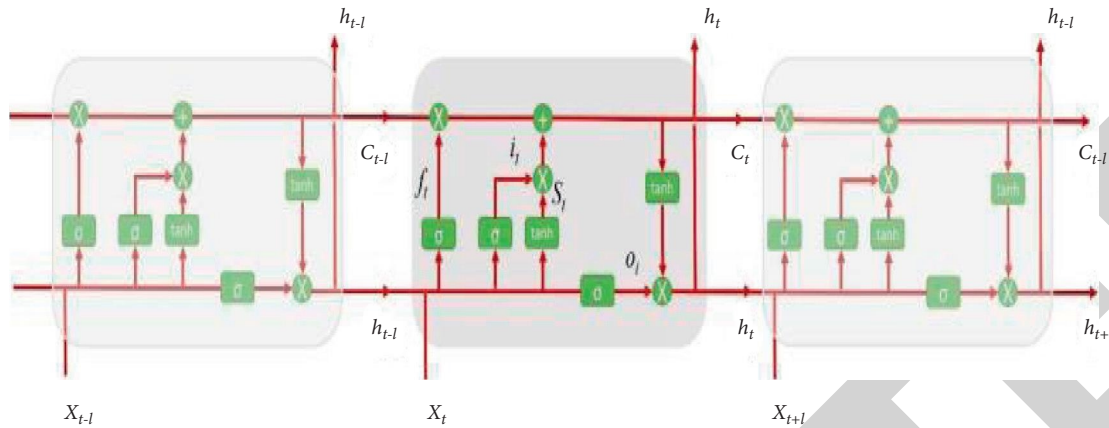


FIGURE 2: Architecture of the LSTM technique.

4.1. Results of Binary Classification. In this section, machine learning and deep learning algorithms are proposed to classify intrusion as normal or attacks.

4.1.1. Machine Learning Algorithm with Binary Classification. In this experiment, binary classifications, namely, QSVM, KNN, linear discriminant, and quadratic discriminant algorithms, were applied to detect intrusion. The binary classifications included two classes (normal and attack packets). Figure 4 shows the instance values of KDD Cup data for normal and attack classes.

The dataset was divided into 70% for training and 30% for testing, and the testing dataset was processed to validate the machine learning algorithms. The evaluation metrics accuracy, precision (%), recall, and F1 score were employed to examine the proposed algorithm to classify intrusion. Table 5 shows the results of the machine learning algorithms. The KNN algorithm achieved high accuracy (98.55%). The quadratic discriminant algorithm obtained lower accuracy (68.91%). Based on these results, we confirmed that the KNN algorithm is an appropriate algorithm for binary classification.

The statistical metrics to find the prediction errors, namely, MAE, MSE, RMSE, and R^2 , were used to measure the relationship between the actual values and predicted values. Table 6 summarizes the prediction errors for machine learning to classify the intrusion. It is noted that the KNN algorithm had a robust correlation between the prediction output and classes; the prediction errors of outputs from the KNN algorithm were MSE (0.01449) and ($R^2 = 94.17\%$).

4.1.2. Results of Deep Learning for Binary Classifiers. In this experiment, the LSTM and autoencoder algorithms were applied to classify intrusion as normal and attack. Table 7 displays the results of deep learning. LSTM achieved good accuracy in detecting intrusion. We observed that the performance of the LSTM algorithm was better than the DAE algorithm. The LSTM approach achieved high accuracy (97.82%).

The performance of the LSTM model to identify intrusion is presented in Figure 5. The accuracy of the LSTM model started at 82% and increased to 98% with 20 epochs. The cross-entropy loss of the LSTM model is shown that validation loss decreased to 0.4.

The training and testing accuracy performance of the DAE algorithm is displayed in Figure 6. The testing accuracy of the DAE algorithm reached 88%. The training loss was 0.114, and the testing loss was 0.106.

4.2. Results of Multiple Classifications. In this experiment, 34 major attacks and normal packets were considered in the KDD Cup for detecting malicious attacks. The machine learning algorithms assessed were the QSVM, KNN, linear discriminant, and quadratic discriminant algorithms. The dataset has four major attacks, namely, DoS, Probe, U2R, and R2L attacks. In the KDD Cup dataset, the DoS attack contains 45570 record packets and was divided into 70% for training and 30% for testing. Table 8 shows the instance values of these attacks. The instance values of each attack are presented in Figure 7.

4.2.1. Machine Learning Algorithm with Multiple Classifications. Table 9 indicates the results obtained using the linear SVM, KNN, linear discriminant, and quadratic discriminant algorithms. From the experimental results, the KNN algorithm achieved 98.28% accuracy for all attacks. Furthermore, the KNN algorithm achieved high accuracy against linear SVM, discriminant, and quadratic discriminant algorithms.

The prediction errors metrics, such as MAE, MSE, RMSE, and R^2 , were employed to measure the performance of the machine learning models. The prediction of machine learning, namely, linear SVM, KNN, linear discriminant, and quadratic discriminant algorithms, is summarized in Table 10. The prediction errors of the KNN model were very low (MSE = 0.050), and the correlation between the actual data and prediction was $R^2 = 95.22\%$. This indicates the strength of the KNN model in detecting attacks, namely, DoS, Probe, U2R, and R2L attacks.

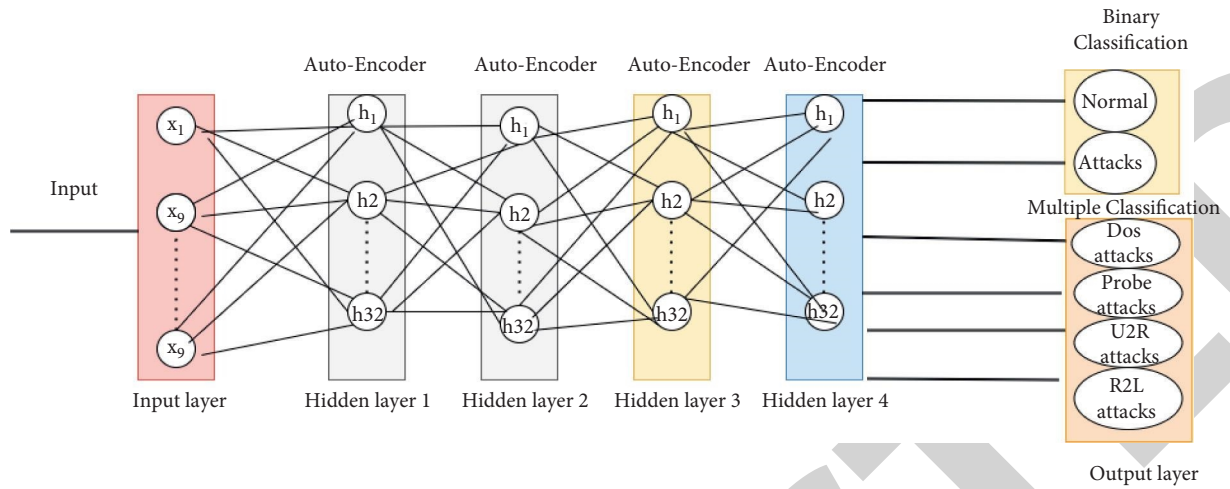


FIGURE 3: The structure of the autoencoder model for an IDS.

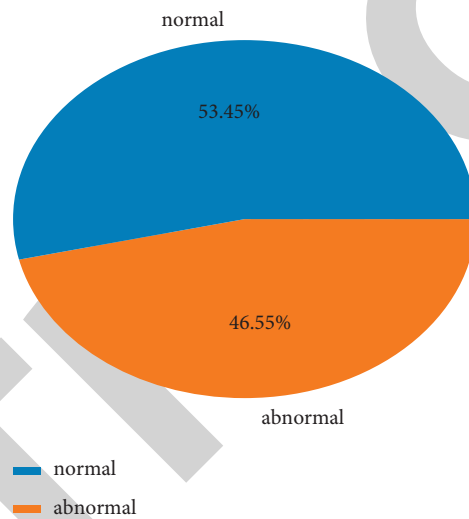


FIGURE 4: Percentage instance values of the KDD Cup data.

TABLE 5: Performance of binary classifiers to detect intrusion.

Models	Network packets	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
QSVM	Normal	95.77	93	92	96
	Attacks		99	99	95
KNN	Normal	98.55	98	99	98
	Attacks		99	98	99
Linear discriminant	Normal	96.77	96	98	97
	Attacks		97	96	97
Quadratic discriminant	Normal	68.91	63	100	77
	Attacks		76	99	86

TABLE 6: Statistical analysis of binary classifiers to predict intrusion.

Models	MAE	MSE	RMSE	R ² %
QSVM	0.0422	0.0422	0.20	83
KNN	0.0144	0.01449	0.120	94.17
Linear discriminant	0.0323	0.0322	0.1796	87.04
Quadratic discriminant	0.3101	0.310	0.55	13.82

TABLE 7: Results of deep learning in binary classes.

Models	Loss	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
LSTM	0.063	97.82	97.25	98.12	97.97
DAE	0.1040	87.40	76.25	98.84	85.71

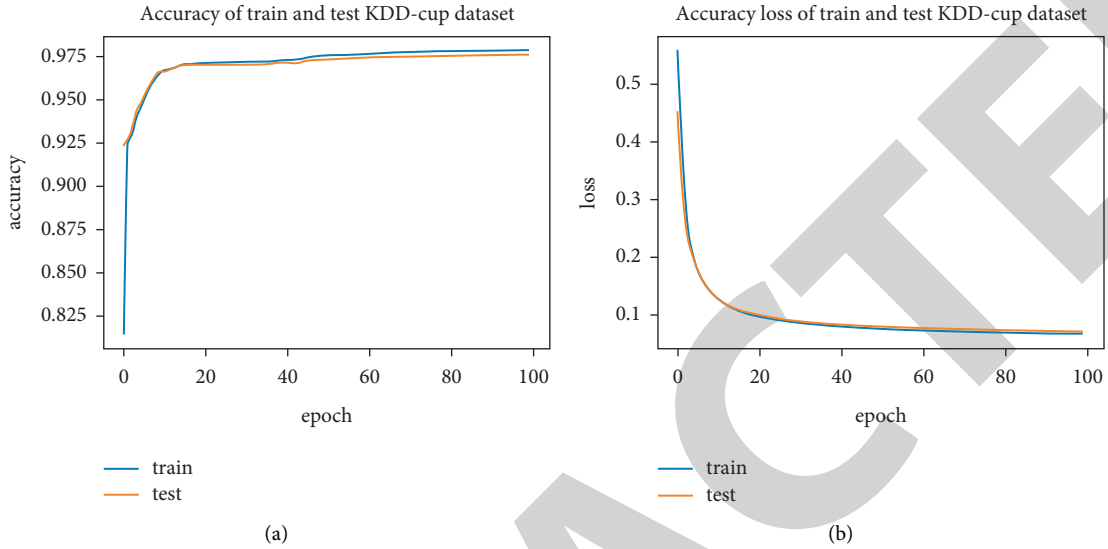


FIGURE 5: Performance of LSTM model on binary classification.

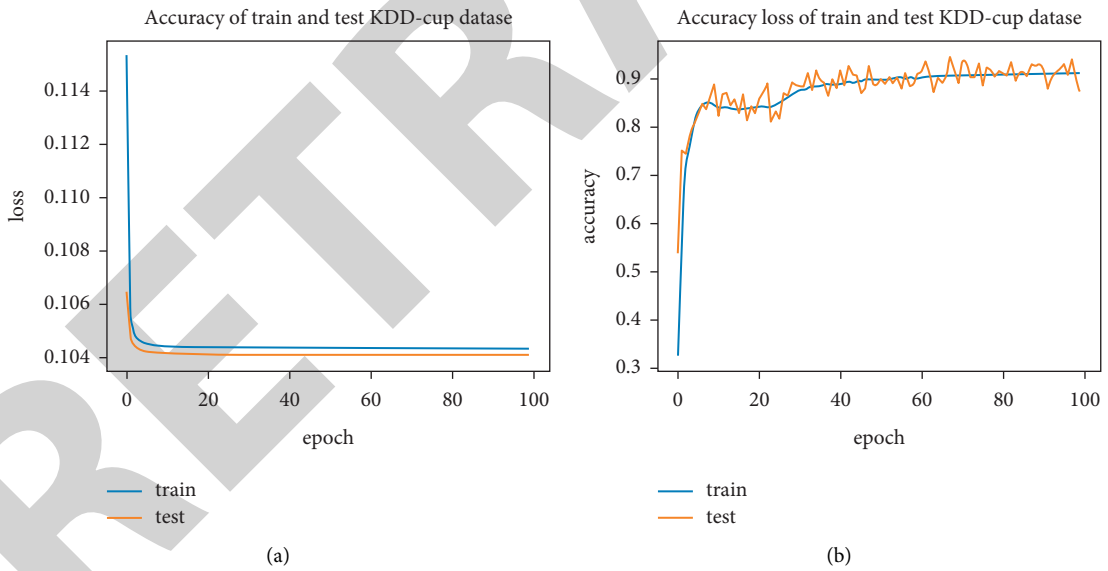


FIGURE 6: Performance of DAE model on binary classification.

TABLE 8: Instance values of attacks.

Attacks	#Instance values
Normal	66810
DoS	45570
Probe	11579
R2L	990
U2R	52

4.2.2. Results of Deep Learning for Multiple Classifications.

The LSTM and DAE algorithms were applied to detect DoS, Probe, U2R, and R2L attacks. Table 11 summarizes the results of deep learning. The LSTM model achieved high accuracy compared with the DAE algorithm. The accuracy percentage of LSTM was 97.07% for the classification of multiple attacks.

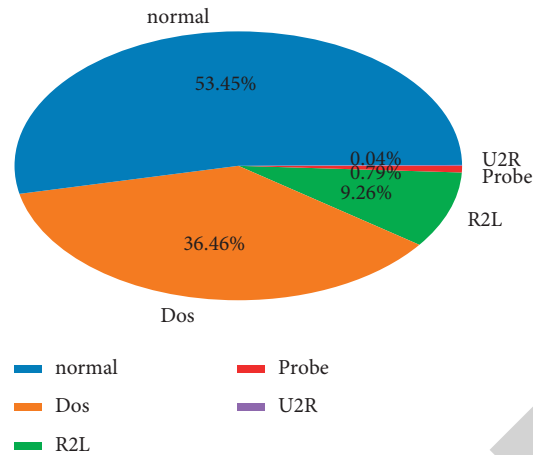


FIGURE 7: Percentage of values of attacks.

TABLE 9: Performance of machine learning algorithms in detecting multiple classes.

Model	Attacks	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
Linear SVM	DoS	95.39	95	96	96
	Probe		87	79	83
	R2L		63	62	62
	U2R		0.00	0.00	0.00
	Normal		97	98	98
QSVM	DoS	92.89	96	94	95
	Probe		97	60	74
	R2L		0.00	0.00	0.00
	U2R		0.00	0.000	0.00
	Normal		91	100	95
KNN	DoS	98.28	99	98	99
	Probe		96	97	96
	R2L		91	80	85
	U2R		57	27	36
	Normal		98	99	99
Linear discriminant	DoS	93.18	94	96	95
	Probe		89	73	80
	R2L		33	88	48
	U2R		0.04	60	0.08
	Normal		97	95	96
	DoS	61.79	94	86	90
	Probe		84	28	42
	R2L		0.03	100	0.06
	U2R		0.00	0.00	0.00
	Normal		75	51	61

The performance of LSTM in the testing and training processes is presented in Figure 8. The performance curve shows that the accuracy started from 40% and reached 97.07%, which indicates the reliability of the LSTM model in detecting multiple attacks, and training loss of the LSTM model is decreased to 1.2.

The performance of the autoencoder algorithm is displayed in Figure 9 the cross-entropy loss of the autoencoder algorithm for training and testing is presented, and it is observed that the performance accuracy of the autoencoder algorithm for 200 epochs was not good.

5. Discussion

Machine learning is a kind of information-driven approach in which the first step is possible when the data are understood. In the present work, we used data on essential ranking attacks. We presented different ways to apply machine learning techniques to design IDSs for various kinds of data. The various kinds of data represent specific attack behaviors, including the behaviors and activities of the host on the network. Server logs reflect host behaviors and network traffic that represent network behaviors. There are

TABLE 10: Statistical analysis of machine learning for multiple classification.

Models	MAE	MSE	RMSE	$R^2(\%)$
Linear SVM	0.097	0.274	0.524	92.41
QSVM	0.20	0.648	0.8050	82.81
KNN	0.050	0.172	0.4158	95.22
Linear discriminant	0.145	0.401	0.633	88.90

TABLE 11: Results of deep learning for multiple classification.

Models	Loss	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
LSTM model	0.088	97.07	97.34	96.86	97.10
Autoencoder model	0.0676	80.01	80	78.23	88.23

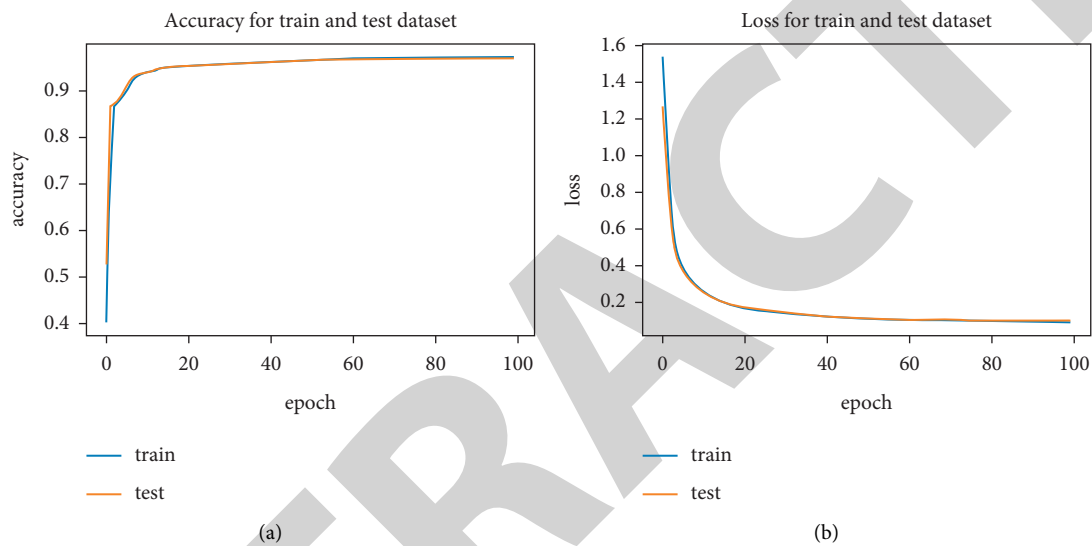


FIGURE 8: Performance of LSTM model on multiple classification.

several types of attacks, and each has a particular pattern. Therefore, it is important to select suitable data sources to detect various attacks as per the features of the threat. One of the main features of the DoS attack, for example, is that it is employed to dispatch several packets in a very short period of time, so data stream is ideal for DOS attack detection. A hidden channel includes a data-leaking operation between two different IP addresses and is best for session data discovery.

Developing intelligent systems based on machine learning and deep learning approaches was the main purpose of this study. The KDD Cup dataset is a common network dataset that contains several attacks that were used to evaluate the proposed intelligent model. In this research, we applied various machine learning and deep learning models to design cybersecurity systems in the IoT environment. During the training of the models, we observed the robustness of each model for detection intrusion.

Two experiments were conducted for binary and multiple classification. The main objective was to use the two experiments to design the signature database for detection intrusion. The empirical results of two experiments showed the appropriate algorithms for detecting

binary and multiple classes. Table 12 shows the comparison of machine learning and deep learning models for binary and multiple classes in terms of accuracy. Among the various machine learning and deep learning algorithms, the KNN and LSTM models were found to be appropriate models for detecting intrusion with binary and multiple attacks.

The KNN and LSTM model achieved high accuracy percentages for binary and multiple classification. The performance of KNN showed 98.55% accuracy, where the accuracy of the KNN for classifying multiple classes was 98.28%. Furthermore, the LSTM showed scores of 97.82% and 97.07% for detecting intrusion by binary and multiple classification, respectively.

Receiver operating characteristic (ROC) curves for the LSTM model with binary and multiple classifications are presented in Figure 10. The ROC graphs show the significance of the LSTM model in classifying multiple classes. The y-axis represents the true-positive rate of the LSTM model, and the x-axis indicates the false-positive rate of LSTM model in detecting normal, DoS, Probe, R2L, and U2R attacks. Overall, the KNN and LSTM models are the best algorithms for detecting attacks of binary and multiple databases.

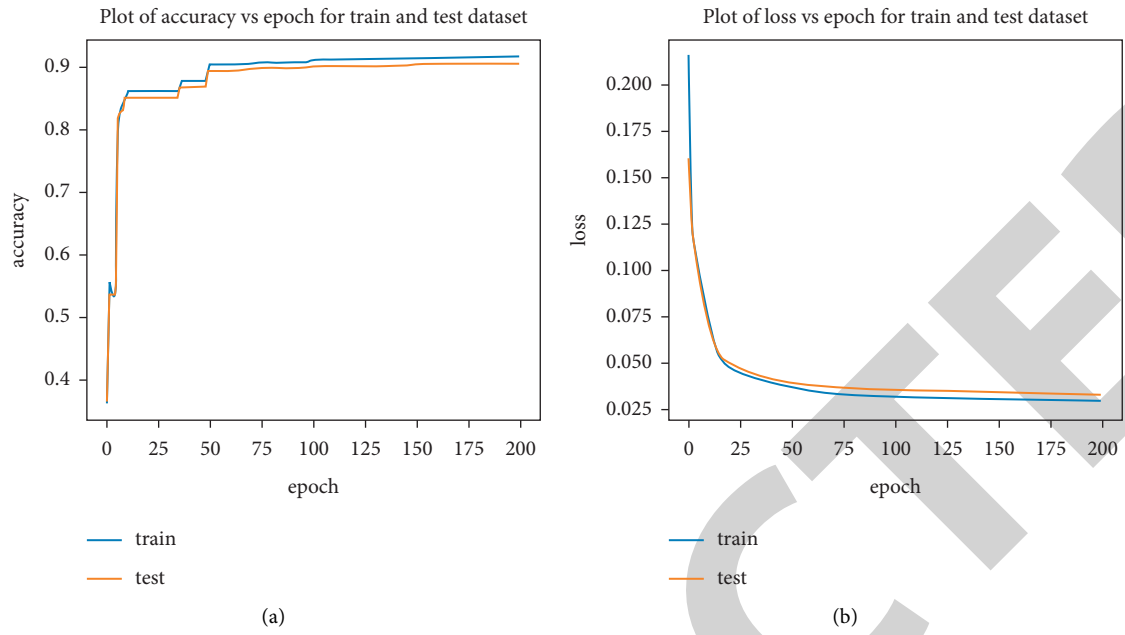


FIGURE 9: Performance of DAE model on multiple classification.

TABLE 12: Significant results of the proposed system.

Model	Accuracy	Experiments
KNN	98.55	Binary classification
LSTM	97.82	Binary classification
KNN	98.28	Multiple classification
LSTM	97.07	Multiple classification

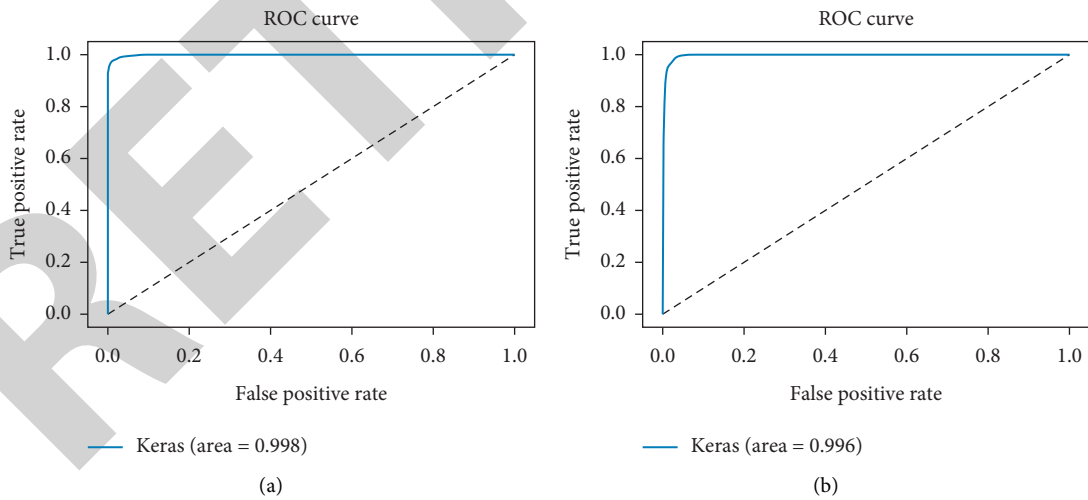


FIGURE 10: Continued.

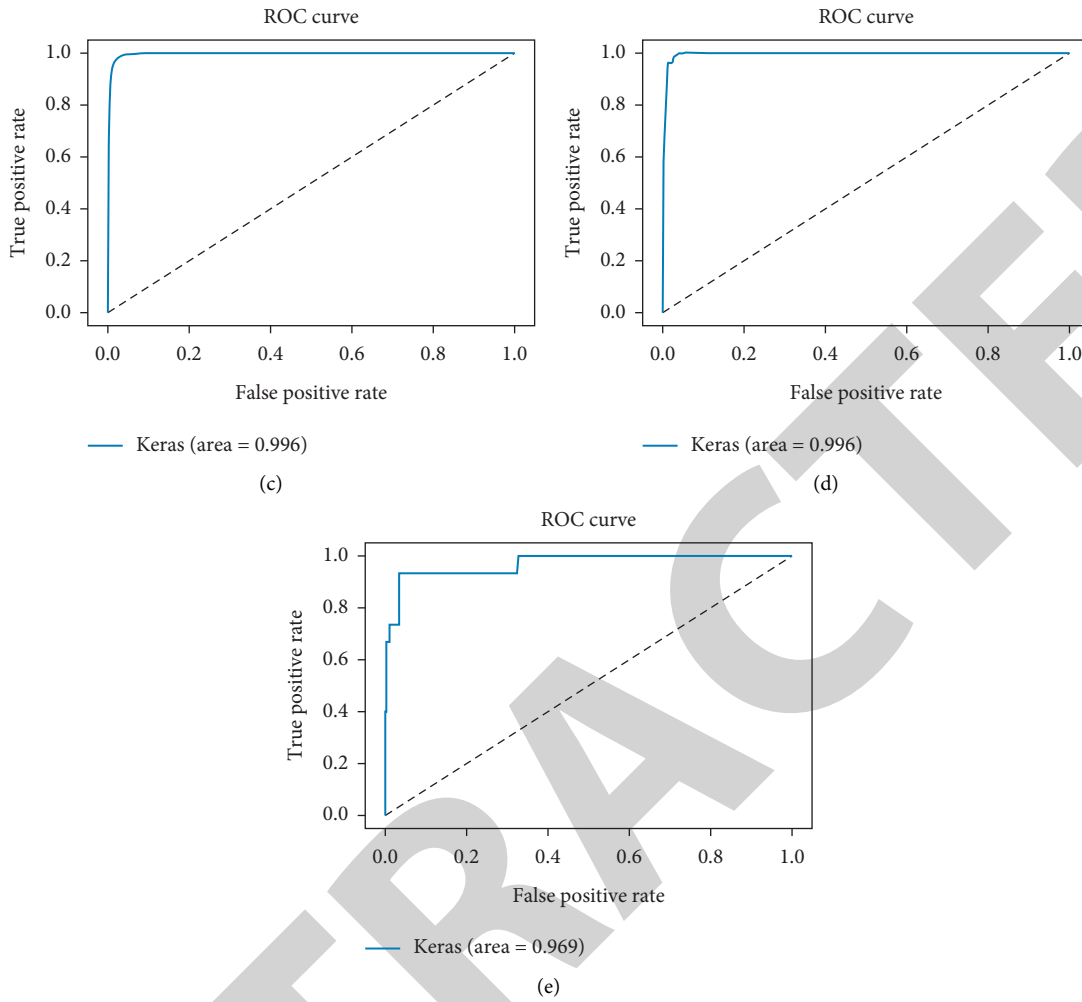


FIGURE 10: ORC of the LSTM model for multiple classification: (a) normal, (b) DoS attack, (c) probe attack, (d) R2L attack, and (e) U2R attack.

TABLE 13: Comparison results of the proposed system against existing security system using artificial intelligence approaches

Ref.	Model	Datasets	Accuracy %
[56]	SAAE-DNN	NSL-KDD test	87.74%
[57]	ICVAE-DNN	NSL-KDD test	85.97%
[58]	Bagging	NSL-KDD test	90.41%
[59]	GAR-forest	NSL-KDD test	90%
Proposed system	LSTM	NSL-KDD test	97.07% with multiple classification and 97.82 with binary classification

The comparison of the classification results of the proposed system against existing security system using artificial intelligence approaches is presented in Table 13.

Overall, the proposed system has achieved highest accuracy than eastings systems (97.07%) by using binary classification, whereas the proposed system with multiple classes has achieved 97.82%.

6. Conclusion

Considering that Web-based businesses manage exceeding amounts of data and business-related secrets, it is necessary

to conduct system movement examinations to achieve appropriate data security.

Therefore, there is a need to develop a smart system to protect IoT networks. Machine learning and deep learning are strategies to detect attacks intelligently. Various machine learning algorithms, namely, QSVM, KNN, linear discriminant, and quadratic discriminant algorithms, were applied, and deep learning algorithms, namely, LSTM and DAE algorithms, were proposed to detect intrusion.

The KDD Cup dataset was employed to test the various machine learning and deep learning algorithms. This dataset has various types of attacks and normal packets. The one-hot

encoding method was used to convert four categorical features into numerical features. The dataset has 41 features for consuming training time and improving the performance of the proposed system. The correlation methods were used to select significant features based on high percentage relationships with classes. These selection features were normalized using the min-max normalization method for scaling the data in the same range, which can help to increase the accuracy.

Machine learning and deep learning algorithms were tested with two databases, namely, binary and multiple classifications. Empirical results showed that the KNN and LSTM models achieved high accuracy in binary and multiple classifications. This study offers a comprehensive summary of the proposed algorithms and gives useful insights into the appropriate machine learning and deep learning models for detecting intrusions in IoT systems and any network. The hybrid CNN-LSTM model will be proposed for improving accuracy of the proposed system.[57], [58], [59].

Data Availability

The data presented in this study are available at <https://www.unb.ca/cic/datasets/nsl.html>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Faisal University for funding this research work through the project no. NA00078.

References

- [1] V. Adat and B. B. Gupta, "Security in internet of things: issues, challenges, taxonomy, and architecture," *Telecommunication Systems*, vol. 67, no. 3, pp. 423–441, 2018.
- [2] Statista Research Department, "IoT: number of connected devices worldwide 2012–2025. Available online," 2020), <https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/>.
- [3] S. Hajiheidari, K. Wakil, M. Badri, and N. J. Navimipour, "Intrusion detection systems in the Internet of things: a comprehensive investigation," *Computer Networks*, vol. 160, pp. 165–191, 2019.
- [4] M. Samaila, M. Neto, D. Fernandes, M. Freire, and P. Inácio, "Challenges of securing internet of things devices: a survey," *Security Point*, vol. 1, 2018.
- [5] S. Bhattarai and Y. Wang, "End-to-End trust and security for internet of things applications," *Computer*, vol. 51, no. 4, pp. 20–27, 2018.
- [6] E. Spafford and P. A. James, "An information security pioneer," *IEEE Secur. Priv.* vol. 6, 2008.
- [7] M. S. Alnaghes and F. Gebali, "A survey on some currently existing intrusion detection systems for mobile ad hoc networks," in *Proceedings of the Second International Conference on Electrical and Electronics Engineering, Clean Energy and Green Computing (EEECEGC2015)*, vol. Volume 12, Antalya, Turkey, 26–28 May 2015.
- [8] P. B. Hari and S. N. Singh, "Security attacks at MAC and network layer in wireless sensor networks," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 11, no. 12, pp. 82–89, 2019.
- [9] J. Pacheco and S. Hariri, "IoT security framework for smart cyber infrastructures," in *Proceedings of the IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, pp. 242–247, Augsburg, Germany, 12–16 September 2016.
- [10] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [11] A. Torkaman and M. Seyyedi, "Analyzing IoT reference architecture models," *International Journal of Computer Systems Science and Engineering*, vol. 5, p. 154, 2016.
- [12] M. A. Chaqfeh and N. Mohamed, "Challenges in middleware solutions for the internet of things," in *Proceedings of the 2012 International Conference on Collaboration Technologies And Systems (CTS)*, pp. 21–26, Denver, CO, USA, 21–25 May 2012.
- [13] N. Moustafa, G. Creech, E. Sitnikova, and M. Keshk, "Collaborative anomaly detection framework for handling big data of cloud computing," in *Proceedings of the 2017 Military Communications and Information Systems Conference (MilCIS)*, p. 1, Canberra, Australia, 14–16 November 2017.
- [14] N. Moustafa, K.-K. R. Choo, I. Radwan, and S. Camtepe, "Outlier Dirichlet mixture mechanism: adversarial statistical learning for anomaly detection in the fog," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 1975–1987, 2019.
- [15] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, "Security, privacy and trust in Internet of Things: the road ahead," *Computer Networks*, vol. 76, pp. 146–164, 2015.
- [16] C. Koliass, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," *Computer*, vol. 50, no. 7, pp. 80–84, 2017.
- [17] M. A. Al-Garadi, A. Mohamed, A. Al-Ali, X. Du, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," arXiv 2018, arXiv: 1807.11023.
- [18] C. Koliass, A. Stavrou, J. Voas, I. Bojanova, and R. Kuhn, "Learning internet-of-things security "Hands-On"," *IEEE Security & Privacy*, vol. 14, no. 1, pp. 37–46, 2016.
- [19] T. Marsden, N. Moustafa, E. Sitnikova, and G. Creech, "Probability risk identification based intrusion detection system for SCADA systems," in *Proceedings of the International Conference on Mobile Networks and Management*, pp. 353–363, Springer, Berlin, Germany, November 2017.
- [20] N. Moustafa, G. Misra, and J. Slay, "Generalized outlier Gaussian mixture technique based on automated association features for simulating and detecting web application attacks," *IEEE Trans. Sustain. Comput.*, 2018.
- [21] C. Modi, D. Patel, B. Borisaniya, H. Patel, A. Patel, and M. Rajarajan, "A survey of intrusion detection techniques in cloud," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 42–57, 2013.
- [22] R. Rizwan, F. A. Khan, H. Abbas, and S. H. Chauhdary, "Anomaly detection in wireless sensor networks using immune-based bioinspired mechanism," *International Journal of Distributed Sensor Networks*, vol. 11, Article ID 684952, 2015.
- [23] I. Butun, S. D. Morgera, and R. Sankar, "A survey of intrusion detection systems in wireless sensor networks," *IEEE Commun. Surv. Tutor.* vol. 16, pp. 266–282, 2013.

- [24] R. Mitchell and I. R. Chen, "A survey of intrusion detection techniques for cyber-physical systems," *ACM Computing Surveys*, vol. 46, p. 55, 2014.
- [25] A. Mishra, K. Nadkarni, and A. Patcha, "Intrusion detection in wireless ad hoc networks," *IEEE Wireless Communications*, vol. 11, no. 1, pp. 48–60, 2004.
- [26] T. Anantvalee and J. Wu, "A survey on intrusion detection in mobile ad hoc networks," in *Wireless Network Security*, pp. 159–180, Springer, Berlin, Germany, 2007.
- [27] S. Kumar and K. Dutta, "Intrusion detection in mobile ad hoc networks: techniques, systems, and future challenges," *Security and Communication Networks*, vol. 9, no. 14, pp. 2484–2556, 2016.
- [28] A. Riahi Sfar, E. Natalizio, Y. Challal, and Z. Chtourou, "A roadmap for security challenges in the Internet of Things," *Digital Communications and Networks*, vol. 4, no. 2, pp. 118–137, 2018.
- [29] M. Keshk, N. Moustafa, E. Sitnikova, and G. Creech, "Privacy preservation intrusion detection technique for SCADA systems," in *Proceedings of the 2017 Military Communications and Information Systems Conference (MilCIS)*, pp. 1–6, Canberra, Australia, 14–16 November 2017.
- [30] K. Zhao and L. Ge, "A survey on the internet of things security," in *Proceedings of the 2013 Ninth International Conference on Computational Intelligence and Security*, pp. 663–667, Leshan, China, 14–15 December 2013.
- [31] J. S. Kumar and D. R. Patel, "A survey on internet of things: security and privacy issues," *International Journal of Computer Application*, vol. 90, 2014.
- [32] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the internet of things: a review," in *Proceedings of the 2012 International Conference on Computer Science and Electronics Engineering*, vol. 3, pp. 648–651, Hangzhou, China, 23–25 March 2012.
- [33] D. E. Kouicem, A. Bouabdallah, and H. Lakhlef, "Internet of things security: a top-down survey," *Computer Networks*, vol. 141, pp. 199–221, 2018.
- [34] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surv. Tutor.* vol. 18, pp. 1153–1176, 2015.
- [35] P. Mishra, V. Varadharajan, U. Tupakula, and E. S. Pilli, "A detailed investigation and analysis of using machine learning techniques for intrusion detection," *IEEE Commun. Surv. Tutor.* vol. 21, pp. 686–728, 2018.
- [36] S. Garg, K. Kaur, S. Batra, G. Kaddoum, N. Kumar, and A. Boukerche, "A multi-stage anomaly detection scheme for augmenting the security in IoT-enabled applications," *Future Generation Computer Systems*, vol. 104, pp. 105–118, 2020.
- [37] S. Garg, K. Kaur, N. Kumar, G. Kaddoum, A. Y. Zomaya, and R. Ranjan, "A hybrid deep learning-based model for anomaly detection in cloud datacenter networks," *IEEE Transactions on Network and Service Management*, vol. 16, no. 3, pp. 924–935, 2019.
- [38] Y. Mirsky, T. Doitshman, Y. Elovici, and S. ., A. Kitsune, "An ensemble of autoencoders for online network intrusion detection," arXiv 2018, arXiv:1802.09089.
- [39] J. McHugh, "Testing Intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory," *ACM Transactions on Information and System Security*, vol. 3, pp. 262–294, 2000.
- [40] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [41] Y. Bengio, "Learning deep architectures for ai," *Found. Trends Mach. Learn.* vol. 2, no. 1, pp. 1–127, 2009.
- [42] L. Yann and B. Yoshua, "Convolutional networks for images, speech, and time-series," *Handb. Brain Theory Neural Netw.*, vol. 10, pp. 2571–2575, 1995.
- [43] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: a comprehensive review," *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [44] S. N. Alsubari, S. N. Deshmukh, M. H. Al-Adhaileh, F. W. Alsaade, and T. H. Aldhyani, "Development of integrated neural network model for identification of fake reviews in E-commerce using multidomain datasets," *Applied Bionics and Biomechanics*, vol. 2021, Article ID 5522574, 11 pages, 2021.
- [45] Y. Liu, L. Guan, C. Hou et al., "Wind power short-term prediction based on LSTM and discrete wavelet transform," *Applied Sciences*, vol. 9, no. 6, p. 1108, 2019.
- [46] S. N. Alsubari, S. N. Deshmukh, A. A. Alqarni et al., "Data analytics for the identification of fake reviews using supervised learning," *CMC-Computers, Materials & Continua*, vol. 70, 2022.
- [47] H. Alkahtani, T. H. Aldhyani, and M. Al-Yaari, "Adaptive anomaly detection framework model objects in cyberspace," *Applied Bionics and Biomechanics*, vol. 2020, Article ID 6660489, 14 pages, 2020.
- [48] A. Mihoub, O. B. Fredj, O. Cheikhrouhou, A. Derhab, and M. Krichen, "Denial of service attack detection and mitigation for internet of things using looking-back-enabled machine learning techniques," *Computers & Electrical Engineering*, vol. 98, Article ID 107716, 2022.
- [49] I. Jemal, M. A. Haddar, O. Cheikhrouhou, A. Mahfoudhi, and 'M-Cnn, "A new hybrid deep learning model for web security," in *Proceedings of the 2020 IEEE/ACS 17th International Conference on Computer Systems and Applications (AICCSA)*, pp. 1–7, Antalya, Turkey, 2020.
- [50] S. Swati, G. Isha, G. Sheifali et al., "Automated detection of diseases from apple leaf images," *CMC-Computers, Materials & Continua*, vol. 71, no. 1, pp. 1849–1866, 2022.
- [51] I. Jemal, M. A. Haddar, O. Cheikhrouhou, and A. Mahfoudhi, "Malicious http request detection using code-level convolutional neural network," *Revised Selected Papers in Proceedings of the Risks and Security of Internet and Systems: 15th International Conference, CRiSIS 2020*, vol. 15, pp. 317–324, Paris, France, November 4–6, 2020.
- [52] H. Alkahtani, H. Theyazn, and H. Aldhyani, "Botnet attack detection by using CNN-LSTM model for internet of things applications," *Security and Communication Networks*, vol. 2021, Article ID 3806459, 23 pages, 2021.
- [53] H. Tahaei, F. Afifi, A. Asemi, F. Zaki, and N. B. Anuar, "The rise of traffic classification in IoT networks: A survey," *Journal of Network and Computer Applications*, p. 154, 2020.
- [54] T. H. H. Aldhyani and H. Alkahtani, "Attacks to automatous vehicles: a deep learning algorithm for cybersecurity," *Sensors*, vol. 22, no. 1, p. 360, 2022.
- [55] V. Anand, S. Gupta, D. Koundal, S. Mahajan, A. Kant Pandit, and A. Zaguia, "Deep learning based automated diagnosis of skin diseases using dermoscopy," *Computers, Materials & Continua*, vol. 71, no. 2, pp. 3145–3160, 2022.
- [56] C. Tang, N. Luktarhan, and Y. Zhao, "SAAE-DNN: deep learning method on intrusion detection," *Symmetry*, vol. 12, no. 10, p. 1695, 2020.
- [57] Y. Yang, K. Zheng, C. Wu, and Y. Yang, "Improving the classification effectiveness of intrusion detection by using

Retraction

Retracted: A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] H. Zhao, X. Dong, T. Huang, and X. Li, "A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9979200, 8 pages, 2022.

Research Article

A Potential Prognostic Biomarker for Glioma: Aldo-Keto Reductase Family 1 Member B1

Hulin Zhao,¹ Xuetao Dong,² Tianxiang Huang ,³ and Xueji Li ⁴

¹Department of Neurosurgery, Chinese PLA General Hospital, Beijing 100853, China

²Department of Neurosurgery, Chuiyangliu Hospital Affiliated To Tsinghua University, Beijing, China

³Department of Neurosurgery, Xiangya Hospital, Central South University, Changsha, China

⁴Department of Neurosurgery, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Correspondence should be addressed to Tianxiang Huang; huang966931guo7533@163.com and Xueji Li; lixueji@cumt.edu.cn

Received 20 January 2022; Accepted 23 February 2022; Published 18 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Hulin Zhao 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.

Aldo-keto reductase family 1 member B1 (AKR1B1) plays a vital role in tumor development and is involved in the tumor immune process. However, its role in glioma cell is poorly studied. This study's aim was to assess the role of AKR1B1 in glioma through bioinformatics analysis. The AKR1B1 expression data and corresponding clinical data of glioma were collected from the Cancer Genome Atlas (TCGA) database. The R packages were used for data integration, extraction, analysis, and visualization. According to the median value of the risk score, all patients were divided into high-risk and low-risk groups to draw the Kaplan–Meier (KM) survival curves and to explore the level of immune infiltration. The expression of AKR1B1 was significantly elevated in glioma tissues compared to normal tissues ($P < 0.001$). The high expression of AKR1B1 was significantly associated with WHO grade ($P < 0.001$), IDH status ($P < 0.001$), 1p/19q codeletion ($P < 0.001$), primary therapy outcome ($P = 0.004$), and age ($P < 0.05$). Kaplan–Meier survival analysis found that OS (HR = 3.75, $P < 0.001$), DSS (HR = 3.85, $P < 0.001$), and PFI (HR = 2.76, $P < 0.001$) were lower in patients with glioma with high AKR1B1 expression than in the group with low AKR1B1 expression. Based on GESA, six pathways (including interferon gamma signaling, signaling by interleukins, cell cycle checkpoints, cytokine receptor interaction, cell adhesion molecules (CAMs), and cell surface interactions) at the vascular wall were identified as significantly different between the two groups. Moreover, highly expressed AKR1B1 was associated with immune cell infiltration. AKR1B1 plays a key role in glioma progression and prognosis and, therefore, serves as a potential biomarker for prediction of patients' survival.

1. Introduction

Glioma is one of the most common classes of brain tumors, comprising 80% of malignant primary brain tumors [1]. Glioma is the most common and aggressive cancer of the neuroepithelial tissue, with an average survival time of only 15 months [2]. Gliomas are highly heterogeneous with multiple genetically distinct clones which makes it difficult to successfully implement clinical treatments, even with the many new therapies now being investigated [3]. Moreover, even with improved surgical techniques and postoperative radiotherapy, the average survival time of glioma patients after neurosurgical radiotherapy is still limited to around 12

months, and the factors affecting the prognosis of glioma patients are complex. The lack of appropriate screening and diagnostic biomarkers has led to failure in the early diagnosis of glioma. Although a variety of biomarkers have been used in the diagnosis of glioma, their reliability remains controversial [4]. Therefore, the discovery of biomarkers relevant to tumor staging and prognosis is extremely important to facilitate the early diagnosis, prognosis, and treatment of glioma.

AKR1B1, aldehyde reductase (AR) protein, is one of the key members of the superfamily of aldehyde reductases (AKRs) [5]. It is an NADPH-dependent, multifunctional enzyme class with the spatial structure of a single-chain

polypeptide containing a sulfhydryl group, which is generally present in organisms as a monomer [6]. AKRs have been implicated in the development of various tumors in humans. Previous study has reported that overexpression of AKR1B1 was found to promote the development and metastasis of breast cancer, suggesting that the overexpression of AKR1B1 could predict poor prognosis in breast cancer patients [7]. In addition, overexpression of AKR1B1 could decrease the expression of β 2-adrenergic receptor (β 2-AR) and increase the expression of p-ERK1/2, thereby promoting the proliferation and metastasis of pancreatic cancer [8]. Moreover, many members of the aldose reductase superfamily have been used as diagnostic markers for tumors in the detection of cancers [9, 10]. There are also members that are potential diagnostic markers for cancers. However, the association of AKR1B1 expression and clinicopathological characteristics, as well as prognosis of glioma patients with OS, has not yet been characterized.

In our study, we tried to determine the association between AKR1B1 and glioma and to explore the prognostic role of AKR1B1 in glioma based on RNA-sequencing data from TCGA. We used the information of glioma patients in TCGA database to analyze the correlation between MIR4435-2HG and clinicopathological features, prognosis, and immune cell infiltration and then selected several possible signaling pathways for further study to help future clinical treatment.

2. Materials and Methods

2.1. Data Collection, Screening, and Analysis. UCSC Xena (<https://xenabrowser.net/datapages/>) RNAseq data in TPM format for TCGA and GTEx were harmonized by the Toil process [11]. Data were extracted from the GBMLGG of TCGA (1157 cases of glioma) and the corresponding normal tissues in GTEx (689 cases). The gene expression data were divided into high and low expression groups according to the median expression of AKR1B1.

2.2. Gene Enrichment Analysis. GSEA was used for analysis, and the c2.cp.v7.2.symbols.gmt dataset was obtained using the gene enrichment analysis website MSigDB database. Enrichment analysis was performed according to the default weighted enrichment statistics, with the number of random combinations set to 1000. The expression level of AKR1B1 was used as the expression marker. Enrichment pathways were ranked for each phenotype using P values and normalized enrichment scores (NES).

2.3. Data Analysis. The R package (v.3.6.1) was used for all statistical analyses. The relationship between AKR1B1 and clinicopathological characteristics was analyzed using the Wilcoxon rank sum test. Cox regression and Kaplan–Meier survival analysis were used to analyze the relationship between clinicopathological characteristics and OS, DSS, and PFI in patients with glioma. $P < 0.05$ was considered a statistically significant difference.

3. Results

3.1. Clinical Characteristics. We collected clinicopathological characteristics of glioma patients from TCGA database, including WHO grade, IDH status, 1p/19q codeletion, primary therapy outcome, gender, race, and age (shown in Table 1). Patients with glioma were divided into high and low-AKR1B1 expression groups according to the median AKR1B1 expression. We subsequently analyzed the relationship between AKR1B1 expression and clinical characteristics by the chi-square test and Wilcoxon rank sum test. The results showed that high levels of AKR1B1 expression were associated with WHO grade ($P < 0.001$), IDH status ($P < 0.001$), 1p/19q codeletion ($P < 0.001$), primary therapy outcome ($P = 0.004$), and age ($P < 0.05$) but not gender and race.

3.2. The Expression of AKR1B1 in Glioma Tissues and Paraneoplastic Tissues. Analysis of 1157 cases of paraneoplastic and 689 cases of glioma tissues in TCGA database revealed that AKR1B1 expression levels were significantly higher in tumor tissues than in normal paraneoplastic tissues ($P < 0.001$, Figure 1(a)). In addition, AKR1B1 overexpression showed promising discriminative power with an AUC value of 0.935 to distinguish tumors from normal tissues (Figure 1(b)).

3.3. Associations between AKR1B1 Expression and Clinicopathologic Variables. Based on the analysis of glioma patients from TCGA database, the Kruskal–Wallis test showed that AKR1B1 overexpression was significantly associated with histological type (Figure 2(a)). The Wilcoxon rank sum test showed that AKR1B1 overexpression was significantly associated with 1p/19q codeletion and IDH status (Figures 2(d)–2(e)). Also, independent sample t -test found that AKR1B1 overexpression was significantly correlated with age (Figure 2(f)).

3.4. The Relationship between AKR1B1 Expression and Glioma Prognosis. To assess the value of AKR1B1 in predicting the prognosis of glioma patients, we analyzed the relationship between AKR1B1 expression and OS, DSS, and PFI. As shown in Figure 3(a), the prognosis of glioma patients in the high AKR1B1 expression group was worse (HR = 3.75 (2.87–4.80), $P < 0.001$) (Figure 3(a)). We further analyzed the DSS and PFI by Kaplan–Meier plotter database and found that both DSS and PFI were lower in patients with glioma with high AKR1B1 expression than in the group with low AKR1B1 expression (Figures 3(b)–3(c)). The altered results suggested that patients with overexpressed AKR1B1 have a worse prognosis.

3.5. AKR1B1-Related Signaling Pathways Based on GSEA. We used GSEA to perform enrichment analysis of high AKR1B1 expression to identify signaling pathways activated in glioma. As shown in Figure 4, six pathways, including interferon gamma signaling, signaling by interleukins, cell

TABLE 1: The characteristics of patients with glioma based on TCGA.

Characteristic	Low expression of AKR1B1	High expression of AKR1B1	P
<i>n</i>	348	348	
WHO grade, <i>n</i> (%)			<0.001
G2	160 (25.2%)	64 (10.1%)	
G3	132 (20.8%)	111 (17.5%)	
G4	17 (2.7%)	151 (23.8%)	
IDH status, <i>n</i> (%)			<0.001
WT	61 (8.9%)	185 (27%)	
Mut	284 (41.4%)	156 (22.7%)	
1p/19q codeletion, <i>n</i> (%)			<0.001
Codel	133 (19.3%)	38 (5.5%)	
Non-codel	215 (31.2%)	303 (44%)	
Primary therapy outcome, <i>n</i> (%)			0.004
PD	55 (11.9%)	57 (12.3%)	
SD	100 (21.6%)	47 (10.2%)	
PR	42 (9.1%)	22 (4.8%)	
CR	96 (20.8%)	43 (9.3%)	
Gender, <i>n</i> (%)			0.146
Female	139 (20%)	159 (22.8%)	
Male	209 (30%)	189 (27.2%)	
Race, <i>n</i> (%)			0.383
Asian	4 (0.6%)	9 (1.3%)	
Black or African American	16 (2.3%)	17 (2.5%)	
White	319 (46.7%)	318 (46.6%)	
Age, <i>n</i> (%)			<0.001
≤60	305 (43.8%)	248 (35.6%)	
>60	43 (6.2%)	100 (14.4%)	
Histological type, <i>n</i> (%)			<0.001
Astrocytoma	110 (15.8%)	85 (12.2%)	
Glioblastoma	17 (2.4%)	151 (21.7%)	
Oligoastrocytoma	81 (11.6%)	53 (7.6%)	
Oligodendroglioma	140 (20.1%)	59 (8.5%)	
Age, median (IQR)	40 (32, 53)	51 (38, 62)	<0.001

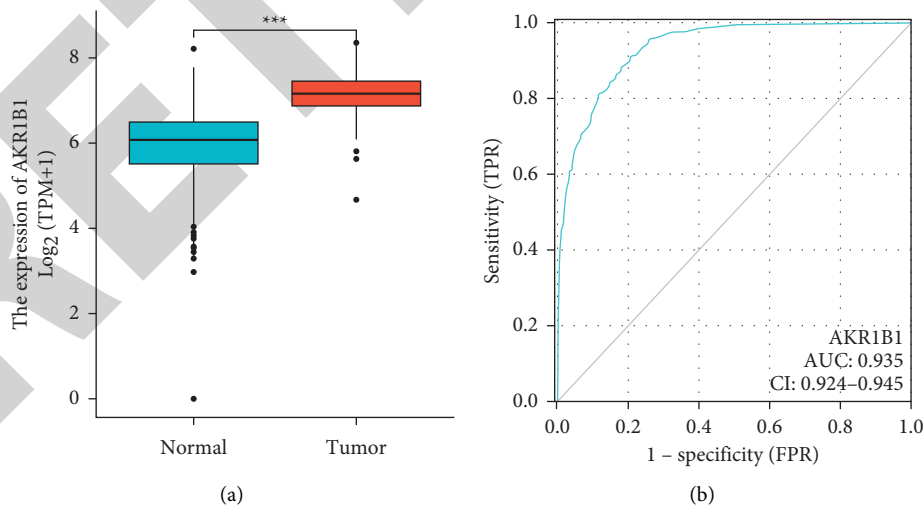


FIGURE 1: AKR1B1 expression between cancer and normal tissues in glioma patients. (a) AKR1B1 expression levels in glioma and normal tissues. (b) ROC analysis of AKR1B1 showed promising discrimination power between tumor and normal tissues.

cycle checkpoints, cytokine receptor interaction, cell adhesion molecules (CAMs), and cell surface interactions at the vascular wall, were identified as significantly different between the two groups.

3.6. *Correlation between Immune Cell Infiltration and AKR1B1.* We analyzed the correlation between AKR1B1 expression level and the degree of immune cell enrichment based on Spearman's correlation coefficient. AKR1B1

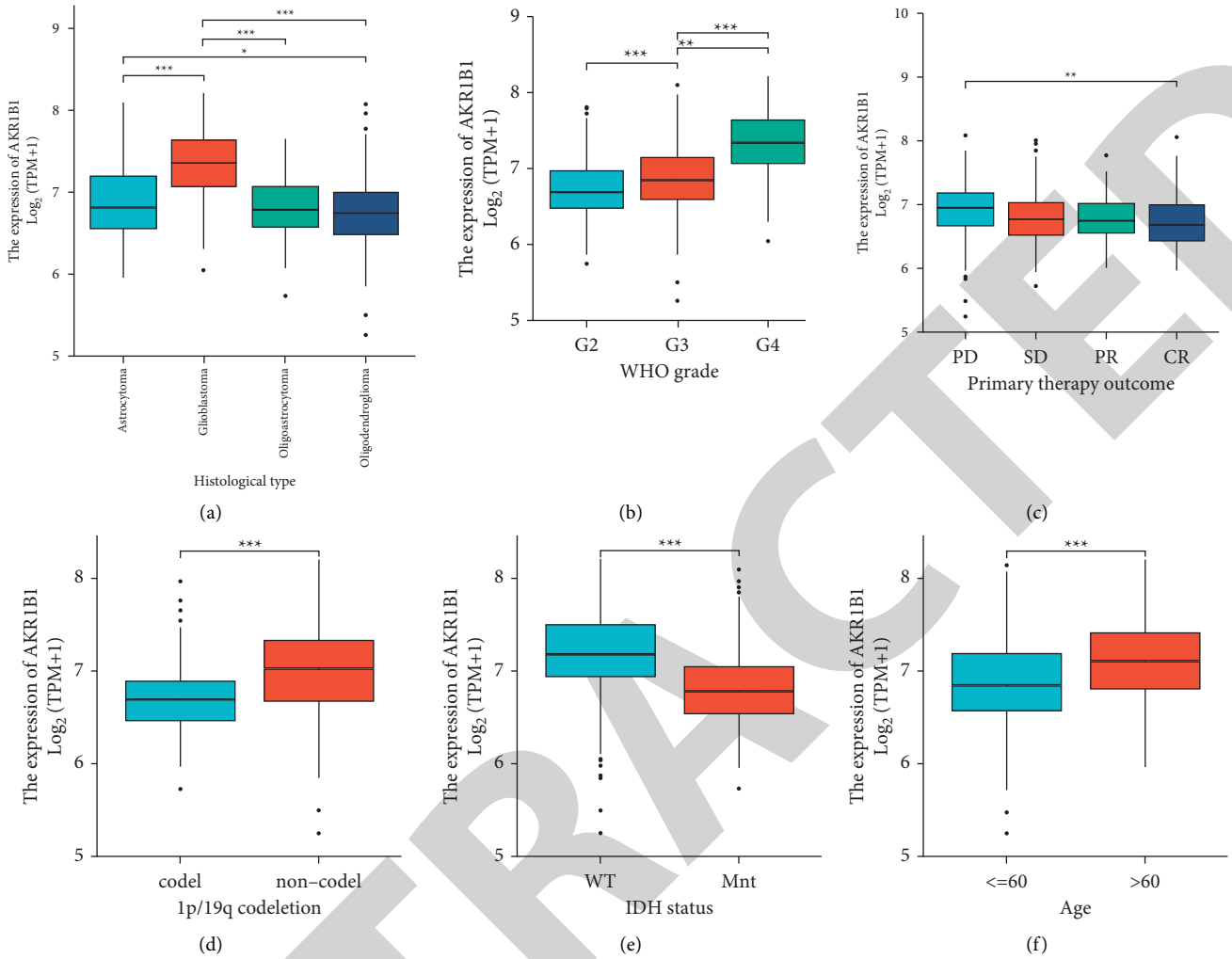


FIGURE 2: The correction of AKR1B1 expression with clinicopathologic characteristics. (a) Histological type. (b) WHO grade. (c) Primary therapy outcome. (d) 1q/19q codeletion. (e) IDH status. (f) Age.

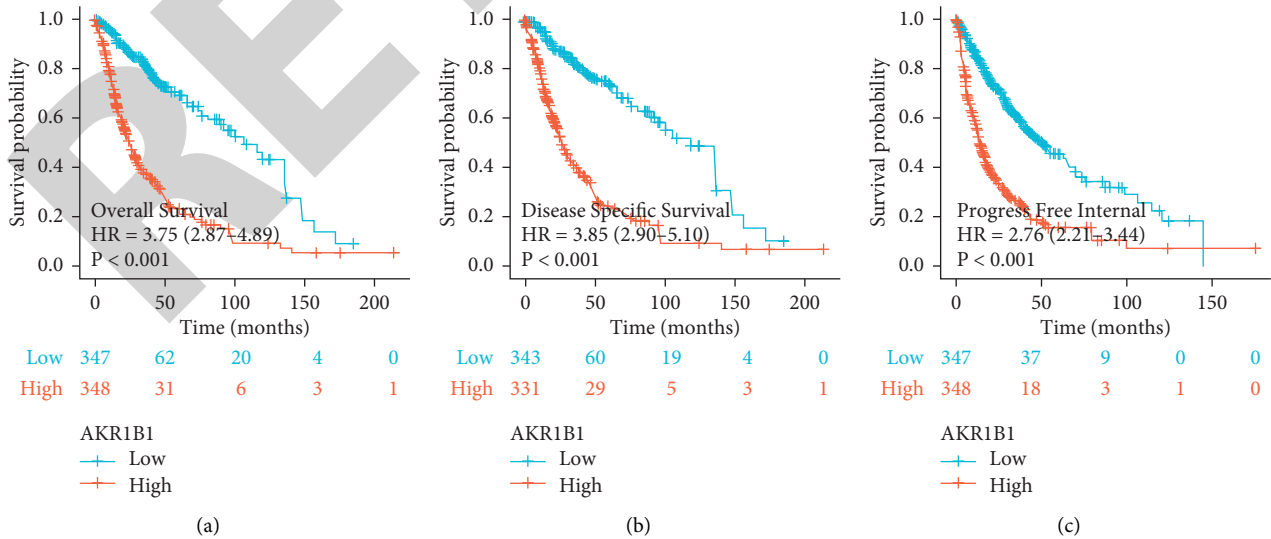


FIGURE 3: Kaplan-Meier survival curves comparing the high expression and low expression of AKR1B1 in glioma. (a) Overall survival. (b) Disease-specific survival. (c) Progression-free survival.

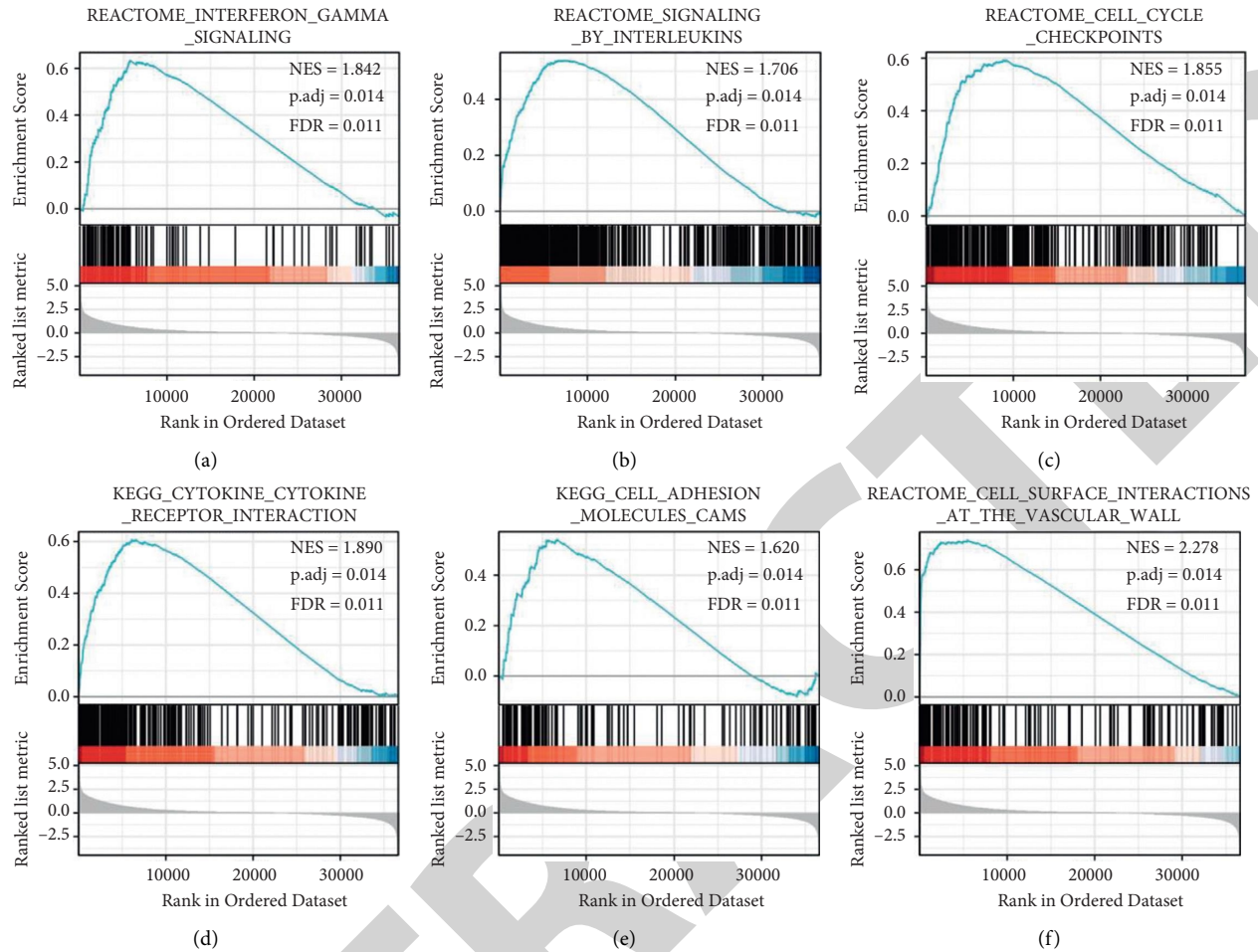


FIGURE 4: Enrichment plots from GSEA. (a) Interferon gamma. (b) Signaling by interleukins. (c) Cell cycle checkpoints. (d) Cytokine receptor interaction. (e) Cell adhesion molecules (CAMs). (f) Cell surface interactions at the vascular wall.

expression levels were negatively correlated with NK CD56bright cells, pDC, Tcm, and Tgd, while they were positively related with the most abundance of macrophages (Figure 5(a)). Moreover, overexpression of AKR1B1 in macrophages was correlated (Figure 5(b)). Therefore, we further analyzed the correlation between AKR1B1 and marker CD163 and TGFB1 in M2 macrophages (Figures 5(c) and 5(d)). The results revealed that AKR1B1 expression levels were significantly and negatively correlated with marker CD163 ($P < 0.001$, $r = 0.520$), TGFB1 ($P < 0.001$, $r = 0.500$), and VSIG4 ($P < 0.001$, $r = 0.520$) in M2 macrophages.

4. Discussion

According to the Central Brain Tumor Registry of the United States (CBTRUS), approximately 2% of adult tumors are brain tumors, and up to 80% of these are gliomas. The current surgical treatment of glioma, radiotherapy, chemotherapy, and molecular targeted therapy have achieved considerable success, but the recurrence rate of surgical treatment is still very high [12]. Radiotherapy and chemotherapy are prone to resistance and metastasis with a poor

prognosis [13]. Therefore, there is an urgent need for new studies to discover more effective diagnostic criteria and diagnostic methods. The literature reported that silencing NORAD inhibited the proliferation, invasion, and migration of glioma cells, while overexpression of AKR1B1 reversed the effects of silencing NORAD on glioma cells [14]. This suggested that AKR1B was involved in the proliferation and progression of glioma cells, but its role and significance in the clinical setting have rarely been investigated, so this study focused on the clinical aspects.

AKR1B1 protein is an important member of the aldo-keto reductase superfamily and has been found to be closely associated with the development of many human tumors [15]. However, the effect of AKR1B1 on tumors may vary depending on the stage, type, and aggressiveness of the tumor. For example, in colorectal cancer, AKR1B1 levels were higher in metastatic SW620 cells than in non-metastatic SW480 [16]. In addition, inhibition of AKR1B1 inhibited the proliferation, migration, and invasion of colon cancer cells, which in turn impeded liver metastasis of CRC [17]. In the literature, knockdown of AKR1B1 in HeLa cells was found to inhibit proliferation, migration, and invasion of HeLa cells [18]. Another study found that AKR1B1 was overexpressed in

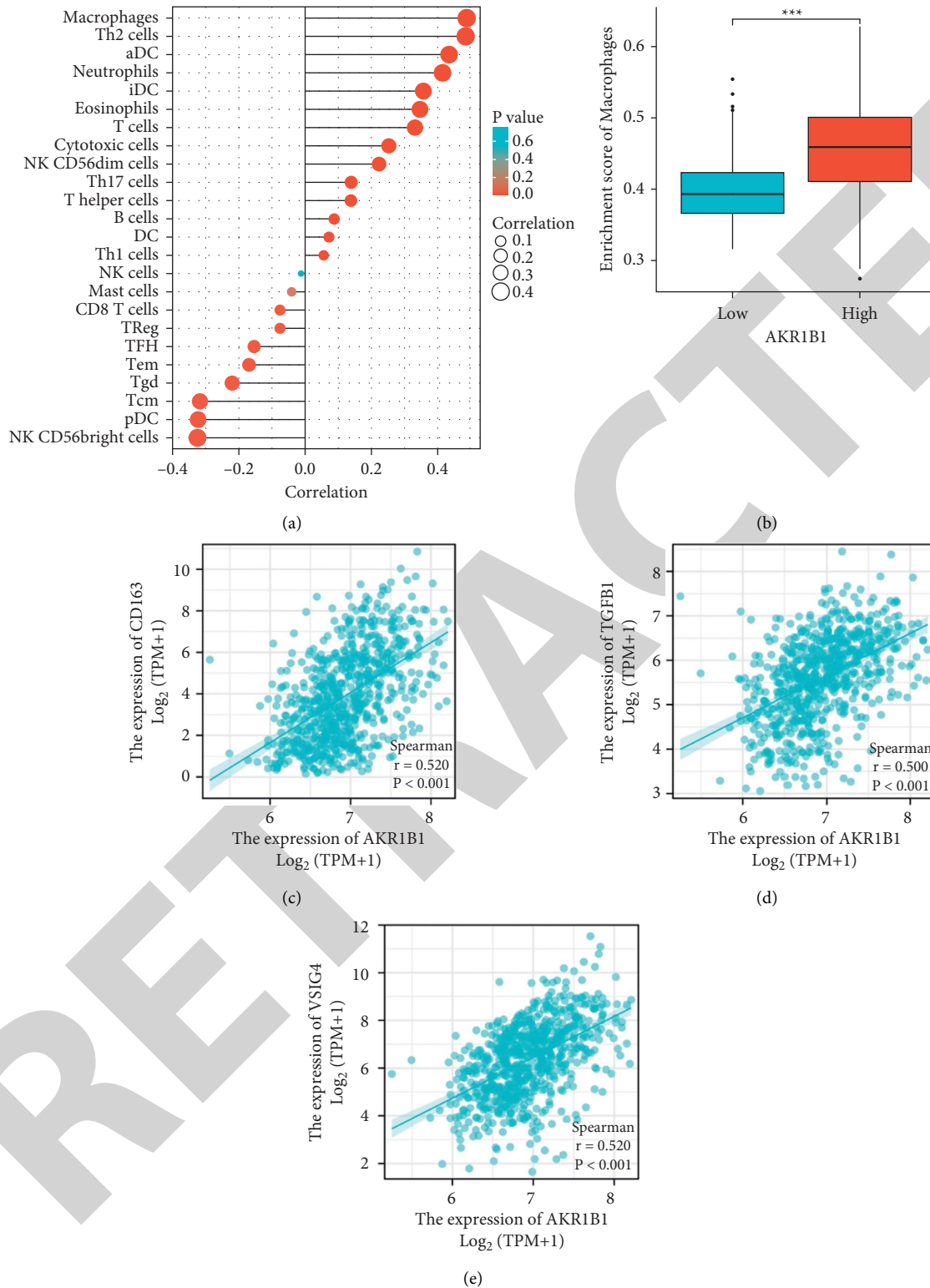


FIGURE 5: Association between immune cell infiltration and AKR1B1 in glioma. (a) Correlation between the relative abundance of 24 immune cells and AKR1B1 expression level. (b) Macrophage infiltration level in the high AKR1B1 expression group and low AKR1B1 expression group in TCGA cohort. (c–e) Correlation between AKR1B1 expression and M2-like macrophage marker.

BLBCs but unchanged in tubular cells [19]. In contrast, some studies found that AKR1B1 was significantly less expressed in breast cancer tissues than in normal paracancerous tissues

[20]. Although the current level of AKR1B1 expression does not indicate its function in breast cancer, AKR1B1 may play an important function in breast cancer and EMT [21].

In this study, we first analyzed glioma-related transcriptomic data from TCGA database and found that AKR1B1 expression was significantly elevated in glioma tissues. Previous literature reported that patients with high AKR1B1 expression had poorer OS, and the prognostic analysis of this study also found that AKR1B1 overexpression predicted poorer OS, DSS, and PFI in glioma patients. We further investigated the function of AKR1B1 in glioma tissues using GSEA and found that interferon gamma signaling, signaling by interleukins, cell cycle checkpoints, cytokine receptor interaction, cell adhesion molecules (CAMs), and cell surface interactions at the vascular wall all showed significantly differential enrichment in high AKR1B1 expression phenotype. The results of the AKR1B1-related signaling pathway enrichment analysis involved several signaling pathways that are closely related to tumors, including interferon gamma signaling, signaling by interleukins, cell cycle checkpoints, cytokine receptor interaction, cell adhesion molecules (CAMs), and cellular surface interactions at the vascular wall. The signaling pathways mentioned above may play a key role in the development and progression of hepatocellular carcinoma. Cell adhesion molecules play a significant role in cancer progression and metastasis [22]. Cell-cell interactions of cancer cells with endothelium determine the metastatic spread. In addition, direct tumor cell interactions with platelets, leukocytes, and soluble components significantly contribute to cancer cell adhesion, extravasation, and the establishment of metastatic lesions. Tumor microenvironment immune cells constitute a key component of tumor tissue [23]. Tumor infiltrating immune cells are an integral component of the tumor microenvironment, and their composition and distribution are thought to be associated with cancer prognosis. Specifically, the level of TAM infiltration accelerates glioma progression. TAM is composed mainly of M2 macrophages and is a complex factor in exposure to the tumor microenvironment [24]. Therefore, we further analyzed the relationship between the expression level of AKR1B and the level of immune cell infiltration in glioma tissues. By analyzing the relationship between AKR1B expression levels and immune cells, we found that an increase in AKR1B expression was significantly and positively correlated with marker of M2 macrophages. The correlation between AKR1B and immune cells suggested that AKR1B played a key role in the regulation of tumor immunity.

There are still several limitations in the current study. One of the major shortcomings of this study is that it only provides ideas for clinical treatment from a genetic perspective, for involving AKR1B1 protein, and the relationship between gene expression and protein expression still needs to be further explored. In addition, to improve the credibility of the results, the sample size should be further expanded to include more clinical factors for a comprehensive evaluation. Also, the phenotypic assays and in vivo animal experiments for the validation of our findings are necessary in the future.

5. Conclusion

In conclusion, our study found that overexpression of AKR1B is associated with poor prognosis in glioma and is considered to be an independent factor in patients with

glioma. AKR1B may play an important role in immune cell infiltration and may represent a valuable prognostic biomarker for glioma.

Data Availability

The datasets used and analyzed during the current study are available from the corresponding authors on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] R. Inano, N. Oishi, T. Kunieda et al., "Voxel-based clustered imaging by multiparameter diffusion tensor images for glioma grading," *NeuroImage: Clinic*, vol. 5, pp. 396–407, 2014.
- [2] T.-Y. Chen, Y. Liu, L. Chen, J. Luo, C. Zhang, and X.-F. Shen, "Identification of the potential biomarkers in patients with glioma: a weighted gene co-expression network analysis," *Carcinogenesis*, vol. 41, no. 6, pp. 743–750, 2020.
- [3] S. M. D. Rizvi, T. Hussain, K. Mehmood, A. Moin, A. S. Alanazi, and G. M. Subaiea, "Molecular docking and dynamic simulation study to explore quercetin as a multipotent candidate against gliomas," *Tropical Journal of Pharmaceutical Research*, vol. 20, no. 4, pp. 815–823, 2021.
- [4] F. H. Hochberg, N. A. Atai, D. Gonda et al., "Glioma diagnostics and biomarkers: an ongoing challenge in the field of medicine and science," *Expert Review of Molecular Diagnostics*, vol. 14, no. 4, pp. 439–452, 2014.
- [5] S. P. Baba, K. Wetzelberger, J. D. Hoetker, and A. Bhatnagar, "Posttranslational glutathiolation of aldose reductase (AKR1B1): a possible mechanism of protein recovery from S-nitrosylation," *Chemico-Biological Interactions*, vol. 178, no. 1-3, pp. 250–258, 2009.
- [6] F. Balestri, V. Barracco, G. Renzone et al., "Stereoselectivity of aldose reductase in the reduction of glutathionyl-hydroxynonanal adduct," *Antioxidants*, vol. 8, no. 10, p. 502, 2019.
- [7] X. Wu, X. Li, Q. Fu et al., "AKR1B1 promotes basal-like breast cancer progression by a positive feedback loop that activates the EMT program," *Journal of Experimental Medicine*, vol. 214, no. 4, pp. 1065–1079, 2017.
- [8] M.-B. Xiao, D.-D. Jin, Y.-J. Jiao et al., " β 2-AR regulates the expression of AKR1B1 in human pancreatic cancer cells and promotes their proliferation via the ERK1/2 pathway," *Molecular Biology Reports*, vol. 45, no. 6, pp. 1863–1871, 2018.
- [9] M. Hojnik, G. S. Frkovic, I. Verdenik, and T. L. Rizner, "AKR1B1 and AKR1B10 as prognostic biomarkers of endometrioid endometrial carcinomas," *Cancers*, vol. 13, no. 14, 2021.
- [10] E. S. Kropotova, R. A. Tychko, O. L. Zinov'Eva et al., "[Downregulation of AKR1B10 gene expression in colorectal cancer]," *Molecular Biology (Moscow)*, vol. 44, no. 2, pp. 243–250, 2010.
- [11] J. Vivian, A. A. Rao, F. A. Nothaft et al., "Toil enables reproducible, open source, big biomedical data analyses," *Nature Biotechnology*, vol. 35, no. 4, pp. 314–316, 2017.
- [12] B. Luo and J. Zhang, "MicroRNA-16 inhibits the migration and invasion of glioma cell by targeting Bcl-2 gene," *Tropical Journal of Pharmaceutical Research*, vol. 19, no. 12, pp. 2499–2504, 2020.

Retraction

Retracted: Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] V. K. Gunjan, Y. Vijayalata, S. Valli, S. Kumar, M. O. Mohamed, and V. Saravanan, "Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3604113, 10 pages, 2022.

Research Article

Machine Learning and Cloud-Based Knowledge Graphs to Recognize Suicidal Mental Tendencies

Vinit Kumar Gunjan ¹, Y. Vijayalata ², Susmitha Valli ², Sumit Kumar ³,
M. O. Mohamed ⁴ and V. Saravanan ⁵

¹Department of Computer Science and Engineering, CMR Institute of Technology, Hyderabad, India

²Department of CSE, GRIET, Telangana, India

³Indian Institute of Management, Kozhikode, India

⁴Mathematics Department, Zagazig University, Faculty of Science, Egypt

⁵Department of Computer Science, College of Engineering and Technology, Dambi Dollo University, Dambi Dollo, Oromia Region, Ethiopia

Correspondence should be addressed to Vinit Kumar Gunjan; vinitkumargunjan@gmail.com and V. Saravanan; saravanan@dadu.edu.et

Received 31 January 2022; Revised 11 February 2022; Accepted 19 February 2022; Published 17 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Vinit Kumar Gunjan 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.

To improve the quality of knowledge service selection in a cloud manufacturing environment, this paper proposes a cloud manufacturing knowledge service optimization decision method based on users' psychological behavior. Based on the characteristic analysis of cloud manufacturing knowledge service, establish the optimal evaluation index system of cloud manufacturing knowledge service, use the rough set theory to assign initial weights to each evaluation index, and adjust the initial weights according to the user's multiattribute preference to ensure that the consequences are allocated correctly. The system can help counselors acquire psychological knowledge in time and identify counselors with suicidal tendencies to prevent danger. This paper collected some psychological information data and built a knowledge graph by creating a dictionary and generating entities and relationships. The Han language processing word segmentation tool generates keywords, and CHI (Chi-square) feature selection is used to classify the problem. This feature selection is a statistical premise test that is acceptable when the chi-square test results are distributed with the null hypothesis. It includes the Pearson chi-square test and its variations. The Chi-square test has several benefits, including its distributed processing resilience, ease of computation, broad information gained from the test, usage in research when statistical assumptions are not satisfied, and adaptability in organizing information from multiple or many more group investigations. For improving question and answer efficiency, compared with other models, the BiLSTM (bidirectional long short-term memory) model is preferred to build suicidal tendencies. The Han language processing is a method that is used for word segmentation, and the advantage of this method is that it plays a key role in the word segmentation tool and generates keywords, and CHI (Chi-square) feature selection is used to classify the problem. Text classifier detects dangerous user utterances, question template matching, and answer generation by computing similarity scores. Finally, the system accuracy test is carried out, proving that the system can effectively answer the questions related to psychological counseling. The extensive experiments reveal that the method in this paper's accuracy rate, recall rate, and F1 value is much superior to other standard models in detecting psychological issues.

1. Introduction

Cloud manufacturing (CMfg) is a new service-oriented and knowledge-based networked and agile manufacturing model [1] that integrates advanced technologies, such as existing

information technology, cloud computing [2], and the Internet of Things [3]. It centrally stores the optimized and integrated manufacturing resources in the cloud pool of the manufacturing system and provides users with various high-quality and fast manufacturing services on demand through

the network. As a result, cloud manufacturing knowledge is a dynamic resource that is infiltrating all cloud manufacturing service life cycles [4]. At the same time, cloud manufacturing knowledge service (CMKS) is a knowledge transfer and sharing service that can effectively support cloud manufacturing. As a result, all manufacturing activities in the manufacturing environment are carried out efficiently, stably, and accurately.

The addition of knowledge services significantly improves cloud manufacturing systems' operational efficiency and problem-solving capabilities [5]. The organic combination of CMfg and knowledge services promotes the extension of "manufacturing as a service" to "knowledge as a service," enabling enterprises to quickly obtain the required manufacturing knowledge and services with the support of the cloud platform, effectively solving the problems in the production and operation process of enterprises. Technical bottleneck problem improves its comprehensive competitiveness. It is a statistical premise test that is acceptable when the chi-square test results are distributed with the null hypothesis. It includes the Pearson chi-square test and its variations. The Chi-square test has several benefits, including its distributed processing resilience, ease of computation, broad information gained from the test, usage in research when statistical assumptions are not satisfied, and adaptability in organizing information from multiple or much more group investigations. At present, scholars at home and abroad have explored the relevant theories and technologies of knowledge services in the cloud manufacturing environment. For example, author [6] proposed a personalized knowledge service technology based on user perception of the cloud manufacturing environment. The platform user task requirements and information actively push knowledge resources related to user manufacturing tasks. To a certain extent, it has promoted the evolution of cloud service platforms from "useable" to "easy to use." Author [4] proposed a cloud manufacturing knowledge service method based on uncertain rule reasoning, which realizes the knowledge service from quantitative to qualitative and then from quantitative to qualitative. The conversion to quantitative provides a new idea for the accurate distribution of cloud manufacturing knowledge services. Reference [7] builds a knowledge service realization model for the machining process of blade parts in complex curved parts in a cloud manufacturing environment, which shortens the time to a certain extent. In this research, we have inferred that LSTM-based models are quite higher than ML algorithms because the LSTM framework for organizations was used to improve the knowledge graph, and a heuristics candidate answer ordering approach was employed to verify the system's performance.

The generation time of the tool path in the processing of blade parts is improved, and the processing efficiency and quality of the blade are improved. The literature [8] proposes an optimization method for matching the knowledge service demander and the provider and uses a fuzzy axiomatic design. The theoretical measurement of matching satisfaction provides a specific theoretical basis for matching the supply and demand of knowledge services. To realize the

acquisition, storage, and sharing of different types of knowledge, literature [9] designs a cloud computing industry alliance based on the behavioral characteristics of knowledge interaction. To a certain extent, the knowledge resource sharing efficiency of the cloud computing industry alliance and the level of knowledge service have been improved. In this research, we have used the graph knowledge method. The limitation of this method is as follows: its evaluation measures concerning the model fit or factor levels cannot be done using graphic approaches since they do not provide certainty ranges for the variables (levels supplied by a correlation tool for all of this type of data are wrong). The above research has promoted the development of CMKS to a certain extent. The correctness and trustworthiness of the knowledge service offered cannot be ensured, resulting in low knowledge service performance. To that purpose, this study uses the user's mental behavioral traits as a judgment element, combining the benefits of the rough set theory in dealing with ambiguity and uncertainty. The CMKS is a knowledge-sharing and sharing service that can help cloud manufacturers succeed. As a consequence, all manufacturing operations in the manufacturing environment run smoothly, consistently, and precisely. Still, the research on the optimal decision-making method of knowledge service is relatively lacking. The accuracy and credibility of the provided knowledge service cannot be guaranteed, resulting in the low efficiency of knowledge service. To this end, this paper takes the user's psychological behavior characteristics as a decision-making factor and combines the advantages of the rough set theory in dealing with uncertainty and ambiguity. The optimal selection of services provides a reference idea and method.

2. Characteristics Analysis of Cloud Manufacturing Knowledge Service

CMKS is a kind of service guided by users' knowledge needs on the cloud manufacturing service platform, aiming at knowledge innovation, solving the manufacturing problems faced by users, and proactively providing users with personalized and specialized knowledge resources [10, 11]. The cloud manufacturing knowledge service (CMKS) is an information-sharing and sharing service that can help cloud manufacturers succeed. As a consequence, all manufacturing processes in the manufacturing environment run smoothly, consistently, and precisely. The integration of knowledge services enhances the operating effectiveness and problem-solving capacities of cloud manufacturing systems dramatically.

Compared with other services, knowledge services in the cloud manufacturing environment have the following salient features:

- (a) Knowledge transformation is complex: there are many types of knowledge resources in the cloud manufacturing environment, including explicit knowledge, such as standards, patents, and documents, and tacit knowledge, such as experience and capabilities. Most of these highly personalized and

challenging to standardize implicit knowledge resources come from different enterprises. Because of the inconsistency of knowledge resource management and processing methods, other users can benefit from the same knowledge resources. Therefore, converting cloud manufacturing knowledge service resources into their proper value is more challenging.

- (b) Dynamic changes in value: in the cloud manufacturing environment, the value of complex manufacturing resources will decrease with aging and wear and tear, while the value of knowledge resources changes in different directions. For example, with the growth of age and the continuous accumulation of knowledge, the value of employee experience will become higher and higher. With the advancement of technology and users' higher requirements for product quality, the value of a specific method or patent will become higher. Cloud platforms need to monitor these knowledge resources and continuously update their value status to provide appropriate services to users at the right time.
- (c) Highly integrated and shared: as an intellectual resource on the cloud manufacturing service platform, knowledge can provide oral, written, electronic, and other means and is minimally affected by geographical location. Therefore, compared with the complex manufacturing resources on the cloud manufacturing platform, the integration and sharing of resources are better.
- (d) Heterogeneity and isomerism coexist: Because of the differences in thinking habits and problem-solving methods, the manifestations of knowledge condensed by people are also complex and diverse. The utility of the same type of knowledge services is also different. Therefore, knowledge services in the cloud manufacturing environment are heterogeneous.

3. Mental Illness

The pace of life in the current society is getting faster and faster, and people are facing more pressure from work and study, which makes them susceptible to mental illness. The new corona epidemic has brought psychological stress, panic, and anxiety to people and increased mental illnesses' prevalence. Mental diseases have become a significant global public health problem [12]. Mental illness requires timely treatment. Otherwise, the long-term accumulation of negative emotions it brings will cause incalculable consequences [13]. For example, Cui Yongyuan, the famous host of CCTV, suffered from depression because of his work troubles, and suicides of college and middle school students are also familiar [14].

On the one hand, this phenomenon is because of people's lack of basic knowledge of mental illness, lack of a clear understanding of the dangers of mental illness, and lack of advanced means to popularize the understanding of the mental illness. However, on the other hand, psychological

counselors are in short supply, and counselors cannot get timely and effective help [15], lacking a scientific theoretical system to solve psychological problems intelligently. Therefore, an intelligent platform is needed to store the relevant psychological counseling knowledge. Then, when the user needs to obtain psychological understanding, the platforms can quickly get feedback from the immense ability to help and guide the mentally ill patients in time. The knowledge graph is a better method for intelligently storing information at present. This method was proposed by Google in 2012 and was quickly used for intelligent semantic search [16]. At present, artificial intelligence technology is gradually becoming mature and has penetrated all aspects of society, especially the development of natural language processing technology. The knowledge graph has more application prospects. The influence of social media on mental health is that it may have an effect on psychological health and behavioral activities that might have potential medical care concerns, and social networking holds an ever expanding route for both our everyday lives and globally. As a result, there is a pressing need to develop a better understanding of the long-term effects of social media on people's health. As an essential application direction of the knowledge graph, the question answering system based on knowledge graph can quickly find the correct answer from the knowledge graph through the input natural language question and present it to the user in the form of natural language, and this question answering system is efficient in response and feedback. To solve the above problems, this study, firstly, constructs a knowledge map of psychological counseling, promotes the knowledge of the mental illness, and provides quick psychological counseling services by establishing an immediate semantic psychological question answering system. The system uses the BiLSTM algorithm to detect suicidal tendencies, identify people who commit suicide, self-mutilation, and harm others in time, help users identify diseases, and help consultants understand relevant knowledge through knowledge quizzes. The BiLSTM method is used to predict the next piece of information based on the previous piece of information, making it ideal for having dealt with contextually related text data like sentences. As there might be delays of undetermined time across critical occurrences in a time series, LSTM methods are well-suited to categorizing, processing, and generating forecasts time series analysis. This research allows counselors to relieve their troubles, make up for the shortage of psychological counseling resources, and improve the work efficiency of psychological counseling.

4. Related Work

Knowledge graphs can be divided into general knowledge graphs and domain knowledge graphs. Typically available knowledge graph representatives include Freebase [8], DBpedia, Yago, Baidu, Google, etc., and they are mainly based on triple fact-based knowledge and have a certain tolerance for the quality of knowledge extraction. Typical domain knowledge graphs include e-commerce, finance, medical care, etc. In the field of e-commerce, take Alibaba

as an example. Its knowledge graph has reached tens of billions, widely supporting commodity search, commodity shopping guides, intelligent question and answer, etc. Knowledge graphs allow investors and financiers to understand investment behaviors more quickly and grasp market conditions in the financial field. Aiming at the lack of financial charts, the author constructed a small financial knowledge graph using the crawled structured data, such as financial stocks and corporate information. At present, knowledge graphs are used primarily for clinical treatment decision support, medical intelligent semantic search, and medical question-answering systems in medicine. Clinical treatment decision support is to automatically generate a treatment plan for each patient according to the patient's situation, combine it with extensive data analysis in the medical field, and provide it to doctors for reference. Medical intelligent semantic search is to combine related entities from the medical knowledge graph. It can be used to query information such as the relationships and attributes to optimize the search results of medical information. Medical question and answer is another form of medical information retrieval. Its returned answers are in the form of natural language.

In terms of question answering system based on the knowledge graph, Tan Gang et al. used the LSTM model for entities/assertions to enhance the knowledge graph, used a heuristic candidate answer ranking method, and finally verified that the system has good performance through experiments. The author proposed a multi-round question answering system based on the knowledge graph of road regulations, which can better identify user intentions; the author with knowledge graph as database support, designed a question-and-answer system, which is based on the e-commerce field and realizes functions, such as question answering and reasoning. The authors are the parts of preprocessing, question classification, question template matching, and answer generation in a question answering system are implemented. The question answering procedures established above have been well-implemented in their respective fields, and the characteristics of the domain are integrated into the question answering system process. However, no scholars have studied question-answering systems in psychological counseling, and few texts have been classified as applied to question answering systems.

5. Question Answering System Framework

Compared with traditional search engines, the question answering system is more targeted, accurate, and more accessible for users to accept. The knowledge graph of psychological counseling realizes the association and integration of various kinds of knowledge in psychological counseling. It expresses understanding semantically in a professional and structured way, which can conveniently manage and query knowledge. The BiLSTM technique is used to predict suicidal inclinations, identify persons who commit suicide, self-mutilation, or damage others in a timely manner, assist users in identifying

illnesses, and assist consultants in understanding necessary knowledge via knowledge assessments. The question-and-answer system constructed in this paper is biased toward acquiring professional knowledge of psychological counseling to help people with psychological problems find the correct answer and incorporate suicidal tendency detection to identify dangerous speeches promptly. The framework of the psychological counseling question answering system based on knowledge graph mainly includes four parts: data acquisition, graph construction, problem understanding, and user interface (Figure 1).

The data acquisition module of this system obtains relevant data of web pages through crawler technology. It combines the open data of Chat opera, organizes it into structured data through data processing, and uses Neo4j's python to drive py2neo to construct a knowledge graph. The question-understanding module uses the HanLP tool for word segmentation, part-of-speech tagging, etc. It then uses the CHI feature selection, uses the established BiLSTM model classifier to classify the question and judge whether it has suicidal tendencies, and finally uses the BERT (bidirectional encoder representation from transformers) model to convert queries into queries word vectors for semantic similarity calculation to match question templates and generate answers. The user interface module is the user's question input, and the system answers feedback and involves the mutual conversion of speech and text.

6. Question Answering System Implementation

6.1. Data Acquisition. The data sources for constructing knowledge graphs and classifying texts are web text data and Chat opera datasets, containing questions and label information. Web text data uses web crawler technology to crawl questions about suicide, self-injury, and injury tendencies and structured data about mental illnesses from Weibo Shudong, Baidu Know, 525 Psychology, Yixin, and Simple Psychology. Chat opera cooperated with some professionals to complete a corpus, which is the first open knowledge question-and-answer corpus in the field of psychological counseling, including 20,000 pieces of psychological counseling data. The dataset includes the annotation information of suicidal tendency, and the annotation information and question information are extracted from this dataset. Question information consists of three columns, cat-id, cat, and question, category (Table 1).

6.2. Knowledge Graph Construction. The knowledge graph is used as the database support of the question answering system. Hence, it is necessary to build the knowledge graph first. The primary purpose of constructing a knowledge map of psychological counseling, referring to the psychological counseling ontology database created by author [16] and consulting relevant scholars, is to analyze and summarize the psychological counseling process and the psychological knowledge involved. The psychological counseling

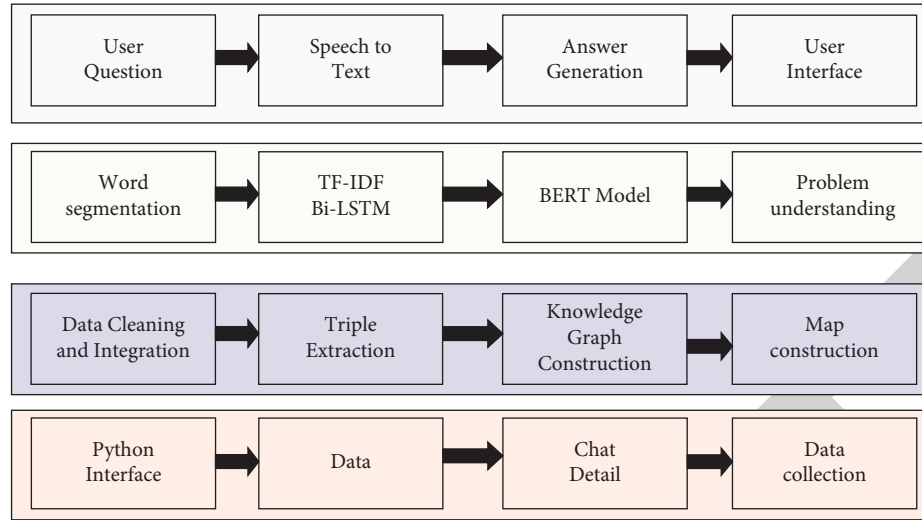


FIGURE 1: The representative framework for psychological question answering system.

TABLE 1: Examples of marked suicidal tendencies and normal questions.

Cat-id	Cat	Question
1	Tendency to self-harm	How to relieve the unreal feeling of depersonalization? In junior high school, my mood fluctuated a lot, and I cut myself with a knife because of some things and even suffered from severe depression
2	Normal	Ever since I reconciled with my boyfriend, he has always been hot and cold, and he does not want to admit that we are together. What does he mean?
3	Tendency to harm	Am I mentally ill? Ever since I was a child, I have had the idea of killing people and killing people around me.
0	Suicidal tendencies	I want to jump off a building, but I am afraid of heights

information is divided into six categories: patients, symptoms, diagnostic criteria, causes of disease, diagnostic results, and treatment methods. Among them, disease causes are divided into biological causes, psychological causes, social causes, and defense mechanisms. There is a causal relationship between the cause of the disease and the diagnosis result. The diagnosis result is divided into the severity of the disease and the name of the mental illness. Diagnosis results and treatment methods are decisive, and treatment methods are divided into psychological treatment, drug treatment, and food treatment. The treatment method has an executive relationship with the patient. The patient needs to adopt the treatment method. The patient mainly includes the patient's identity information, such as age, height, occupation, gender, etc., and the patient's past medical history, which will affect the treatment of the consultant. There is a relationship between patients and symptoms, and symptoms mainly include severity and symptoms. Symptoms and diagnostic criteria are subordinate, and the diagnostic criteria include disease course criteria, severity criteria, and exclusion criteria. It shows that these elements involve various aspects of mental illness, and a question and answer usually involves parts of multiple aspects of mental illness. There are often causal relationships and dependencies between these elements. The protégé tool is used to build an entity-relationship diagram in the field of psychological counseling (Figure 2).

6.3. Instance Layer Construction. The crawled data is organized into structured data and is divided into 12 fields, which correspond to the entity-relationship of psychological counseling designed above, and they are divided into diseases and disease-related entities and attributes. The entities are diseases, aliases, symptoms, complications, drugs, and foods, and the details are susceptible populations, examination methods, treatment methods, cure periods, costs, and causes of disease.

The instance layer construction is divided into entity extraction, relation extraction, attribute extraction, triple construction, and knowledge storage. Here, the python-driven py2neo tool of the Neo4j tool is used to create the knowledge graph. Entity extraction is to extract entities according to the concept of the schema layer and save the entity field information in the structured data after sorting into a dictionary. Relation extraction finds the relationship between entities and saves the relationship between diseases and other entities and attributes as a dictionary. Attribute extraction is to extract the attribute information of some entities, i.e., to save the attribute field information in the sorted, structured data into a dictionary. The construction of triples is to organize the data into the form of (entity, relationship, entity), create nodes without attribute fields and disease nodes with attributes, and then use disease nodes as start-nodes and attribute nodes as end-nodes, with query = "match(p:%s), (q:%s), where p.name = "%s" and

q.name = "%s" create(p)-[rel:%s{name: "%s"}]->(q)" % (start_node, end_node, p, q, rel_type, rel_name) command, and create a triple by creating a node relationship edge through the previously established relationship dictionary. Knowledge storage uses the platform to store the constructed knowledge graph. Here, the Neo4j platform stores the constructed small knowledge graph of psychological counseling Figure 3.

6.4. Word Segmentation Processing. Firstly, several mainstream word segmentation tools, such as Jieba, HanLP, and Chinese Academy of Sciences word segmentation NLPiR, are tested. The question-and-answer sentences of the psychological counseling QA corpus are selected as the data source. Before the test, psychology professionals were required to choose 100 data records and manually annotate them to achieve word segmentation and part-of-speech tagging as the actual value of the experiment. During the test, the 100 data records were processed with three-word segmentation tools, respectively, and the processing results were compared with the annotation results of professionals. The three-word segmentation tools obtained evaluation indicators, such as the accuracy rate, recall rate, F1 value, and word segmentation time (Table 2). Through the evaluation of the three-word segmentation tools, it can be concluded that NLPiR has the fastest word segmentation efficiency, however, HanLP has the highest accuracy. Combined with the evaluation results and psychological counseling needs, this study selected the HanLP tool for text segmentation. It imported the dictionary of disease information constructed above into HanLP's custom dictionary to make word segmentation more accurate.

6.5. Question Category. This paper combines the obtained psychological counseling question-and-answer data and consults professional psychological counselors to divide the frequently asked questions into five categories: disease identification, factual questions, method questions, list questions, and other questions. Disease identification questions mainly answer "what disease," real questions mainly answer "what," method questions mainly answer "how to do," list questions mainly answer "what are," and suicidal tendency questions are input by the user.

Since the principle of question classification is the same as that of suicidal tendency detection, the detailed process of suicidal tendency text classification is an example to illustrate. This article divides suicidal tendencies into four categories: suicidal tendencies, self-harm tendencies, harming others, and standard classes. Firstly, select 1500, 1200, 1000, and 9000 items from each dataset type as training data and then select another 150, 120, 100, and 900 items as test data. Then, using the HanLP word segmentation tool combined with stop words to perform word segmentation and afterword segmentation, the characteristic terms of suicidal tendency, self-harm tendency, injury tendency, and normal tendency were obtained, and the word cloud map of each category was drawn.

$$\text{chi} = \frac{n(ad - bc)^2}{(a + c)(b + d)(a + b)(c + d)} \quad (1)$$

Equation (1) is used to calculate the CHI value of all words after HanLP segmentation, which is used as the basis for feature selection of problem classification and suicidal tendency classification.

In the formula, N is the total number of labelled questions, A is the total number of documents used to record the word t in a specific category, B represents the documents that do not belong to a particular category but also contain the word t , and C represents a category that does not have the word t . In t documents, D represents documents that do not belong to a specific category and do not contain t . The CHI value denotes the degree of distinction between the two categories. The larger the CHI value, the more the word can represent a specific category.

Sort all the words according to the CHI value from large to small. Then, by selecting the 150 words with the most oversized CHI in each category, the selected 1000-dimensional feature list is obtained, and the feature vector weight of each question sentence is calculated, i.e., the word is in the question. If it appears in the sentence, it is assigned a value of 1. Otherwise, it is 0, and the final output is a feature vector that the classifier can recognize.

6.6. Problem Classification Model Selection. In text classification, five models of Naive Bayes, decision tree, SVM, XGBoost, and BiLSTM were used. Each model's training effects were compared, and the classification model with the best result was selected for the system. When testing each model, the precision rate, recall rate, and F1 value of each type of question will be obtained first, and then the precision rate, recall rate, and F1 value of all kinds of questions will be averaged. In the information extraction system, the BiLSTM model is tuned for suicidal propensity identification. The BiLSTM method is used to diagnose suicidal inclinations, identify those who commit suicide, self-mutilation, or damage others in a timely manner, assist users in identifying illnesses, and assist consultants in comprehending necessary facts through knowledge quizzes. Counselors can use this study to ease their stress, compensate for a lack of psychological treatment resources, and improve the efficiency of their job. In the information extraction system, the BiLSTM model is tuned for suicidal propensity identification. These values represent different models—the classification accuracy of (Table 3). From the test results in Table 3, it can be seen that the BiLSTM model is more accurate in the classification effect.

The BiLSTM model can predict the subsequent information using the previous information, which is suitable for dealing with contextually related text data, such as sentences. Taking the classification process of the BiLSTM model for suicidal tendency problems as an example, the HanLP word segmentation tool is selected for word segmentation, and the loss function and accuracy function image are obtained after the classification is completed (Figures 2–4).

TABLE 2: Evaluation results of three-word segmentation tools.

Word segmentation tool	Accuracy	Recall	F1 value	Word segmentation time (s)
Jieba	0.90	0.89	0.89	12.480
HanLP	0.91	0.90	0.90	04.478
Academy of Sciences word segmentation NLPPIR	0.81	0.75	0.78	30.0598

TABLE 3: Classification performance of model.

Model	Accuracy	F1-score	Precision
Naïve Bayes	79.65	80.56	69.56
Decision tree	74.52	81.26	65.25
SVM	75.62	76.85	70.62
XGBoost	80.54	82.54	72.56
BiLSTM	85.63	89.85	75.65

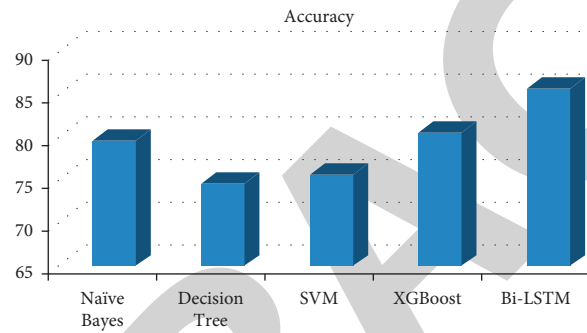


FIGURE 2: represents the accuracy of tendency classification model.

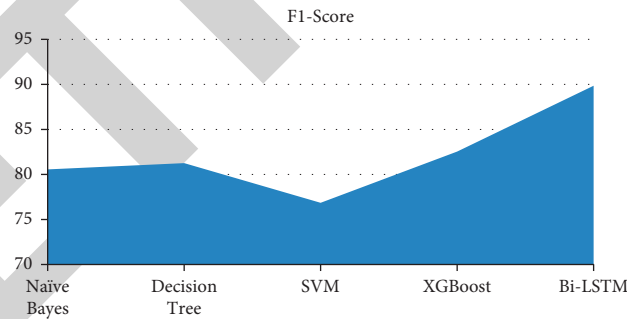


FIGURE 3: The F1-Score of tendency classification model.

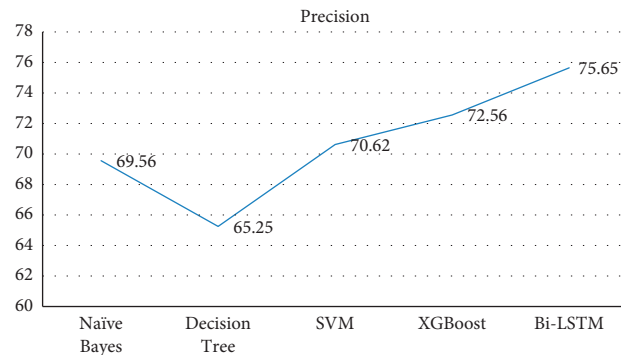


FIGURE 4: Representation of the precision of the tendency classification model.

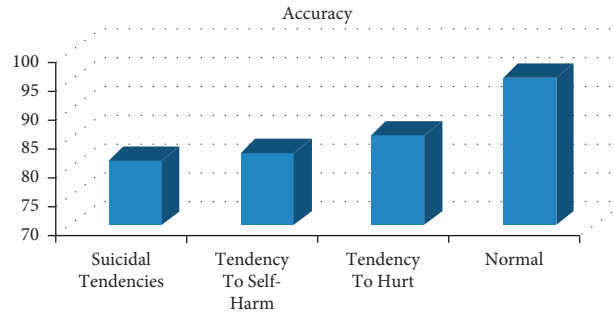


FIGURE 5: The accuracy of suicidal tendency classification by BiLSTM model.

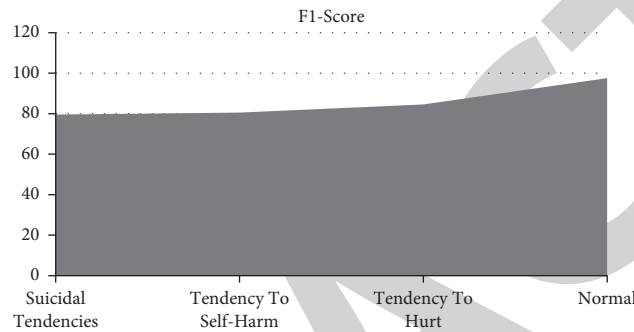


FIGURE 6: The F1-Score of suicidal tendency classification by BiLSTM model.

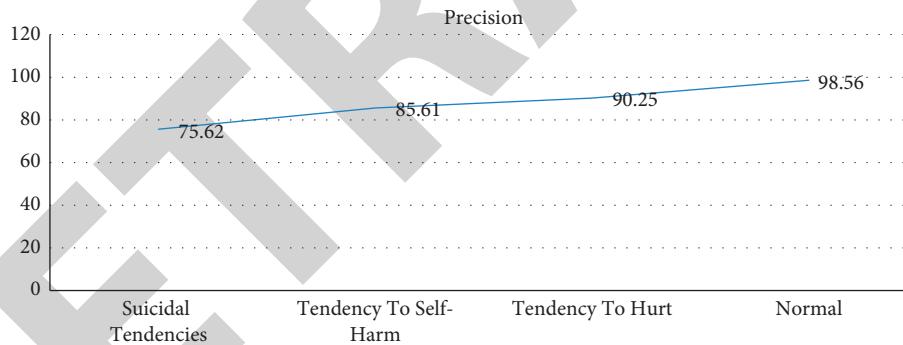


FIGURE 7: Representation of the F1-Score of suicidal tendency classification by BiLSTM model.

In Figures 5–7, the train represents the training set. The test means the test set. The horizontal axis of the function is the number of training iterations, and the vertical axis represents the loss value and the accuracy value, respectively. When the number of iterations becomes more extensive, the loss rate gradually decreases and starts to level off when the number of iterations is 14. The training set tends to be between 0.1 and 0.2, and the test set tends to be between 0.2 and 0.3. The accuracy gradually increases and starts to level off when the number of iterations is 14. The training set tends to be 0.975, and the test set tends to be between 0.925 and 0.950.

By testing the precision rate, recall rate, and F1 value of each type of question to evaluate the model (Table 4), it can be concluded that the BiLSTM model has a better classification effect than the previous four models.

TABLE 4: Suicidal Tendency classification by BiLSTM Model.

Question	Accuracy	F1-score	Precision
Suicidal tendencies	81.26	79.56	75.62
Tendency to self-harm	82.62	80.54	85.61
Tendency to Hurt	85.68	84.56	90.25
Normal	95.62	97.56	98.56

Comparing the questions answered correctly by the system with the questions answered incorrectly by the system, it can be seen that most of the questions answered correctly have similar characteristics to the question template, and some questions have the same entity and semantics as the question template. However, the expression methods are different, proving that the question answering

system has a high semantic understanding ability. For questions that are not answered correctly, it has two characteristics: the system can recognize the entities in these questions but cannot recognize the semantic information that does not appear in the template, such as “What is the general age of onset of bipolar disorder?,” “Can bipolar disorder be cured?,” etc. Another feature is that the entity attribute information involved is not comprehensive enough, for example, what side effects of drugs exist and other attribute information should be added to the knowledge graph.

7. Conclusion

Today’s society is facing the problem of knowledge explosion, especially the intermixing of knowledge in various fields, which affects people’s acquisition and understanding of knowledge. This study is geared on gaining the professional understanding of psychological counseling to assist people with psychological issues in determining the proper answer, as well as including suicidal inclination detection to quickly identify risky remarks. The advantage of this research is that this study uses the user’s psychological behavior traits as a decision-making element, combining the benefits of rough set theory in responding with uncertainty. A comparable notion and approach are provided by the best service choice. The vertical domain knowledge graph stores professional knowledge in a particular area, which significantly facilitates users to understand the knowledge in this field and use this knowledge to reduce losses or create a more excellent value. The psychological counseling knowledge graph and question answering system constructed in this paper has the following characteristics: (1) It provides a platform to store the knowledge of mental illness, which is more semantic than traditional databases. (2) The built question answering system realizes a complete set of word segmentation, question classification, question template matching, and answer generation, which can help users’ judgmental illnesses and help users query relevant knowledge about mental illnesses research significance and value. (3) The BiLSTM model is optimized for suicidal tendency detection in the question answering system. The experimental results show that the accuracy rate, recall rate, and F1 value of the model in this paper are significantly better than other traditional models in the detection of suicidal tendency.

Data Availability

The data shall be made available on request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] W.-C. Chiang, P.-H. Cheng, M.-J. Su, H.-S. Chen, S.-W. Wu, and J.-K. Lin, “Socio-health with personal mental health records: suicidal-tendency observation system on Facebook for Taiwanese adolescents and young adults,” in *Proceedings of the 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services*, pp. 46–51, Columbia, MI, USA, June 2011.
- [2] S. B. Hassan, S. B. Hassan, and U. Zakia, “Recognizing suicidal intent in depressed population using NLP: a pilot study,” in *Proceedings of the 2020 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 0121–0128, Faridabad, India, November 2020.
- [3] P. Gupta and B. Kaushik, “Suicidal tendency on social media: a case study,” in *Proceedings of the 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, pp. 273–276, February 2019.
- [4] C. Wu, P. Lu, F. Xu, J. Duan, X. Hua, and M. Shabaz, “The prediction models of anaphylactic disease,” *Informatics in Medicine Unlocked*, vol. 24, no. 100535, Article ID 100535, 2021.
- [5] G. Manimala, V. Kavitha, P. Pranav, and G. Vishnu Prasad, “Monitoring mental Heal TH using physiological signals,” in *Proceedings of the 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, pp. 1–9, IEEE, Chennai, India, December 2020.
- [6] J. Yuan, X. Lu, Y. Liu, D. Shi, T. Pan, and Y. Li, “Depressive tendency recognition using the gated recurrent unit from speech and text features,” in *Proceedings of the 2021 International Conference on Asian Language Processing (IALP)*, pp. 42–46, Yutai, China, October 2021.
- [7] A. Tiwari, V. Dhiman, M. A. M. Iesa, H. Alsarhan, A. Mehbodniya, and M. Shabaz, “Patient behavioral analysis with smart healthcare and IoT,” *Behavioural Neurology*, vol. 2021, Article ID 4028761, 9 pages, 2021.
- [8] L. Chang, A. Cassinelli, and C. Sandor, “Augmented reality narratives for post-traumatic stress disorder treatment,” in *Proceedings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 306–309, Recife, Brazil, November 2020.
- [9] A. Dev, N. Roy, M. K. Islam et al., “Exploration of EEG-based depression biomarkers identification techniques and their applications: A systematic review,” *IEEE Access*, vol. 10, pp. 16756–16781, 2022.
- [10] N. Q. Anayan and V. L. Penuela, “Coping mechanism of students below poverty line towards continuous education amidst COVID 19 pandemic,” in *Proceedings of the 2021 IEEE International Conference on Educational Technology (ICET)*, pp. 226–229, Beijing Shi, China, June 2021.
- [11] H. D. Calderon-Vilca, W. I. Wun-Rafael, and R. Miranda-Loarte, “Simulation of suicide tendency by using machine learning,” in *Proceedings of the 2017 36th International Conference of the Chilean Computer Science Society (SCCC)*, pp. 1–6, Arica, Chile, December 2017.
- [12] S. Liu, J. Shu, and Y. Liao, “Depression tendency detection for microblog users based on SVM,” in *Proceedings of the 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, pp. 802–806, Dalian, China, June 2021.
- [13] A. B. Rahmat and K. Iramina, “Classification of multiclass EEG signal related to mental task using Higuchi fractal dimension and 10-Statistic Parameters - support Vector Machine,” in *Proceedings of the TENCON 2015 - 2015 IEEE Region 10 Conference*, pp. 1–6, November 2015.
- [14] S. Saleque, G. A. Z. Spriha, R. Ishraq Kamal, R. Tabassum Khan, A. Chakrabarty, and M. Z. Parvez, “Detection of major depressive disorder using signal processing and machine learning approaches,” in *Proceedings of the 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1032–1037, Kristiansand, Norway, November 2020.

Retraction

Retracted: Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty

Computational Intelligence and Neuroscience

Received 23 November 2022; Accepted 23 November 2022; Published 28 December 2022

Copyright © 2022 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.

Computational Intelligence and Neuroscience has retracted the article titled “Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Chief Editor.

References

- [1] X. Zhang, Y. Shen, S. Yang, B. Wang, and J. Chen, “Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1054299, 6 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Research Article

Botulinum Toxin Type A Injection Improves the Intraperitoneal High Pressure in Rats Treated with Abdominal Wall Plasty

Xue Zhang, Yingmo Shen, Shuo Yang, Baoshan Wang, and Jie Chen 

Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

Correspondence should be addressed to Jie Chen; chenjiehernia@163.com

Received 30 December 2021; Revised 24 January 2022; Accepted 10 February 2022; Published 15 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Xue 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 aim of the study is mainly to study the subject of BoNT-A injection to improve IAH in rats undergoing abdominal angioplasty. The study problem in surgery, especially in ICU, burn, and trauma centers, intra-abdominal hypertension (IAH), and abdominal compartment syndrome (ACS) are common complications. At present, there are various treatments for IAH. The intramuscular injection of Botulinum toxin type A (BoNT-A) into the abdominal wall has received a lot of attention. Based on this, this study proposes a method for measuring abdominal pressure, applies BoNT-A to reduce abdominal pressure in the IAH state of abdominal wall angioplasty, and explores a way to increase the compliance of the abdominal wall under the premise of maintaining the sealing of the abdominal cavity, so as to realize the expansion of the abdominal cavity. A method is achieved to reduce intra-abdominal pressure and eliminate or alleviate ACS. The results of the experiment showed that when the rats in the control group were injected with the same amount of normal saline as the rats in the experimental group, the IAP was significantly higher than that in the experimental group ($P < 0.05$). This shows that BoNT-A increases the compliance of the abdominal wall while maintaining the closure of the abdominal cavity, thereby increasing the volume of the abdominal cavity and alleviating the state of IAH in rats.

1. Introduction

In recent years, due to the gradual increase in mortality of ACS and IAH, effective treatment of them has attracted increasing attention [1, 2]. So far, abdominoplasty is still an effective measure for the treatment of IAH [3]. The therapeutic botulinum toxin type A is artificially extracted from botulinum toxin type A. The efficacy and safety of BoNT-A injection therapy have been affirmed by the academic community and the industry [4, 5]. Therefore, exploring BoNT-A injection to improve IAH in rats undergoing abdominal angioplasty has important research significance.

Regarding the study of BoNT-A and IAH, many scholars have conducted multiangle explorations. For example, the overactivity of the whole body joints caused by Noonan syndrome leads to children's trapezius spasm and causes neck pain and craniocervical headaches; Tofts LJ studied the effect of BoNT-A on its treatment [6]; Bittar studied the efficacy of BoNT-A in the treatment of postherpetic

neuralgia [7]; Park studied the effect of Botox type A through the JNK signaling pathway that inhibits the pro-fibrosis effect of hypertrophic scar fibroblasts [8]. It can be seen that there are few studies on the effect of BoNT-A on IAH in rats. In this paper, BoNT-A injection is used to improve IAH in rats undergoing abdominal angioplasty.

The study contribution takes IAH as the research object to study the mechanism of BoNT-A on improving IAH in rats undergoing abdominal wall angioplasty. This article first presents a method for detecting abdominal pressure in rats, as well as a model of IAH in an animal. Different doses of BoNT-A were set in the control and experimental groups to evaluate the effect of BoNT-A on decreasing abdominal pressure in the IAH state of abdominal wall angioplasty. BoNT-A can increase abdominal wall compliance while preserving abdominal cavity closure and reduce IAH in rats, according to the results of the experiments.

This study found that injecting BoNT-A into the abdominal wall muscles of rats increased the abdominal

cavity's volume and decreased its pressure, providing BoNT-A for the clinical treatment of abdominal compartment syndrome A reliable basis.

2. Measurement Design of Intra-Abdominal Pressure in Rats

2.1. Making a Simple IAP Meter. Materials needed: desktop sphygmomanometer, disposable infusion set, three-way tube, disposable intravenous infusion needle ($0.9 \times 28\text{TWLB}$), 50 ml syringe, nitrogen gas storage bag.

Production method: cut out the infusion tube about 40 cm in length from the disposable infusion set, save one end of the filter, connect the disposable intravenous infusion needle ($0.9 \times 28\text{TWLB}$), connect the other end to the tee tube, and unplug the desktop sphygmomanometer, and the connecting tube between the cuffs connects the other ends of the tee tube with the infusion tube to the sphygmomanometer and the other tee tube with a 50 ml syringe and an oxygen bag for storing nitrogen.

2.2. IAH Animal Model Preparation. Female/male Sprague Dawley rats are randomly selected weighing about 220 g-250 g. The rats were anesthetized by intraperitoneal injection of 3% pentobarbital sodium (30 mg/kg) to maintain the rats' spontaneous breathing. After the rat is anesthetized successfully, place the rat on the rat board, prepare the skin on both sides of the groin, cut the skin, and routinely disinfect, and then, place a disposable intravenous infusion needle ($0.9 \times 28\text{TWLB}$) at a 45-degree angle with the abdominal wall in the left lower abdomen. Slowly puncture the needle toward the upper part of the abdomen and place it into the abdominal cavity. Connect the sphygmomanometer and the nitrogen gas storage bag through the three-way tube, and connect the gas storage bag to a 50 ml syringe. Inject the nitrogen-connected insufficiency needle into the abdominal cavity of the rat, slowly and continuously increase the amount of gas into the abdomen, and continuously monitor IAP dynamically [9].

2.3. Mortality of Rats with Different IAP Levels at Various Time Points. The mortality of rats with different IAP levels at various time points is shown in Table 1: Group 1 rats died within 6 h of IAH and died at 4 different time points of IAH (IAH 1 h, 2 h, 4 h, and 6 h). The change trend of the rate is not obvious. The mortality of rats in the second group showed a gradual upward trend at the above 4 different time points. All 10 rats in the third group died within 2 hours, and 9 of them died within 1 hour.

Figure 1 shows that IAP10 mmHg for more than 2 h can cause abnormal abdominal function and IAP20 mmHg for more than 1 h can cause abnormal abdominal function and cause death. The higher the IAP, the higher the mortality of rats.

3. Experimental Design

3.1. Experimental Animals. There are 50 Sprague Dawley rats, male and female randomly, weighing about

220 g-250 g. Experimental animals that have been kept in the experimental animal center for at least 7 days were selected. All experimental animals were given free drinking water and granular rat feed according to the standards of the Experimental Animal Center and maintained a light-dark cycle for 12 h-12 h every day. Fasting was carried out for 12 h-16 h before surgery, and there was no restriction on drinking water.

3.2. Experiment Grouping and Specific Operations

3.2.1. Group Experiment Setting. Each pair of rats was randomly divided into the experimental group and control group, with 25 rats in each group. After fasting for 8 hours, ketamine hydrochloride injection (100 mg/kg) was used for intraperitoneal anesthesia. After successful anesthesia, the rats were fixed in the supine position on the operating table. After preparing and disinfecting the abdomen, the skin is incised along the midline of the lower abdomen and enters the white line into the abdominal cavity. The incision is about 1.5 cm long. A pressure-measuring balloon and a water injection tube are inserted into the abdominal cavity. The balloon catheter and the water injection tube are fixed on the abdominal wall.

3.2.2. Abdominal Pressure Measurement. The direct pressure measurement method is used for abdominal pressure measurement. During the pressure measurement process, it needs to pay attention to keeping the position of the pressure measurement tube and the pressure sensor at the same level. And the water injection pipe is connected to 500 ml 0.9% sodium chloride warm solution through a disposable infusion set, a small opening is cut on the upper side of the sodium chloride injection infusion bag to ensure that the pressure of the liquid plane is equal to the atmospheric pressure, and the flow regulator is adjusted. The gas in the drip hopper of the infusion set is completely discharged, and the liquid dripping speed is maintained at 40 ml/min. When the pressure in the abdominal cavity of the two groups of rats stabilizes at about 15 mmHg, the flow regulator is closed and the injection of each rat at this time water volume, about 350 ml/kg, is recorded. After completing the measurement, the water is released in the abdominal cavity.

3.2.3. Inject BoNT-A and Measure Intra-Abdominal Pressure

(1) Experimental Group. 0.8 ml of BoNT-A solution was injected into the abdomen of rats in the experimental group, and then, the intra-abdominal water injection tube and balloon piezometer were taken out, and the abdominal incision was sutured. Two days later, when the rats in the experimental group were injected with about 400 ml/kg of normal saline again, the pressure value at this time was less than 15 mmHg, and the water injection was continued to 15 mmHg. At this time, the water injection volume was about 510 ml/kg. The height of the infusion stand is adjusted to make the infusion. The height of the liquid level in the bag

TABLE 1: The mortality of rats with different IAP levels at various time points (%).

Sequence	Group	N	Time			
			1 h	2 h	4 h	6
IAP10 mmHg	1	10	0	0	0	20%
IAP15 mmHg	2	10	20%	40%	50%	75%
IAP20 mmHg	3	10	90%	100%	—	—

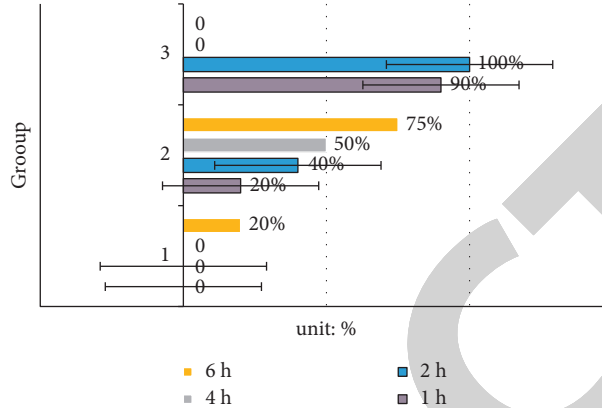


FIGURE 1: The mortality of rats with different IAP levels at various time points (%).

is about 20 cm from the height of the rat horizontal plane, and the intra-abdominal pressure is maintained at about 15 mmHg. After 3 hours, the normal saline in the abdominal cavity was released, the water injection tube and balloon piezometer tube in the abdomen were taken out, and the abdominal incision was sutured.

(2) *Control Group*. The control group was injected with the same amount of normal saline at the same point. First, the intra-abdominal pressure is recorded when about 400 ml/kg of normal saline is injected into the abdomen of the control group rat; after 2 days, the intra-abdominal water injection is continued until the total amount reaches 510 ml/kg, records the intra-abdominal pressure at this time, adjusts the height of the infusion stand to keep the intra-abdominal pressure stable, releases the normal saline in the abdominal cavity after 3 hours, takes out the intra-abdominal water injection tube and balloon piezometer, and sutures the abdominal incision.

The rats in the two groups underwent left femoral artery puncture, and blood draw 1 h before the start of the experiment, 3 h after IAH, and 3 h after IAH was removed. 15 rats were randomly selected for blood collection at each time point, and blood was collected about 1.5 ml each time. ELISA method was used to detect the content of superoxide dismutase (SOD) and malondialdehyde (MDA) in rats.

3.3. *Statistical Analysis*. All data in this experiment are statistically analyzed by GraphPad Prism software, and the experimental results are expressed as mean \pm standard deviation (mean \pm SD). *T*, *Z*, and heterogeneity test methods are used to verify whether the differences in experimental results are statistical significance.

3.3.1. *T, Z Inspection*. Statistical test of combined effect size $SE(T)$: to test whether the combined effect size of multiple independent original studies is statistically significant. The *z* test is a commonly used method, and the calculation method is shown as follows:

$$SE(T) = \frac{1}{\sqrt{\sum w_i}}, \quad (1)$$

$$z = \frac{\bar{T}}{SE(T)}. \quad (2)$$

Among them, W_i represents the weight value of the *i*-th study; \bar{T} represents the combined effect size. The *P* value under the probability of this statistic can be obtained from the *z* value. When $P < 0.05$, the result is considered to have a significant difference. When $P > 0.05$, it means \bar{T} is not statistically significant.

3.3.2. *Heterogeneity Test*. The basic idea of this method of testing is that all studies come from the same population, so the true effect is consistent in different studies, and the difference between the observed effect sizes is due to sampling errors. The commonly used test is to calculate the chi-square value; that is, the *Q* test is shown as follows:

$$Q = \sum w_i (T_i - \bar{T}). \quad (3)$$

Among them, T_i represents the effect size of the *i*-th original study. When the result of the *Q* test is $P > 0.1$, it means that multiple original studies have homogeneity; when the result of the *Q* test is $P < 0.1$, it means that there are multiple original studies, and there is no homogeneity between studies.

3.4. Relationship between Intraperitoneal Water Injection and IAP. The comparison results of IAP between the two groups of rats with the same three intraperitoneal injections are shown in Table 2: when the first injection of water, the abdominal pressure of the two groups of rats can reach 12–15 mmHg. During the second water injection, the rats in the experimental group had an average IAP of only 10.72 mmHg; during the third injection, the IAP reached 14.49 mmHg; while in the control group during the second injection, the IAP reached 14.96 mmHg, compared with the experimental group of rats. The difference was statistically significant ($P < 0.01$); when the third water injection, the average IAP reached 17.63 mmHg, and compared with the experimental group of rats, the difference was statistically significant ($P < 0.01$).

Figure 2 shows that the injection of BoNT-A into the abdominal muscles of rats can effectively reduce IAH. Two days after the rats in the experimental group were injected with BoNT-A intramuscularly, they were injected with the same physiological saline as the first water injection, and the IAP was significantly reduced, indicating that BoNT-A increased the compliance of the abdominal wall while maintaining the sealing of the abdominal cavity. And it can expand the volume of the abdominal cavity and relieve the IAH state of rats.

3.5. Comparison of Serum MDA Content. The comparison results of the serum MDA content of the two groups of rats are shown in Table 3.

It can be seen from Table 3 and Figure 3 that the serum MDA content of the two groups of rats was higher than before the establishment of the IAH model ($P < 0.05$) at the two time points of IAH 3 h and 3 h after IAH withdrawal. Moreover, the serum MDA content of rats in the control group increased more than that in the experimental group, and T test of experimental group and control group data $P < 0.05$.

3.6. Rat Serum SOD Content. The comparison results of the serum SOD content of the two groups of rats are shown in Table 4: the serum SOD content of the two groups of rats has a significant decrease in the two time points of IAH 3 h and 3 h after IAH withdrawal ($P < 0.05$). Moreover, the serum SOD content of rats in the control group was lower than that in the experimental group ($P < 0.05$).

According to Figure 4, when IAH is maintained for 3 hours and pressure is withdrawn for 3 hours, the serum SOD content of rats in the two experimental groups has a greater decrease, which indicates that the ischemia/reperfusion injury of rats in the experimental group is significantly less than that in the control group. It is said that BoNT-A reduces the ischemia/reperfusion injury while reducing the body's IAH.

4. BoNT-A Injection Improves IAH in Rats Undergoing Abdominal Angioplasty

This study found that the two groups of rats had different degrees of peroxidation damage after IAH3h and IAH3h

TABLE 2: Comparison of IAP when three intraperitoneal water injections are equal (mmHg).

Animal grouping	N	Time		
		The 1st time	The 2nd time	The 3rd time
Test group	15	12.19 ± 0.43	10.72 ± 0.67*	14.49 ± 0.22*
Control group	15	14.74 ± 0.27	14.96 ± 0.54	17.63 ± 0.64

(Note: * $P < 0.01$, # $P < 0.05$).

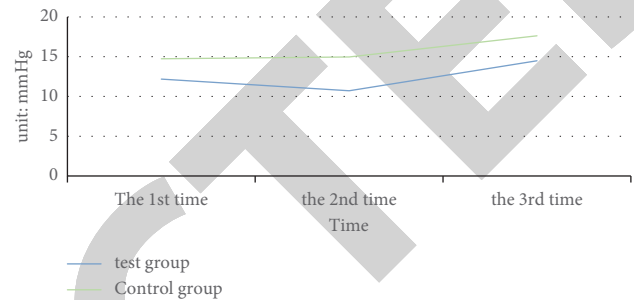


FIGURE 2: Comparison of IAP when three intraperitoneal water injections are equal (mmHg).

were withdrawn. Compared with the normal abdominal pressure, the two groups of rats also showed increased serum MDA levels ($P < 0.05$) and decreased SOD activity ($P < 0.05$) within the same period of time. The two indicators changed after 3 hours of withdrawal more significantly ($P < 0.05$). This shows that the body has different degrees of ischemia/reperfusion injury during IAH and after decompression.

IAH may cause pressure on the blood vessels of the abdominal organs and cause ischemic injury. Ischemia can affect vasodilation and defecation. Furthermore, hypoxia may affect the blood perfusion of specific organs in the abdomen, potentially leading to vascular reperfusion and oxidative stress [10, 11] infliction. When the IAH state is relieved, the blood flow of the kidneys increases. The reperfusion of long-term ischemic and hypoxic tissues and organs produces a large number of oxygen-free radicals in a short period of time. As a result, the ischemia/reperfusion injury becomes more serious.

Therefore, the degree of changes in serum MDA and SOD levels will decrease more significantly after IAH is relieved [12].

At the same time, two days after the rats in the experimental group were injected with BoNT-A intramuscularly, they were injected with the same normal saline as the first water injection, and the IAP was significantly reduced while the rats in the control group received the same amount of intra-abdominal injection as the rats in the experimental group. In normal saline, IAP was significantly higher than that in the experimental group. This indicates that BoNT-A can increase the compliance of the abdominal wall under the premise of keeping the abdominal cavity closed, thereby increasing the abdominal cavity volume and helping to alleviate the state of IAH in rats.

TABLE 3: Comparison of serum MDA concentration in two groups of rats (nmol/L).

Grouping	N	Time		
		1 h before surgery	IAH 3 h	IAH decompression 3 h
Test group	15	0.77 ± 0.24	1.32 ± 0.41*#	1.63 ± 0.47*#
Control group	15	0.81 ± 0.22	1.55 ± 0.67#	2.17 ± 0.24#

(Note: compared with the control group, *P < 0.05, and compared with preoperatively, #P < 0.05)

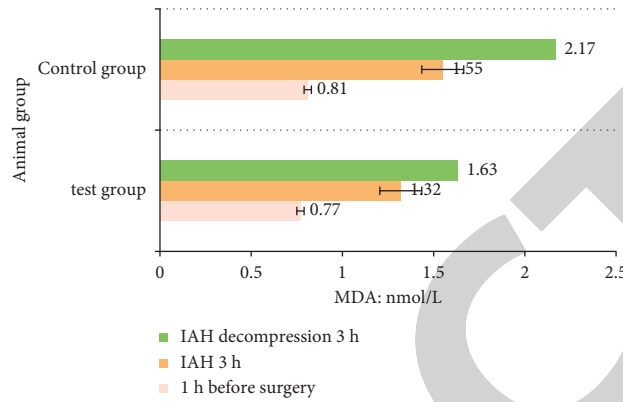


FIGURE 3: Comparison of serum MDA concentration in two groups of rats (nmol/L).

TABLE 4: Comparison of serum SOD content in the two groups of rats (ng/ml).

Grouping	N	Time		
		1 h before surgery	IAH3 h	IAH decompression 3 h
Test group	10	88.07 ± 18.06	83.04 ± 3.52*#	63.02 ± 4.05*#
Control group	10	88.07 ± 18.06	70.90 ± 3.20#	47.61 ± 3.21#

Note. Compared with the control group, *P < 0.05; compared with preoperatively, #P < 0.05.

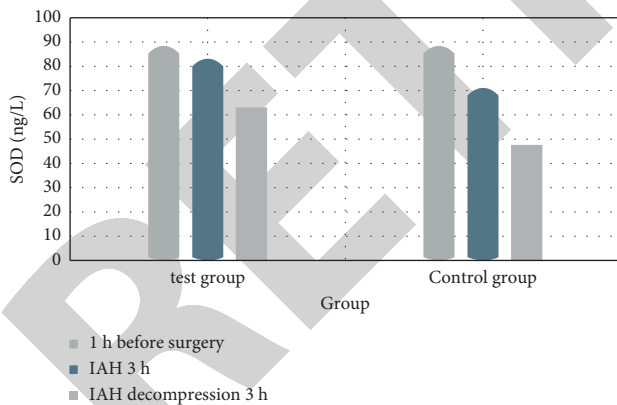


FIGURE 4: Comparison of serum SOD content in the two groups of rats (ng/ml).

5. Conclusion

This study found that BoNT-A injection into the abdominal wall muscles of rats increased the capacity of the abdominal cavity and reduced the pressure of the abdominal cavity, consequently eliminating or easing abdominal compartment syndrome. The foundation is solid. As a result, injecting BoNT-A into the abdomen wall intramuscularly to prevent and treat IAH has a favorable impact on lowering intra-abdominal pressure and can be utilized in clinical IAH therapy.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that there are no potential conflicts of interest.

Authors' Contributions

All authors have seen the manuscript and approved it to submit to your journal.

References

- [1] A. Skervin and M. Mobasheri, "Abdominal compartment syndrome," *Surgery*, vol. 37, no. 10, pp. 588–594, 2019.
- [2] A. di Natale, U. Moehrlen, H. R. Neeser et al., "Abdominal compartment syndrome and decompressive laparotomy in children: a 9 year single-center experience," *Pediatric Surgery International*, vol. 36, no. 4, pp. 513–521, 2020.
- [3] J. Sun, H. Sun, Z. Sun, Y. Xin, Z. Shuhua, and W. Jianxin, "Intra-abdominal hypertension and increased acute kidney injury risk: a systematic review and meta-analysis," *Journal of International Medical Research*, vol. 49, no. 5, Article ID 3000605211016627, 2021.

Retraction

Retracted: Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] N. An and J. Chuo, "Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3440437, 7 pages, 2022.

Research Article

Walking and Activeness: The First Step toward the Prevention of Strokes and Mental Illness

Ning An  and Jing Chuo 

College of Physical Education, Taiyuan University of Technology, Taiyuan 030024, China

Correspondence should be addressed to Jing Chuo; chuojing@tyut.edu.cn

Received 9 January 2022; Revised 3 February 2022; Accepted 7 February 2022; Published 14 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Ning An and Jing Chuo. 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.

Physical activity, especially routine walking, is an imperative factor for the prevention of strokes, mental illness, and cardiovascular diseases (CVDs). The NIH (National Institute of Health) has also acknowledged walking as the most important factor of the stroke rehabilitation program. Many research studies are presented by physicians and researchers in the literature that highlight the positive impacts of walking on human health (physical and mental). This paper has the objective of studying the impact of regular walking, especially on mental illness, CVDs, and strokes. The C-reactive protein (CRP), P-selectin protein, and homocysteine biomarkers are considered to decide the improvement in the health of an individual with respect to CVDs and strokes. The other parameters considered for the recommendations of physicians and healthcare experts for mental health are PSS (perceived stress score) and ESS (Epworth sleepiness score) that control mental illness. The values are measured for the participating subjects before participating in the walking activity and after the end of the walking schedule to see the impact on individuals. The overall mental and physical health of an individual contributes to the chances of occurrence of CVDs, mental illness, and strokes in individuals aged between 40 years and 55 years, as per the study presented in this paper. The results show that the PSS and ESS scores are improved after the walking activity. Eventually, it improved recovery from many kinds of mental illness and also reduced the chances of strokes. Similarly, the levels of the biomarkers that determine the chances of an individual having CVD or stroke also improved. Walking can impact our overall health in many ways, however, in this paper, the focus is given to ailments, such as strokes, CVDs, and mental illness. The results show that stress and improper sleepiness can impact mental health negatively. The research outcome is measured by adding walking in a routine life so that every individual can get rid of many physical and mental ailments. The results presented in the paper reveal that the 90-day walking program has created a good impact on the health of individuals by improving their physical and mental health.

1. Introduction

A stroke is a medical emergency that occurs when a blood artery that supplies oxygen and nutrients to the brain is blocked, generally because of a blood clot [1]. Stroke symptoms include difficulty in walking, talking, and understanding, along with paralysis or numbness of the face, arm, or leg. Clinical manifestations include the stiffness or paralysis of the face, hand, or limb, along with trouble while walking, talking, and comprehending. In the United Kingdom, it is the third-largest risk factor for death rate and the main factor of acquired impairment. Every year, over 130,000 people experience their first stroke [2]. In addition

to easily available pharmacological therapy for stroke prevention, stroke avoidance and rehabilitative treatments are necessary. Moreover, establishing such techniques is a major undertaking of 21st-century medicine. The role of physical exercise in lifestyle modification is of significant interest because of the sheer high morbidity, mortality, and huge burden of stroke-related comorbidities on patients and health care resources. The use of physical exercise is similar to a therapeutic method.

There is a lot of data to promote the use of endurance training in patients with risk factors, both quantitative and qualitative. In the rehabilitation of stroke victims, there is a huge amount of research that endorses the objective of

maximum treatment outcomes. Indoor activities, such as playing computer games, watching Television, accessing computers, and staying in one spot for long periods of time, contribute to a sedentary lifestyle. It has a number of negative consequences for our bodies. Muscular endurance deteriorates, and nutrient composition decreases, causing bones to deteriorate [3]. Carbohydrates and lipids are tough for the body to break down. The circulatory system is harmed, the cells are inflamed, and there is a hormonal imbalance. Overweight, cardiovascular disease, high cholesterol, heart problems, stroke, metabolic disorders, hypoglycemia, malignancy, anxiety, stress, cancers, and other disorders connected with these maladies are all induced by physical inactivity [4]. To lower the chance of developing such diseases, one should adopt a healthier lifestyle and incorporate physical activity into one's daily routine [5].

The correlation between walking and cerebrovascular disease is complicated, and it interests doctors, physiotherapists, and epidemiologists to figure out all the factors that impact positively and negatively on such kinds of diseases. Physical inactivity was identified as one of the five key risk factors in stroke-related research, accounting for approximately more than 80% of the worldwide stroke incidence rates, as per the statistical data [6, 7]. Walking and physical activities are sharply promoted as a facet of stroke treatment programs and mental health improvement programs because of the rising pile of empirical studies demonstrating its advantages in the treatment of these diseases [8, 9]. This paper presents a comprehensive review of the studies on the relationship between physical exercise and CVD-related illnesses, including strokes, with an emphasis on stroke prevention. Figure 1 demonstrates the lowered impacts of total serum and cholesterol level, lipids, overall fat percentage, and inflammatory responses, all of which contribute to issues like stroke.

1.1. Related Study

1.1.1. Impact of Regular Walk and Physical Activity on Reducing the Chances of Stroke, CVDs, and Mental Health. According to the authors, the cumulative effect of the nutritional interventions investigated potentially lowered the risk of cardiovascular diseases by 12% and stroke by 11% [6]. This forecast is based on the assumption of nutritional improvements. It is estimated that one-third of all deaths because of strokes and heart attacks are caused by inadequate regular physical exercise [10]. Physical activity promotes endothelial function, which improves blood vessel vasodilation and vasomotor function [11]. Several studies have been conducted to explore the correlation between workouts and cardiovascular diseases in the ability to forecast stroke [12]. Exercise decreased the risk of overall heart diseases, sudden cardiac death, and stroke in a drug concentration pattern, according to a study of female investigations [13]. There is a direct correlation between exercise programs and cardiovascular disease that is material up to a specific threshold of fitness, according to research. In

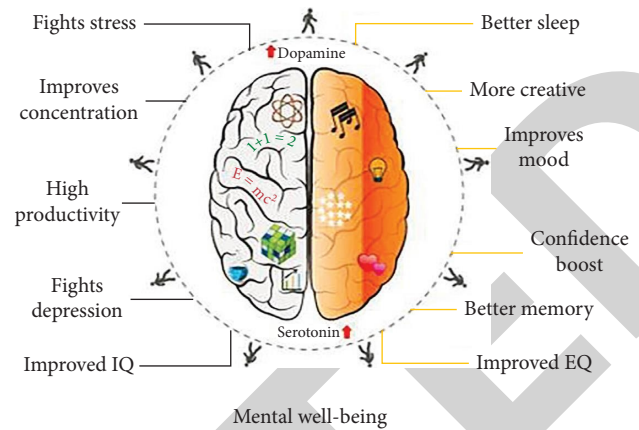


FIGURE 1: Benefits of walking.

the case of stroke, a similar relationship has been discovered. In [14], authors performed a study that included 15 peer groups and 5 case-control studies and found that people who are mildly active, highly active, and fit have a lower risk of stroke and mortality. It was observed that the people who are active, the significant decline in stroke occurrence was predicted to be around 20% to 27%. A concept of empirical evidence from the peer group and instance group studies that looked at the influence of workplace and relaxation workout on stroke risk found additional evidence in favor of exercise's preventive effect. The moderate levels of regular exercise at work were attributed to a 36% reduced risk of stroke when compared to being sedentary at the workplace [15].

1.1.2. Impact of Alcohol on the Occurrences of Strokes. Various investigations have demonstrated a U- or J-shaped association between fatality and liquor use. The light or average drinkers die at a rate lower than those who do not drink, and those who excessively drink die at a greater rate [16]. Heavy drinkers had a greater risk of death from any cause and from pulmonary illness, particularly sudden cardiac arrest and blood clots. Excess liquor use can also contribute to mental, social, and physical problems [17].

1.1.3. Psychosocial Factors That Can Reduce the Chances of Stroke and Mental Illness. Several social variables, including anxiety, stress, depression, social isolation, feelings of loneliness, and demanding work settings, have been demonstrated to be the major causes behind the development of stroke, cardiac diseases, and mental illnesses [18, 19]. If people can make their lives stress-free by adding a walking routine, then the chances of mental illnesses can be reduced remarkably.

1.1.4. Blood Pressure Control Can Reduce the Chances of Stroke and CVDs. Hypertension is still the most significant contributing variable for both types of stroke attacks. High pressure of the blood is thought to be the cause of nearly 700,000 early deaths globally. It is a key risk factor for CVD

and stroke. There is a strong and well-documented link between hypertension and stroke risk [20]. In healthy, normotensive persons, physical exercise is correlated with treating hypertension and a decreased likelihood of developing hypertension, thus changing a significant factor to stroke likelihood [21, 22].

1.1.5. The Effect of Gender on the Chances of Occurrence of Strokes and Mental Illness. Physical activity lowered the likelihood of an acute stroke by 23 percent in women and 26 percent in men, according to a study of data from 35 prospective studies and 11 randomized controlled investigations on the impact of physical fitness on the stroke-related cause of death and disability. A potential peer group study of Women's Health Study participants [23] looked at the impact of different lifestyle elements on the risk of stroke in females and found a weak correlation between the risk of stroke and physical fitness alone. The literature reveals that combining routine exercise with smoking cessation, a low body mass index (BMI), a nutritious diet, and consuming alcohol in lower quantities was associated with a lower risk of stroke [8, 24, 25]. As many of the trials excluded females, there is less data on the function of exercise in improving the outcomes in females to make judgments and provide guidelines.

All these studies explore the impact of many factors on the occurrence of strokes, mental health, and CVDs. There is not a single piece of research work that can show the impact of walking on all these ailments. Hence, it has motivated us to conduct a study on the impact of walking on mental health and strokes. This paper explores the impact of walking on strokes, mental health, and CVDs using a research-based study. The research endeavors of the paper are mentioned below.

- (1) This paper attempts to highlight the impact of walking on human health, especially focusing on mental illness, CVDs, and strokes
- (2) To carry out this research, three groups of subjects are created, irrespective of gender. The subjects of the walking activity are divided into three groups, namely Group-1 walkers, Group-2 walkers, and Group-3 walkers
- (3) The activity of walking was derived based on the individual interests and ability to walk based on their physical fitness/health
- (4) All the participants were recommended to join Telegram and Kakao Talk Messenger to share their data routinely
- (5) The records are collected and analyzed based on the data collected from smart devices and mobile apps, along with the medical records shared by the participant subjects

The paper has four sections. The very next section provides highlights on the methodology used for this research study. The third section describes the results, and the final section summarizes the work.

2. Suggested Research Method

To meet the objectives of the proposed research study, we have designed an innovative method as described in this section. The study is supported by the physicians as the parameters considered for the study are suggested by physicians and health practitioners. The standardized parameters are considered for the proposed study based on the recommendations of physicians and healthcare practitioners. People remain active during their young age, which may be considered below 30, and as they get older, their physical activities start reducing, and health issues start emerging because of a sedentary lifestyle. In sports, military, and police department, people remain active for approximately 40 years, and in our study, we have taken the age group between 40 years and 55 years for measuring the impact of walking on human health and diseases, such as cardiovascular diseases and strokes. The research study is performed on 10,000 subjects between 40 years and 55 years of age. IoT (Internet of Things)- and AI-based smart devices are utilized to collect the data from the subjects. The data includes the medical problems of CVDs prior to the use of walking as daily exercise, and the data is collected after 90 days of walking routine.

It is important to motivate individuals who have earlier worked with the armed forces and individuals who were into sports to start with the regular walking routine along with the individuals who have worked in other sectors. The following steps have been carried out to inculcate the regular walking activity schedule of individuals for 90 days. The subjects have faced one or other cardiovascular diseases earlier and may likely have a stroke in the new future. The attempt was made to save them from strokes. A total of 4300 women and 5700 males have participated as subjects. For sampling, we have particularly focused on 40 plus people, whose participation in physical activities was almost zero because of their office and other engagements in life.

Below is the method devised to indoctrinate the habit of walking in the participants for 90 days.

- (1) To start with the activity, three groups of subjects are created, irrespective of gender. Telegram and another mobile app (KakaoTalk Messenger) are used to share the instructions for their participation.
- (2) The walking activity is divided into three groups, namely Group-1 walkers, Group-2 walkers, and Group-3 walkers.
- (3) Group-1 walkers have the plan for covering 6 miles in a week through a casual walk.
- (4) Group-2 walkers opt for the plan to cover 2 miles a day by opting for a casual walk but not a brisk walk.
- (5) Group-3 walkers opt for the plan to cover 3 miles in an hour or less than an hour in a day.
- (6) The liberty for walking was given based on individual interest and ability to walk based on their physical fitness/health.
- (7) All subjects were advised to join Telegram and Kakao Talk Messenger to share their data every day.

The smartwatches used by people were GOQii Smart and Xiaomi. The data were also collected through mobile apps, such as YodoRun and Codoon.

- (8) The records are prepared based on the data collected from smart devices and mobile apps, along with the medical records shared by the participant subjects.
- (9) Other factors that are also studied along with the impact of walking on the participants are the hazardous impacts of smoking and alcohol consumption on cardiac and mental health.
- (10) The participants were segregated into three groups as per their choice of walking in kilometers.
- (11) The participating subjects are coached to be a part of the walking activity for 90 days. Walking was planned as per miles in a day or in a week.
- (12) All the participating subjects were allowed to share their data on Telegram and Kakao Talk Messenger.

The fields considered for research study are given below in Table 1.

By adopting the above method, an attempt is made to reduce the chances of cardiovascular diseases and strokes in the population aged between 40 years and 55 years. The next section elaborates the results and their respective interpretations.

3. Results and Interpretations

3.1. Impact of Walking on Alcohol Consumers with Respect to Strokes and CVDs. Patients of cardiovascular diseases who consume around 90 ml of alcohol every day or more than 90 ml on average are at a higher risk of having a stroke and other cardiovascular diseases. Alcohol consumption impacts liver function and raises blood pressure levels in the body, which indirectly induces the problems of cardiovascular diseases and increases the chances of stroke. In our sample population, there are 1800 alcoholic subjects who consume 90 ml or above alcohol on average every day. The reduction in the intake of alcohol to 60 ml from 90 ml and a regular walk of 30 minutes improved the health of subjects (sample population) remarkably and reduced the chances of having strokes by 20%. In this sample population, only males are considered who consume fixed amounts of alcohol every day. We have found it difficult to measure the changes. However, the impact of walking on health has been noticed well in the percentage scale for the occurrence of strokes in those who consume alcohol occasionally or in undetermined quantities. The facts are shown in Figure 2.

3.2. Impact of Walking on Improving Mental Health. In a sample population of 5000, the people are judged against the PSS and ESS scales. The values are obtained for 2000 participants who have opted for a casual walk for 2 miles a day, 1900 participants who have opted for 3 miles of brisk walk in an hour, and 1100 participants who have opted for a casual walk of 3 miles in a week. The PSS and ESS scores have improved remarkably for the participants who have taken an

initiative to walk to improve their mental and physical health. Stress is also a major cause of mental illness and strokes. Improvements in the scorecards of PSS and ESS can certainly reduce the risk of CVDs and strokes. Group-1 represents casual walkers who walk 6 miles in a week, Group-2 shows casual walkers who walk 2 miles in a day, and Group-3 belongs to active participants who walk 3 miles every day in an hour or less. The scorecard is self-explanatory. Figure 3 shows that Group-3 is the most beneficial group as its subjects are found to have improved PSS and ESS scores, while Group-2 is the next beneficiary. The third group (Group-1) also benefitted from the walk but not as much as Group-2 and Group-3, as shown in Figure 3. It depicts that walking not only reduces the chances of strokes but also improves mental health by reducing stress levels and by improving the sleeping score.

3.3. Improvement in Mood Swings after Regular Walk. In a population of 5000, 4100 people have been registered as unhappy with the problem of mood swings. A lack of confidence is also found to be a common problem among the participants. After participation in the regular and casual walk for 2 miles a day for 90 days, they found that their bad mood changed to a happy mood and the mood swings reduced remarkably. Besides, their confidence has increased as well. The results are given in percentage, as shown in Figure 4. Based on the outcomes of the study, efforts are made to investigate the reasons behind mood swings and how walking has improved mood swings to alleviate the mental health of people aged between 45 years and 55 years. It was found that mood swings were common in women close to menopause, and walking helped them a lot in improving their mood and mental health. Even men get mood swings as they lose their confidence, as they are getting older.

The results presented in this section reveal that the 90-day walking program has created a good impact on the health of individuals by improving the physical and mental health.

3.4. Interpretations. The results presented in this section reveal that walking is certainly a solution for improving mental health as it minimizes mood swings, reduces stress levels, improves the sleeping index, and improves overall health. The PSS and ESS scores have improved significantly in the participants who have taken an initiative to walk for improving their mental and physical health. The results shown in Figure 4 depict that after participating in a regular and casual walk for 2 miles a day for 90 days, the participants have found their bad mood to have changed to a happy mood and their mood swings to have reduced significantly. In addition to it, the confidence is also alleviated of the participants. Group-1 comprises casual walkers who walk 6 miles in a week, Group-2 comprises casual walkers who walk 2 miles in a day, Group-3 comprises active participants who walk 3 miles every day in an hour or less. The scorecard is self-explanatory. Figure 3 shows that Group-3 is the most benefited group as its participants have improved PSS and

TABLE 1: The parametric values taken for research study.

Attributes	Description
Subject_No	Each subject is given an ID
Group_id	Each subject is allocated a group ID
CVD_problem	This field mentions one CVD problem out of four types
History of strokes	This field contains bullion values in “yes” or “no”
Calorie chart	This field records calories burnt in a week after walking
C-reactive protein (CRP) initial	This field records CRP values before starting walking
Final_CRP	This field records CRP values after ending the walking routine after 90 days
PSS_initial	Perceived stress score is taken for study to check the stress levels of the participating subjects
PSS_final	The PSS of subjects was measured after 90 days of the walking routine
ESS_initial	Epworth sleepiness score is measured to check the stress levels before walking routine
ESS_final	ESS was measured again after the walking activity program of 90 days
Homocysteine_initial	Homocysteine levels depict the chances of having CVD or stroke, and this field records the initial levels of homocysteine before the participation of the subjects in the walking activity
Homocysteine_final	The final levels of homocysteine after the participation of subjects in the walking activity
P-selectin_initial	P-selectin protein measures are taken to see the chances of having strokes in the subjects
P-selectin_final	The levels of P-selectin are monitored to check the impact of walking activity
Walking_success	The field is used to record a bullion value in “Y” or “N,” where “Y” specifies that the program was successful and “N” specifies the inverse to it

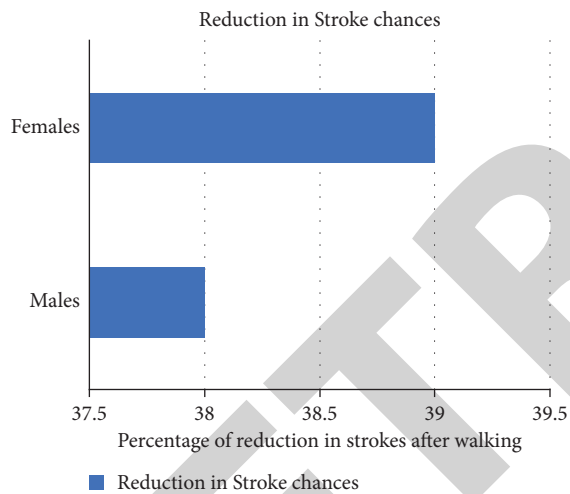


FIGURE 2: Walking impact on strokes as per the genders who smoke regularly.

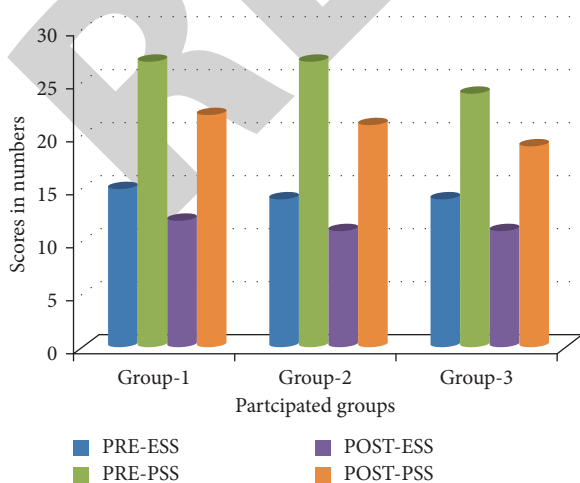


FIGURE 3: Walking impact on stress levels to avoid strokes and improve mental health.

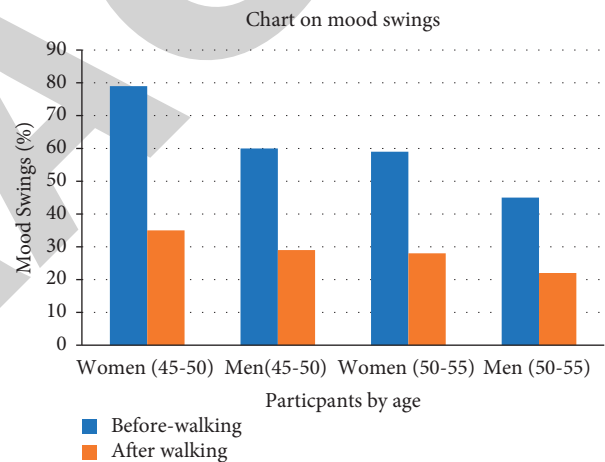


FIGURE 4: Walking impact on mood swings for better mental health.

ESS scores. Group-2 is the next beneficiary. The third group (Group-1) is also benefitted from the walk but not as much as Group-2 and Group-3, as shown in Figure 3. It concludes that walking not only minimizes the chances of strokes but also improves mental health by reducing stress levels and by improving the sleeping score.

4. Conclusions

This paper presents a study on the impact of walking on strokes, mental health, and cardiovascular diseases. Mental health can certainly be improved by adding walking in the routine life, which eventually reduces stress, improves sleeping index, and improves the mental health of a person. The stress levels can be handled well by walking regularly, and a reduction in stress can lead to an improvement in mental health. It can be summarized that it is very important to be active after the age of 40 to get rid of physical and mental ailments. By changing their

lifestyle and adding walking as a regular activity in life, individuals can control cardiovascular diseases and reduce their chances of having strokes. To conduct the research study, social media apps, such as Telegram and other fitness mobile apps, are used, as mentioned in the paper to record the data of the participants. The results show that individuals who even smoke and drink can improve their health by adding walking to their routine life. The scores of PSS and ESS are improved by 10% or more on average, which shows that people are able to reduce stress and improve their sleeping score by walking. Sound sleep is a must for avoiding strokes, improving mental health, and preventing cardiac diseases. The biomarker values are also recoded to prove the viability of the study, and the levels of CRP, P-selectin protein, and homocysteine are found to be improved, which shows that by adding walking exercise to one's routine, one can reduce the chances of having a stroke by 30% or more, depending on the routine followed by the individual. In the future, more parameters will be studied that can reduce the chances of having CVDs and strokes and improve the mental health of the individuals.

Data Availability

The data will be made available on request by the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors are thankful to the research groups that helped carry out this research work and "Research on the Development of Shanxi Nao Yang Movement from the Perspective of Interpretive Anthropology." The authors are also thankful to Shanxi Philosophy and Social Science Project-Folk Sports Era Value Research.

References

- [1] S. Marzolini, A. Balitsky, D. Jagroop et al., "Factors affecting attendance at an adapted cardiac rehabilitation exercise program for individuals with mobility deficits poststroke," *Journal of Stroke and Cerebrovascular Diseases*, vol. 25, no. 1, pp. 87–94, 2016.
- [2] R. L. Sacco, S. E. Kasner, J. P. Broderick et al., "An updated definition of stroke for the 21st century," *Stroke*, vol. 44, no. 7, pp. 2064–2089, 2013.
- [3] E. S. Ford, S. J. Smith, D. F. Stroup, K. K. Steinberg, P. W. Mueller, and S. B. Thacker, "Homocyst(e)ine and cardiovascular disease: a systematic review of the evidence with special emphasis on case-control studies and nested case-control studies," *International Journal of Epidemiology*, vol. 31, no. 1, pp. 59–70, 2002.
- [4] S. Gallanagh, T. J. Quinn, J. Alexander, and M. R. Walters, "Physical activity in the prevention and treatment of stroke," *ISRN Neurology*, vol. 2011, pp. 1–10, Article ID 953818, 2011.
- [5] M. J. Field, N. Gebruers, T. Shanmuga Sundaram, S. Nicholson, and G. Mead, "Physical activity after stroke: a systematic review and meta-analysis," *ISRN Stroke*, vol. 2013, pp. 1–13, Article ID 464176, 2013.
- [6] T. N. Turan, A. Nizam, M. J. Lynn et al., "Relationship between risk factor control and vascular events in the SAMMPRIS trial," *Neurology*, vol. 88, no. 4, pp. 379–385, 2017.
- [7] Y. Dong, "Research on the mechanism of physical exercise affecting residents' self-rated health: empirical analysis based on CFPS2018 data," in *Proceedings of the 2021 International Conference on Health Big Data and Smart Sports (HBDSS)*, pp. 9–20, Guilin, China, October 2021.
- [8] J. Z. Willey, Y. P. Moon, R. L. Sacco et al., "Physical inactivity is a strong risk factor for stroke in the oldest old: findings from a multi-ethnic population (the Northern Manhattan Study)," *International Journal of Stroke*, vol. 12, no. 2, pp. 197–200, 2017.
- [9] G. Wang, "Correlation between aerobic exercise and improvement of physical sub-health based on multiple linear regression," in *Proceedings of the 2020 12th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, pp. 1039–1043, Phuket, Thailand, February 2020.
- [10] E. J. Brunner, K. Rees, K. Ward, M. Burke, and M. Thorogood, "Dietary advice for reducing cardiovascular risk," *Cochrane Database of Systematic Reviews*, vol. 4, Article ID CD002128, 2005.
- [11] I.-M. Lee, C. H. Hennekens, K. Berger, J. E. Buring, and J. E. Manson, "Exercise and risk of stroke in male physicians," *Stroke*, vol. 30, no. 1, pp. 1–6, 1999.
- [12] S. A. Moore, N. Hrisos, D. Flynn, L. Errington, C. Price, and L. Avery, "How should long-term free-living physical activity be targeted after stroke? A systematic review and narrative synthesis," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 15, no. 1, p. 100, 2018.
- [13] Y. Oguma and T. Shinoda-Tagawa, "Physical activity decreases cardiovascular disease risk in women," *American Journal of Preventive Medicine*, vol. 26, no. 5, pp. 407–418, 2004.
- [14] M. Quan, P. Xun, R. Wang, K. He, and P. Chen, "Walking pace and the risk of stroke: a meta-analysis of prospective cohort studies," *Journal of Sport and Health Science*, vol. 9, no. 6, pp. 521–529, 2020.
- [15] F. Sofi, A. Capalbo, F. Cesari, R. Abbate, and G. F. Gensini, "Physical activity during leisure time and primary prevention of coronary heart disease: an updated meta-analysis of cohort studies," *European Journal of Cardiovascular Prevention & Rehabilitation*, vol. 15, no. 3, pp. 247–257, 2008.
- [16] A. C. Smith, D. H. Saunders, and G. Mead, "Cardiorespiratory fitness after stroke: a systematic review," *International Journal of Stroke*, vol. 7, no. 6, pp. 499–510, 2012.
- [17] R. H. Fagard, "Exercise intensity and blood pressure response to endurance training," *Hipertensión Y Riesgo Vascular*, vol. 28, no. 1, pp. 20–23, 2011.
- [18] S. C. Larsson, N. Drca, and A. Wolk, "Alcohol consumption and risk of atrial fibrillation," *Journal of the American College of Cardiology*, vol. 64, no. 3, pp. 281–289, 2014.
- [19] K. J. Mukamal, K. M. Conigrave, M. A. Mittleman et al., "Roles of drinking pattern and type of alcohol consumed in coronary heart disease in men," *New England Journal of Medicine*, vol. 348, no. 2, pp. 109–118, 2003.
- [20] K. Berger, U. A. Ajani, and C. S. Kase, J. M. Gaziano, J. E. Buring, R. J. Glynn, and C. H. Hennekens, "Light-to-Moderate alcohol consumption and the risk of stroke among

Retraction

Retracted: A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] G. Devi S, S. Vairavasundaram, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3432330, 15 pages, 2022.

Research Article

A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts

Gayathri Devi S,¹ Subramaniaswamy Vairavasundaram,¹ Yuvaraja Teekaraman ,² Ramya Kuppusamy ,³ and Arun Radhakrishnan ⁴

¹School of Computing, SASTRA Deemed University, Thanjavur 613401, India

²Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 3JD, UK

³Department of Electrical and Electronics Engineering, Sri Sairam College of Engineering, Bangalore City 562 106, India

⁴Faculty of Electrical & Computer Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

Correspondence should be addressed to Yuvaraja Teekaraman; yuvarajastr@ieee.org and Arun Radhakrishnan; arun.radhakrishnan@ju.edu.et

Received 9 October 2021; Accepted 12 February 2022; Published 11 March 2022

Academic Editor: Syed Ahmad Chan Bukhari

Copyright © 2022 Gayathri Devi S 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.

Tamil is an old Indian language with a large corpus of literature on palm leaves, and other constituents. Palm leaf manuscripts were a versatile medium for narrating medicines, literature, theatre, and other subjects. Because of the necessity for digitalization and transcription, recognizing the cursive characters found in palm leaf manuscripts remains an open problem. In this research, a unique Convolutional Neural Network (CNN) technique is utilized to train the characteristics of the palm leaf characters. By this training, CNN can classify the palm leaf characters significantly on training phase. Initially, a preprocessing technique to remove noise in the input image is done through morphological operations. Text Line Slicing segmentation scheme is used to segment the palm leaf characters. In feature processing, there are some major steps used in this study, which include text line spacing, spacing without obstacle, and spacing with an obstacle. Finally, the extracted cursive characters are given as input to the CNN technique for final classification. The experiments are carried out with collected cursive Tamil palm leaf manuscripts to validate the performance of the proposed CNN with existing deep learning techniques in terms of accuracy, precision, recall, etc. The results proved that the proposed network achieved 94% of accuracy, where existing ResNet achieved 88% of accuracy.

1. Introduction

For future generations, knowledge from many academics was conserved in both written and oral forms. India went through several ages of speech, drawing, and painting. Writing evolved from drawing through a sequence of natural steps [1]. Different sorts of writing resources, such as leaves, wood, stones, barks, metals, and so on, were found as writing increased. Palm leaves were frequently employed among all accessible materials due to their abundant availability and ability to endure harsh preparing. Palm leaf scripts are an important component of India's written history [2–4]. Palm leaf manuscripts in various Indian languages and written scripts may be found all

across the nation in temples, libraries, and museums. An average life of Palm leaves has 300–400 years, despite their conditioning capabilities and many processing, preservation, and conservation procedures. Numerous palm leaf manuscripts written three centuries ago are now in a state of deterioration, and we risk losing the rich material contained on them [5]. These valuable manuscripts provide in-depth knowledge of Astronomy, Literature, Ayurveda, Construction, and Fine Arts written in many vernaculars and scripts from diverse times. In ancient times, the only method to repair and convey this precious material to future age group was to rewrite it. Digitization is being utilized to generate digital pictures of texts to preserve their information as technology advances.

Tamil is among the world's oldest and most recognized languages, having a rich literary heritage. In the ancient period, poets employed Palm leaves, notably in Tamil Nadu, to conceal information [6]. Sangam literature, masterpieces, Vaishnava, Saiva, medicinal works, gastronomy, astrology, Vaastu, gems, music, dance and theatre, and Siddha are all included in the ancient literature. The value of ancient medical texts in Tamil, as well as the necessity to preserve them, has piqued the interest of numerous academics in the recent decade [7]. The conserved old medical writings in Tamil by saints such as Agathiyar were generated for the first phase of a digitalization process to preserve medical materials, and around 10,000 manuscripts were successfully scanned [8]. For digitizing historical documents, handwritten character identification apps used three key methods: statistical, structural or syntactic, and neural network-based techniques.

Optical Character Recognition (OCR) is a method for converting various sorts of documents, such as scanned papers. This method can be classified as either offline or online. Offline handwriting recognition often uses a scanner to acquire the writing optically, and the entire writing is provided as an image [9–11]. Because the characters are identified as they are written, online handwriting recognition is also known as real-time recognition. Online OCR employs pen-based input devices to record the sequence of coordinate points as the character is typed [12]. Handwritten character recognition becomes complex due to various variations in the shape of characters, different writing styles, and overlapping and interconnection of neighboring letters. It also depends on the individual because we do not write the same character in the same manner. It is challenging to create an OCR system with good recognition accuracy for Tamil script [13]. The primary goal of this method is to detect Tamil cursive letters in palm leaf manuscripts. This is accomplished by categorizing the characters into relevant categories based on characteristics collected from each one. The steps below will assist in achieving greater accuracy.

- (i) Image Preprocessing
- (ii) Feature Extraction
- (iii) Segmentation
- (iv) Classification using CNN.

The leaf characters are segmented by using text line slicing, where noises are removed, and morphological operations are carried out in the work. Finally, the palm leaf characters are classified by using CNN model. The experiments are conducted on the collected dataset by using various parameter metrics. The remaining paper is organized as follows: Section 2 consists of Literature Review. The proposed methodology is presented in Section 3. The validation of projected methodology along with existing techniques on the collected cursive Tamil palm leaves manuscripts is presented in Section 4. Lastly, the research work conclusion with its future work is given in Section 5.

2. Literature Review

Tamil, one of the world's oldest languages, is gaining popularity among scholars throughout the world due to its historical relevance and capacity to persist for centuries. Individual interests and ideologies of academics have shaped study on palm leaves manuscripts in Tamil poetry over the periods and in a variety of ways. The most prevalent topic of research in the actual world is cursive character recognition of ancient Tamil characteristics, and it focuses on the documentation of Tamil characters to guarantee that a large amount of data is gathered. However, due to improper care, palm leaf manuscripts have been destroyed, as numerous people held them by default. As a result, a rigorous examination of Tamil palm leaf texts, translation, and cataloguing of their publishing were necessary. Thousands of palm leaf texts exist in the hands of institutes and specific healers, but they must be digitized, and proper catalogues were developed for the future. Similarly, numerous researches on the cursive handwriting recognition of the Tamil language are included in the survey. There is, however, no universal technique for identifying cursive Tamil letters (with sufficient accuracy) in their whole. As a result, various approaches have been used in each stage of the recognition process.

As such, this work is one of the first attempts to build cursive training datasets by manual and automated segregation of Tamil palm leaf scripts. Future researchers could make use of the data set to develop expert systems that could then be used for a range of purposes, such as a character classification founded on the period they have changed and the documentation of extinguished characters, and the documentation of characters whose shape could have altered. The study aimed to separate old Tamil palm leaf writings to provide a huge quantity of information about the cursive Tamil character. There have been no previous investigations on the identification of cursive characters in Tamil palm leaf texts. The ThimuNet of CNN is used in this research project to classify cursive Tamil characters, and its performance is confirmed. In the future, the produced training data might be utilized as input and fed into other deep learning systems for further categorization and finding of the cursive letters contained in manuscripts written in Tamil.

Ali and Joseph [8] developed a CNN perfect for dispensation real-time input pictures including Malayalam characters and the job of segmenting words and typescripts from an image and attractiveness prediction using the CNN model. The feature extraction job in this model is done implicitly in CNN by the gradient descent technique. This technique is efficiently utilized for digitizing Malayalam script, which comprises 36 consonants and 13 vowels, is approved out in stages, and has obtained an accuracy of 97.26 percent for the training dataset.

Balakrishnan and Pavithira [9] suggested a method for optimizing CNN using Simulated Annealing, demonstrating the efficiency of character appreciation. They discussed

several deep learning approaches, the capabilities of CNN, and ways for training CNNs. According to the definition, “character recognition is the process of classifying and distinguishing characters in an input picture and converting them to ASCII or any similar machine-accessible form.” The suggested approach assessed the OCR accuracy of multiple language texts from numerous books, revealing that the accuracy of CNN by SA is higher than the unique CNN.

Hossian and Afrin [10] provided TEMPLATE MATCHING, a method for extracting characteristics from pictures and identifying precise characters to generate written documents. They claimed, “OCR operates similarly to the human capacity to read.” The picture glance at by the scanner will be the source for the OCR scheme in the same manner as an image viewed by human eyes is the source for the brain in the natural process.” They said that this methodology works by comparing the picture’s derived image characteristics and the template for each conceivable displacement. Techniques for word identification and finding in images and videos may be classified as connected component-based methods and texture-based methods. We informed numerous ideas regarding Template Matching, OCR, and the many stages required in Template Matching implementation. They tested the scheme by associating the outcomes of taught datasets to those of unprepared datasets, with accomplished datasets providing 100% accuracy.

According to Baskar [11], the palm leaf was a frequent writing material in primeval India, with the most basic surviving specimens. Furthermore, beginning in the sixth century, the Palmyra palm was employed for document production. Analysts and preservationists have recently been interested in the preservation of records and texts. There has been some awareness in this respect, and their preservation is important. There was a separate custom for palm leaf manuscripts, in which old texts that had become irreversible had to be encrypted on fresh leaves, and the old ones were disposed of in a river. Furthermore, religious laws prohibited pundits from deporting the manuscripts to secure detention facilities. Lack of finances, insufficient apparatus, resources, and qualified staff in document sources are the mechanisms that impede effective care of palm leaf works.

Sageer and Francis [14] suggested a digital library for rare texts as a method for preserving palm leaves. The rate of deterioration to endangered records is higher in a variety of ways, and such leaflets are typically not constrained to the limited bounds of traditional library schemes. They are more commonly found in private groups, official records, and institutional collections. The risk of harm is greater with private groups. Making a digital assortment for endangered materials significantly alleviates this situation. This research examines the different procedures of creating a digital archive for endangered materials and user attitudes toward usability; it focuses especially on user attitudes toward the digitization of palm leaf manuscripts.

According to Devika and Vijayakumar [12], traditional wisdom helps establish long-term interactions between humans and nature. The purpose of this study was to differentiate the contents of palm leaves that can be digitized, the details for digitization, and the various techniques of scanning a palm leaf document. Statistical techniques were

used to analyze the collected data. The study provides useful information on the current collection of palm leaf organizations in Tamil Nadu, such as public libraries and private organizations, and nongovernmental administrations. It is critical to have a sufficient quantity of knowledge on the various approaches offered in Tamil Nadu’s palm leaf library. This study indicated that manuscript conservation by digitalization is one of the most successful and helpful approaches; however, it is time-consuming and expensive.

Various knowledge schemes were documented in ancient India using palm leaf scrolls, according to Narenthiran and Ravichandran [13]. The wisdom was primarily conveyed through the traditional Sanskrit language in various scripts. This study is based on five document libraries in Tamil Nadu, which are a mix of libraries of varying statuses with superior manuscripts. This research concentrated mostly on the cataloguing of manuscripts and the operations associated with digitalized these archives. The study was based on a national goal, and the papers were recognized.

Sabeenian et al. [15] evaluated localized binarization techniques on Tamil palm leaf scripts, which were recorded to hold a huge quantity of info that is highly useful in everyday life. To protect the data on the palm, digital pictures of each leaf were evaluated and preserved. When necessary, these pictures can be accessed. However, the storage volume required for each leaf is considerable, which raises the cost. These pictures must be binarized in order to acquire only the segmentation and storage. Several computer scientists have been stymied by poor reported picture binarization. This research focuses on palm leaf manuscript pictures that have been binarized using the Sauvola, Niblack, Bradley, and Nick binarization methods.

According to Kiruba et al. [16], one of the most essential phases in manuscript recognition is character segmentation in Tamil palm script. Tamil is India’s most frequently used script, and palm scripts contain consonants, modifiers, and vowels. In addition, the accuracy of the recognition approach is influenced by separate letters; therefore, suitable segments are necessary. This paper shows the image segmentation from palm leaf manuscripts of Tamil handwritten letters. The technique consists of three steps: background removal using the Otsu algorithm to separate texts, character segmentation, and line segmentation. This paper proposes a simple histogram-based method for segmenting Tamil palm script characters and discusses many obstacles to Tamil script segmentation.

As Ghosh et al. [17] note out that palm leaf works show a significant role in India’s prized national legacy, particularly in relations of their extraordinary accumulation of famous knowledge. Prior to the invention of paper, palm leaves were most commonly utilized resources for encrypting messages and literature. These books have survived the centuries and are still highly valuable today. The elements that deteriorate palm leaf documents have been investigated, with their conservation strategy, using both ancient and modern approaches. Because palm leaf works are important for knowledge conservation, the research suggested that they be conserved as a first step. The research also suggested that the most essential task of today’s archivists and libraries is to

find efficient methods for the conservation of palm leaf works.

Rendering to Challa and Mehta [18], there are numerous businesses in India dedicated to protecting the earliest palm leaf scripts, owing to the importance of preserving the valuable texts of knowledge. Both natural and artificial processes influence palm leaves throughout time. With the advancement of technology worldwide, this research attempted to digitize palm leaf scripts required in a university library. One specific goal was to create an efficient method of palm leaf texts for efficient information retrieval. Segmentation, picture augmentation, processing, acquisition, and compression are only a few of the techniques available. All of these technologies use distinct algorithms that are effectively applied and produce the intended results. Given that ideal identification rates are not feasible with warped media and noise, selecting a proper image processing approach is important. This research focuses on image processing systems and algorithms utilized on palm leaf texts projected by many scholars for effective data recovery.

Vinoth et al. [19] investigated a method that aids in the conservation of palm leaves. The Tamil language has been identified as one of the world's classical languages. Tamil Nadu was discovered to be a location where palm leaf writings have been preserved. However, there are difficulties encountered with the protection of palm leaf writings. As a result, this research absorbed on the significance of conserving palm leaf work by converting it to digital text format. Digital gadgets are currently used for this resolution of conservation. This digital technology aims to accurately transform the characters and preserve the documents. Overall, the first step is to scan the palm leaf text and convert it to an image layout before saving it in a record. Following that, the pictures should be converted into a digital text format.

The major goal of Sonam and Poornima's [20] study was to create a system that allowed Tamil character identification from palm leaves, as well as caption through the captured pictures and preserving them for future use. Confident training devices have been carried out using a variety of approaches; however, the procedure of distinguishing the postures of Tamil characters is difficult. It is also worth noting that the Tamil language is considered complex compared to other languages due to the existence of curves, slopes, pits, and twists, resulting in variations in different writing styles. Many researches necessitate the adaptation of ancient Tamil characters to current characters to certify that the goal of developing computerized schemes for improving human knowing comprehension is increased. This suggested research is thought to be useful for segmenting Tamil characters and keeping them in an orderly folder, as well as for subsequent image processing. To enumerate the statistical characteristics of the segmented characters, GLCM matrix feature extraction is used. At this phase, utilizing GLCM characteristics, the segmented Tamil characters are distinguished from those in palm leaf manuscripts.

Convolutional neural networks can be used to recognize Tamil palm-leaf characters, as demonstrated by Sabeenian et al. [21]. Five layers of CNN were used in this study: convolution, pooling, activation, fully connected, and softmax classifiers. Scannable photographs of palm-leaf manuscripts were used to compile the database of character sets. There are 15 separate classes in the database, and each class has roughly 1000 samples. Generally speaking, CNN Classifier's recognition is determined to be between 96.1 and 100 percent. For each CNN layer, we extract a big quantum, resulting in an improved prediction rate.

3. Proposed Methodology

This section explains the workflow of the projected scheme, where the four major steps are included. The first step is the collection of cursive Tamil palm leaves scripts, followed by background removal, feature extraction, and classification as the further steps. Figure 1 shows the workflow of the proposed methodology. The following section describes the significant things that are applied in the proposed model.

3.1. Dataset. The cursive Tamil palm leaf manuscripts are collected from the online images and stored in the database to document the target characters. In total, 100 images of cursive palm leaf texts are collected. Figure 2 displays the sample collected cursive palm leaf texts. This data set is enhanced and experimented by different methods, which are represented in the following section.

3.2. Preprocessing. Digitized palm leaf manuscripts image has dark leaf color as background and text in the foreground. The presence of pickup noise in palm leaf manuscript images that happens while scanning or taking photo using the digital camera can either be reduced by sharpening or using any morphological methods. The sample image of the digital noise occurred image is shown in Figure 3. The noise removal is a necessary step to obtain useful information from digital text images. In preprocessing, the background has converted as black and foreground text as white in colors from the range of 0 to 255 into 0 and 1 only so that the characters are very clear and easy for processing. The background removal and morphological operations are the methods to promote the images to be suitable for text line segmentation. According to the shapes of the leaf, the images are processed by the morphological operations, where structuring element is applied to an input image and obtained the output image with same sizes. In MATLAB, a structuring element called *strel* is included for this operation, and finally, the output is obtained. By using this operation, noise and texture that are caused by simple thresholding are removed efficiently.

3.2.1. Background Removal. In most of these systems, binarization is the scanned gray level image (labeling each pixel) print or background in document image analysis. Binarization is a process of assigning 0s and 1s using fixed

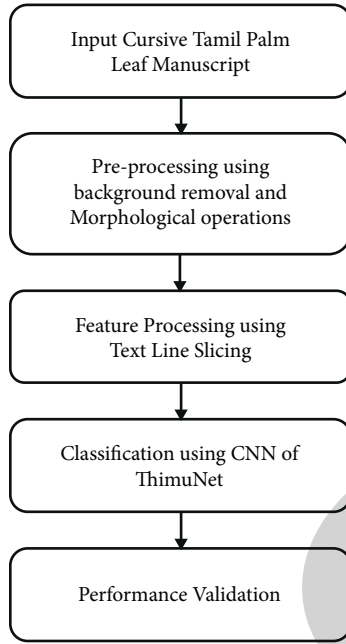


FIGURE 1: Working procedure of proposed methodology.

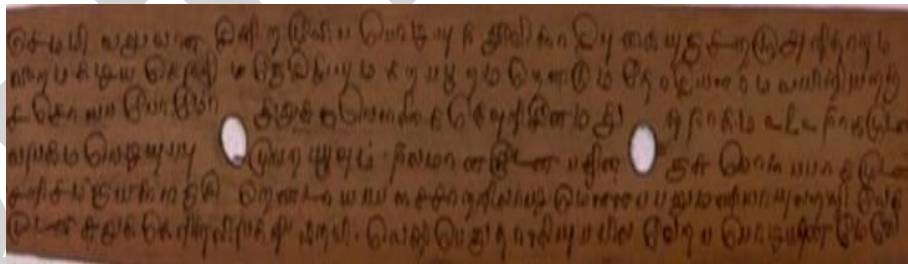
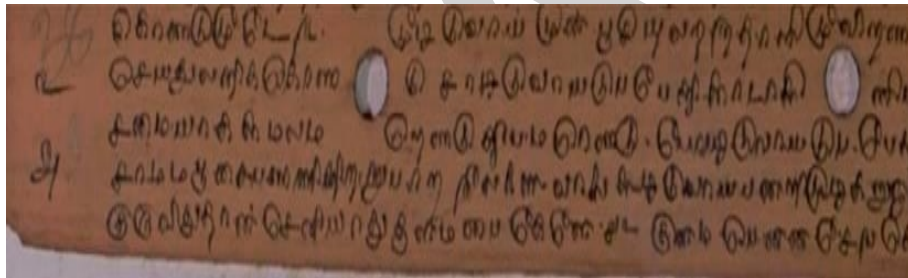


FIGURE 2: Sample cursive Tamil palm-leaf manuscript.



FIGURE 3: Sample noise occurred image before preprocessing.

threshold value. The fundamental idea of the fixed binarization method [21] is expressed in terms of relation. The background of palm leaf manuscripts can be taken as black by 0 and the foreground text as white by 1. Here, T shows global threshold value 50. After background removal, the preprocessed sample output image is shown in Figure 4(b) for the considered input image in Figure 4(a).

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) \geq T, \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where “ x and y ” is an input image sampled from the input data distribution, and “ g ” is the ground truth image corresponding to the input image.

3.2.2. Morphological Operations. In this operations, dilation and erosion are the fundamental operations. Addition of pixels with the boundaries of text objects in an image is known as dilation. Dilation is used to add pixels to the edges of regions or to fill in gaps in the image [22]. Erosion is the inverse of dilation. While dilatation raises the size of borders and fills holes, erosion diminishes the size of boundaries and enlarges holes. The reversal of this operation, i.e., extricating the pixels from the text object boundaries, is termed erosion [20]. In order to process the text-image in palm leaf manuscripts, the pixels may be added or removed depending on the size and shape of the text. In grayscale morphology, the images are mapped into the Euclidean space or grid $ES \cup \{r, -r\}$, where the grayscale erosion of the palm leaf image i by text boundaries b is given as in the following relation:

$$(i \ominus b) = \int_{y \in B}^{\infty} [i(x + y) - b(y)]. \quad (2)$$

3.3. Feature Processing. The text line segmentation in cursive Tamil palm leaf scripts is a Herculean task, and it influences till the end of the character recognition process. An absence of text line segmentation process is not possible for the successful character segmentation and character recognition in Tamil palm leaves. The TLS is applied on the preprocessed binary palm leaf text images to segment the text lines. The new way of approach in-text line segmentation of Tamil palm leaf images is to determine whether the obstacle is present between the text lines. Whenever the strokes of the character exceed from the text zone and extend in the space between the lines, then it is considered to be an obstacle in this case. The same is depicted in Figure 5.

3.3.1. Space without Obstacle. The TLS, the text lines of Tamil palm leaf texts where the text line has enough space to the subsequent text lines or an elongation of character does not reach the below text line as in (Figure 6, are considered as space without obstacle or standard category. TLS can segment these text lines without any complication.

3.3.2. Space with Obstacle. The presence of an obstacle in the space between the text lines can be categorized into two by the length of an obstacle that helps decide whether it is touching or overlapping text line. In Tamil character, an obstacle is an important part of deciding the character. The length of an obstacle that extends and reaches the subsequent text line is known as touching text lines (Figure 7). The first line character “ u ”/you” is touching with the next line character “ th /tha”. In Tamil, the text line segmentation is complicated because if we ignore an obstacle of the character “ u ”, it becomes “ u ”, and if we cut an obstacle in a fixed length, the second line character “ th /tha” becomes “ th /thi”. The overlapping text lines also have the same wrong prediction problem as touching text lines when we precede by existing text line segmentation algorithms. The proposed TLS solves the problem of touching and overlapping text lines by fixing the cutting edge at the end of an obstacle and also prevents wrong predictions of the character.

The purpose of text line segmentation is to precede the character segmentation. The touching and overlapping text lines complicate the text line segmentation and make the further process unproductive. An overlapping text line builds complication in-text line segmentation. An obstacle pervades the text zone of subsequent lines and mixed up with the character strokes that may precede wrong character prediction or different from expected character.

3.3.3. Text Line Slicing (TLS). The proposed TLS line segmentation algorithm identifies an extension of character strokes using four variables such as Vertical space (Vs), Horizontal space (Hs), Vertical Track (VT), and Horizontal Track (HT). The variable Vs used to count zeros vertically to know the stroke of a character exists in text line of binarized Tamil palm-leaf manuscript image. The total columns count of zeros is assigned to the variable VT and compared with the threshold. The value 1 denotes that the space has no obstacle, and 0 represents an obstacle. The variable Hs is used to count the zeros horizontally and the total value assigned to HT, and then, it is compared with a threshold value. Three values are used to decide whether an obstacle is present or not; zero (0) defines the space between the character in text line; one (1) defines that the character has an obstacle; and two (2) defines that the space is not found, which means that the character exists concurrently.

The obstacle creates touching text lines that can be defined by Connected Component (CC). The connectivity of the character is calculated by the weight of the character using CC. The touching and overlapping text line characters are considered single characters when they are connected to each other (Figure 8). An algorithm implementation (Figure 8(a)) proves obstacle identification in the space between the text lines, defines the category of connected characters by connected component, and calculates the weight for the character. When the minimum weight of the character identifies the TLS, the algorithm implements cutting edge (Figure 8(b)) to segment the connected text lines. Cutting edge is a breaking point of touching characters in text lines. The CC provides continuation of the character

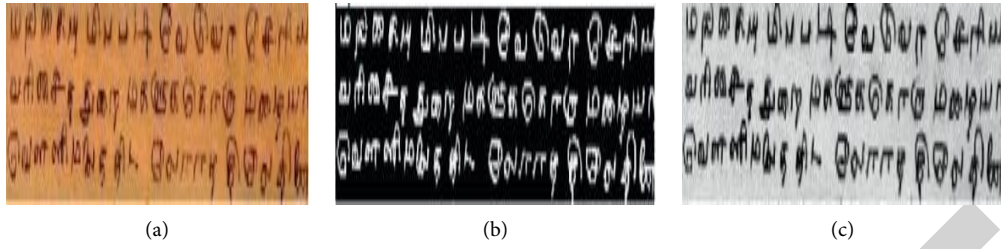


FIGURE 4: Preprocessed sample output image. (a) Input image. (b) Background removal. (c) Morphology.



FIGURE 5: Sample obstacle.

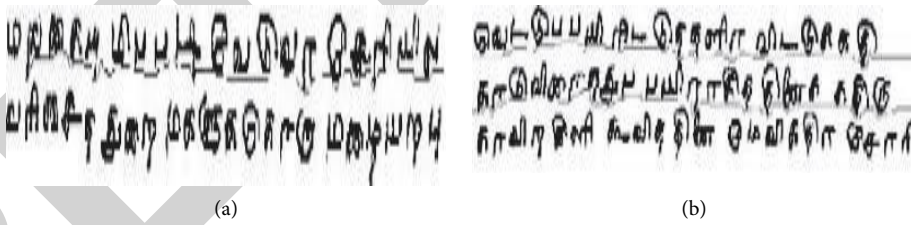


FIGURE 6: Some of the sample text lines. (a) No obstacle, (b) obstacle not reached to the below text line.

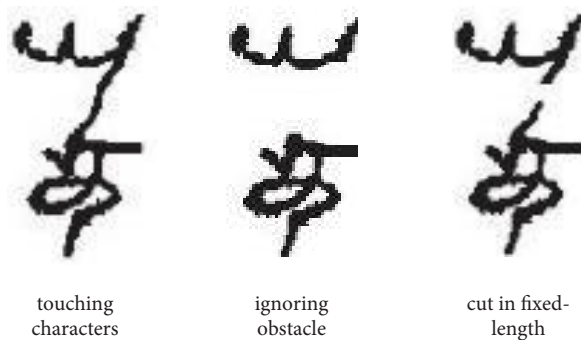


FIGURE 7: Sample obstacle defined in the text line.

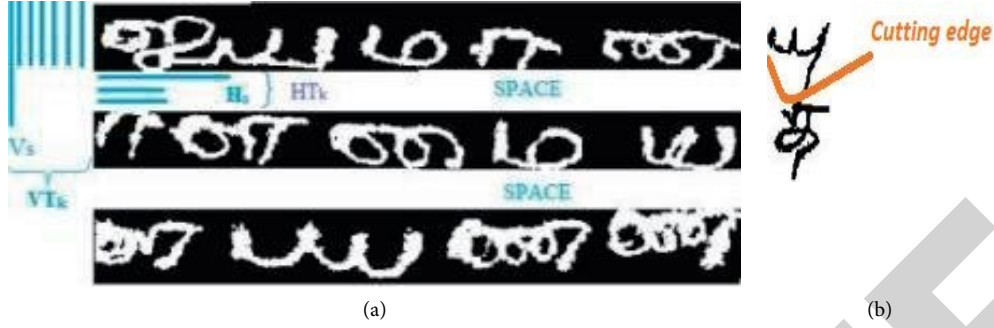


FIGURE 8: Sample text line slicing. (a) Implementation on Tamil palm-leaf manuscript, (b) cutting edge.

strokes and also vertical stroke values. The minimum value of the character stroke is known as the end of an obstacle that must be fixed as a cutting edge for the text lines. The sample images are shown in Figure 9.

3.4. Convolutional Neural Networks. Segmented and feature scaled images are given to the classifier model, which is CNN. Despite their popularity, ANNs were incapable of handling big datasets in recognition/classification responsibilities. To overwhelm these problems, profound learning is a new learning machine paradigm. It is a stacked, multilayered neural network. Earlier versions of the neural network, including the earliest perceptron, were superficial with one input and one output layer and a hidden layer between them. Each node layer in a deep-learning network trains on a range of features depending on the earlier layer output. An ideal will be well-organized if hidden layers can learn complex characteristics from seen data. Deep neural networks outperform previously unknown data. CNNs utilize a variant of the multilayer perceptron to meet the criterion of minimum preprocessing. CNNs are made up of an automated feature extractor and a trainable classifier with several layers, such as

- (i) Convolutional Layer (CL),
- (ii) Pooling Layer (PL),
- (iii) Fully Connected Layer (FCL).

3.4.1. CNN Architecture. CLs and PLs are used in basic CNN models, which offer a common architecture. CNNs apply a sequence of convolution operations on the input, with or without pooling and a nonlinearity activation function, and then send the output to the next layer. The CL employs filters (F) to extract important characteristics from the input picture for further processing. Each filter provides a unique property for accurate prediction. To maintain the image's size, the (zero padding) same padding is utilized; otherwise, its assistances decrease the sum of features. Each CL's convolutional output may be represented as

$$\text{Out} = \frac{(L_{in} + 2 \times \text{Pad} - F)}{S} + 1, \quad (3)$$

where L_{in} is identified as an input length, Out is identified as a length of the output, S is identified as a stride to filter slide, and Pad is identified as a padding.

A CL comprises generally 3-dimensional input ($H_{in} \times W_{in} \times C_{in}$), where the height and width of the input are H_{in} and W_{in} and the channels of the input are C_{in} . The calculation of the output function for each layer is exactly the same. The CLs produce parameters, neurons, and various connections, as mentioned earlier.

$$P = W_t + B, \quad (4)$$

where B is a Bias, W_t is a weight of the CLs, and P is a parameter. A CL weight can be calculated as

$$W_t = C_{out} \times (H_{in} \times W_{in}) \times C_{in}, \quad (5)$$

where C_{out} is output channels of the previous layer.

CLs and PLs are used in basic CNN models to offer a common architecture. CNNs use a succession of convolution operations to the input, coupled with/without pooling and a nonlinearity activation function, before sending the output to the subsequent layer. The filters (F) are used in the CL to extract important characteristics from the input picture for further processing. Each filter provides a unique property for precise estimate. To keep the picture size the same, identical padding (zero padding) is used; then, valid padding is used since its assistance decreases the sum of features. Each CL's convolutional output may be written as

$$\begin{aligned} W_{out} &= \frac{(W_{in} - F)}{S} + 1, \\ H_{out} &= \frac{(H_{in} - F)}{S} + 1, \end{aligned} \quad (6)$$

where W_{out} and H_{out} are height and width of the output, and W_{in} and H_{in} are height and weight of the input.

The weights and parameters change their values when data flows over a deep network, occasionally making the data too huge or too small. This is called the problem of the "interior covariate shift." A modification in the distribution of the domain of a function is called a covariate shift. Each layer entry is affected in deep networks by its parameters, which means that a slight network change can affect the whole network. Such internal layer modifications may cause the deep network to experience an internal covariate shift

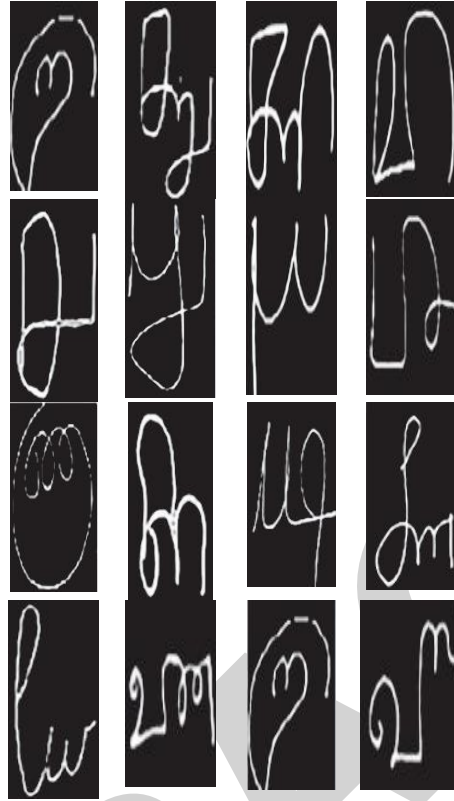


FIGURE 9: Illustration of the unique palm-leaf manuscript representing characters attained as an output with the typescripts.

issue. This problem is generally overcome by normalizing the data in each mini-batch. The general formulae of Batch normalization (BN) are presented in the next section. It is used to allow every network layer to do learning more independently and used to normalize the output of the previous layer. In addition, overfitting is minimized, because it has a slight regularization effect. BN operates on 4D inputs, which may be represented as a small batch of 3D inputs. During training, this layer retained a running evaluation of its calculated mean and variance with a momentum of 0.1.

$$\mu[x_i] = \frac{1}{n} \sum_{i=1}^n x_i, \quad (7)$$

where $\mu[x_i]$ is the mini-batch mean value, x_i is represented as the value of the mini-batch elements, and

$$\sigma^2[x_i] = \frac{1}{n} \sum_{i=1}^n (x_i - [x_i])^2, \quad (8)$$

where $\sigma[x_i]$ is the mini-batch variance values:

$$\hat{x}_i = \frac{x_i - \mu[x_i]}{\sqrt{\sigma^2[x_i]}} + \epsilon, \quad (9)$$

where \hat{x}_i is the mini-batch normalization:

$$Y_{\text{out}} = \gamma \hat{x}_i + \beta, \quad (10)$$

where Y_{out} is the mini-batch normalized value, and γ and β are the learnable parameters:

$$f(x) = \max(x, 0), \quad (11)$$

where $f(x)$ denotes the nonlinear activation function. The goal of FCL is to use the characteristics of the CLs and PLs to classify pictures into distinct classes based on the training dataset.

As previously mentioned, a computationally simple and efficient cursive character specific neural network system is the aim. To accomplish this, a six-layer architecture is proposed, with each layer essentially including CL > Batch Normalization (BN) > ReLU > Max Pooling Layer (Max-PL) (see Figure 10). PLs are employed in the projected model to lower the spatial dimensions, which implies that they will lessen the number of parameters inside the model, with a process known as downsampling or subsampling. Following each PL layer, BN is employed to deal with the interior covariate shift difficult. As the network becomes deeper, it may become trapped in the saturation area, resulting in a vanishing gradient issue/problem. To address this, the planned network employs Rectified Linear Unit (ReLU). The complex patterns in the input data are learned by the network using activation function, and ReLU is considered as an activation function in this model, where it does not activate all the neurons at the same time. The goal of FCL is to use the convolutional and PL features to categorize the input picture into multiple classes depending on the training dataset.

The subsections that follow provide a brief explanation of the suggested CNN model construction.

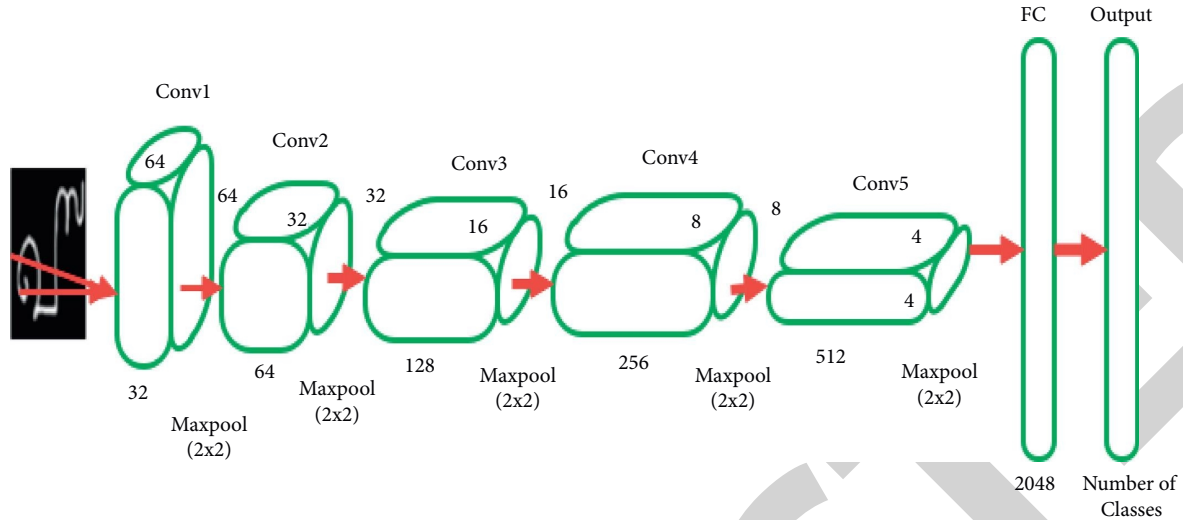


FIGURE 10: Proposed modified CNN architecture—ThimuNet.

(1) *ThimuNet Architecture.* In the preceding sections, a fundamental introduction of CNN architecture was presented. This section discusses the proposed CNN architecture “ThimuNet,” which is meant to identify handwritten cursive characters. The goal is to create a CNN model that can learn differentiating features for handwritten character identification in a fraction of the time and space required by present models. Figure 10 depicts ThimuNet’s architecture, which consists of six levels. The input image has a resolution of 64×64 pixels. First, the supplied image is scaled to (64×64) pixels in size. The first layer then receives picture pixels as input. Each CL alternates with a subsampling or pooling layer, which takes the pooled maps as input.

CL, Max-PL, and FCL are denoted as C_x , M_x , and FC_x correspondingly, in the following discussion, where x is the layer index. The input for the first CL, C_1 , will be $64 \times 64 \times 1$. The output of C_1 (“C” after numerals represents layer number), ReLU, and batch normalization is applied. As a consequence, 22 subsampling processes were performed. M_2 is a max-PL with 32 feature maps of 32×32 dimensions. Table 1 shows the calculation of trainable characteristics and trainable links. The C_3 layer has 64 feature maps. C_3 ’s output feature maps are linked to M_4 . ReLU and BN are implemented after each CL layer. C_5 , C_7 , and C_9 , like the previous convolutional layers, convolute the prior feature maps. Max-PLs are linked to each subsequent CL.

The output volume will be $512 (2 \times 2 \times 512)$ after the final layer. As a result, they can extract characteristics that are more resistant to local changes of the input pictures. This FC layer has been flattened to $1 \times 1 \times 2048$, used to compute the class scores. This entire layer scheme is similar to a conventional feedforward network by overlapping the map element and kernel element input on each specific layer and adding results together for the output to that particular layer after calculation. Gradients of error have been calculated via backpropagation. Gradient descent was employed to update all network weights (equation (5)) and parameter values

(equation (4)) to reduce output errors. A concise architectural indication is as follows: $32C_1 - BN - RELU - 2P_2 - 64C_1 - BN - RELU - 2P_2 - 128C_1 - BN - RELU - 2P_2 - 256C_1 - BN - RELU - 2P_2 - 512C_1 - BN - RELU - 2P_2 - 2048FC$ OUTPUT, where XC_y denotes a CL with a total of X kernels and a stride of y pixels in this representation, MP_n signifies a Max-PL 8 with a $M \times M$ pooling window and a stride of n pixels, BN signifies a Batch normalization 18, and NFC signifies a FC layer with n neurons.

4. Results and Discussion

As previously stated, the objective of this study is to provide a superior CNN model than the current models for the gathered cursive dataset. To build the suggested neural network design, PyTorch7 was used as the Python-based framework. On these gathered datasets, evaluation of many modern models such as LeNet5, ResNet (18/34/50), AlexNet, DenseNet121, InceptionNet v3, and others was also conducted to have a comparative analysis. All of these trials are carried out on a system equipped with Intel Core i3 CPUs, 16 GB of RAM, and an NVIDIA graphics card with 4 GB of internal memory and 768 CUDA cores.

4.1. Result Evaluation of Proposed ThimuNet. This section validates the proposed CNN performance with other techniques in terms of various parameter metrics. The major parameters such as Precision, Balanced Classification Rate (BCR), Recall, Sensitivity, Misclassification Penalty Metric (MPM), Specificity, Balanced Error Rate (BER), F-measure based on sensitivity and specificity, Peak Signal to Noise ratio (PSNR), and Distance Reciprocal Distortion (DRD) can be computed for the Tamil Palm leaf manuscripts. However, in this work, only a few parameters are considered for the validation process; the reason is that it is a collected cursive Tamil Palm Leaf manuscript and cannot apply all major parameters.

TABLE 1: Summarization of different parameters of the proposed ThimuNet.

Convolutional layer	Parameters	Neurons	Connections	Activations
Layer 0- INPUT (64 × 64) 3 channels	0	0	0	64 × 64 × 3 A = 12,288
Layer 1- FILTER (5 × 5) 32 outputs	Wt = 32 × (5 × 5) × 3 B = 32 P = (2400 + 32) = 2,432	(64 × 64) × 32 N = 1,31,072	C = 32,76,800 (64 × 64) × 32 × (5 × 5)	Conv = 64 × 64 × 32 pool = 32 × 32 × 32 BN = 64 × 64 × 32 A = 294,912
Layer 2- FILTER (5 × 5) 64 outputs	Wt = 64 × (5 × 5) × 32 B = 64 P = (51200 + 64) = 51,264	(64 × 64) × 64 N = 2,62,144	(64 × 64) × 64 × (5 × 5) C = 65,53,600 (64 × 64) × 128 × (5 × 5)	Conv = 32 × 32 × 64 BN = 32 × 32 × 64 pool = 16 × 16 × 64 A = 147,456 Conv = 16 × 16 × 128 BN = 16 × 16 × 128
Layer 3- FILTER (5 × 5) 128 outputs	Wt = 128 × (5 × 5) × 64 B = 128 P = (204800 + 128) = 2,04,928	(64 × 64) × 128 N = 5,24,288	C = 1,31,07,200 (64 × 64) × 256 × (5 × 5)	pool = 8 × 8 × 128 A = 73,728 Conv = 8 × 8 × 256 BN = 8 × 8 × 256
Layer 4- FILTER (5 × 5) 256 outputs	Wt = 256 × (5 × 5) × 128 B = 256 P = (819200 + 256) = 8,19,456	(64 × 64) × 256 N = 10,48,576	C = 2,62,14,400 (64 × 64) × 512 × (5 × 5)	pool = 4 × 4 × 256 A = 36,864 Conv = 4 × 4 × 512 BN = 4 × 4 × 512 pool = 2 × 2 × 512 A = 18,432
Layer 5- FILTER (5 × 5) 512 outputs	Wt = 512 × (5 × 5) × 256 B = 512 P = (3276800 + 512) = 32,77,312	(64 × 64) × 512 N = 20,97,152	C = 5,24,28,800	
Fully connected layer (FC layer)—46 class outputs	Wt = 512 × (2 × 2) × 46 B = 46 P = 94,208	0	0	A = 46

$$\begin{aligned}
 \text{Accuracy} &= \frac{(xp + xn)}{(xp + yp + xn + yn)}, \\
 \frac{\text{Recall}}{\text{Sensitivity}} &= \frac{xp}{(xp + yn)}, \\
 \text{Precision} &= \frac{xp}{(xp + yp)}, \\
 \text{F1 - Measure} &= 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}},
 \end{aligned} \tag{12}$$

where xp is denoted as True positive, xn is denoted as true negative, yp is described as false positive, and yn is represented as false negative.

The reason for comparing various pretrained models on CNN is that they are significantly more accurate than the custom-built model of CNN. In addition, the pretrained models can effectively train on large datasets and directly use the weights and architecture obtained to detect the cursive characters of Tamil on palm leaves. LeNet is the base for all ConvNets, which is mainly used for detecting the handwritten characters. ResNet is used to solve the vanishing gradient problem by making the CNN to construct with more than thousands of convolution layer without the increase of training error percentage and outperform the shallower networks. AlexNet is developed by using eight layers with learnable parameters, which has the ability to leverage GPU for training and being able to train with vast numbers of parameters. In DenseNet, each layer is connected to every other layer, and it also alleviates the vanishing gradient problem. Moreover, the features are reused in this model, feature propagation is strengthened and minimized the number of parameters. Initially, the proposed CNN performance is validated with the existing technique in terms of PSNR and BER [23], tabulated in Table 2 and Figure 11.

The performance of the technique is improved when the PSNR level is high and BER is low. For instance, LeNet and AlexNet have low PSNR values (5.77 and 7.08), whereas ResNet and DenseNet achieved nearly 16 of PSNR. But the proposed ThimuNet achieved 17.74 of PSNR. The existing techniques such as LeNet and AlexNet achieved high BER (i.e., nearly 28.5), where ResNet and DenseNet achieved 11.9 and 7.52 of BER. The proposed ThimuNet achieved very low BER, i.e., 6.06, when compared with other existing techniques. The reason is that the proposed ThimuNet is well adapted with all input features. Table 3 and Figure 12 show the experimental results of proposed ThimuNet on precision, recall, sensitivity, specificity, and F-measure.

In the precision experiments, the LeNet achieved 63%, AlexNet achieved 84%, ResNet achieved 91%, and DenseNet achieved 94%, but the proposed ThimuNet achieved only 92%. The reason is that some features are wrongly classified, and some collected cursive writings are not properly removed by using preprocessing techniques. After that, recall and sensitivity experiments are carried out to validate the performance of the proposed ThimuNet. The existing technique LeNet achieved 65%, ResNet achieved 79%, AlexNet achieved 47%, and DenseNet achieved 86%, whereas the proposed ThimuNet achieved 89% recall and sensitivity. This proves that the proposed net achieved better performance than existing techniques. Finally, specificity and F-measure experiments are performed. All the existing techniques achieved nearly 85% to 90% of F-measure and 93% to 99% of specificity, where the proposed ThimuNet achieved 90.51% of F-measure and 99% of specificity. Table 4 and Figure 13 provide the comparative analysis of proposed ThimuNet with existing techniques in terms of accuracy.

The above table and Figure clearly prove that the proposed ThimuNet achieved higher performance than existing techniques in terms of accuracy. For example, the LeNet achieved 71% accuracy, ResNet achieved 88% accuracy,

TABLE 2: Validated results of proposed model with existing techniques.

Network type	PSNR	BER
LeNet	5.77	28.42
ResNet	15.69	11.9
AlexNet	7.08	29.74
DenseNet	16.08	7.52
ThimuNet	17.74	6.06

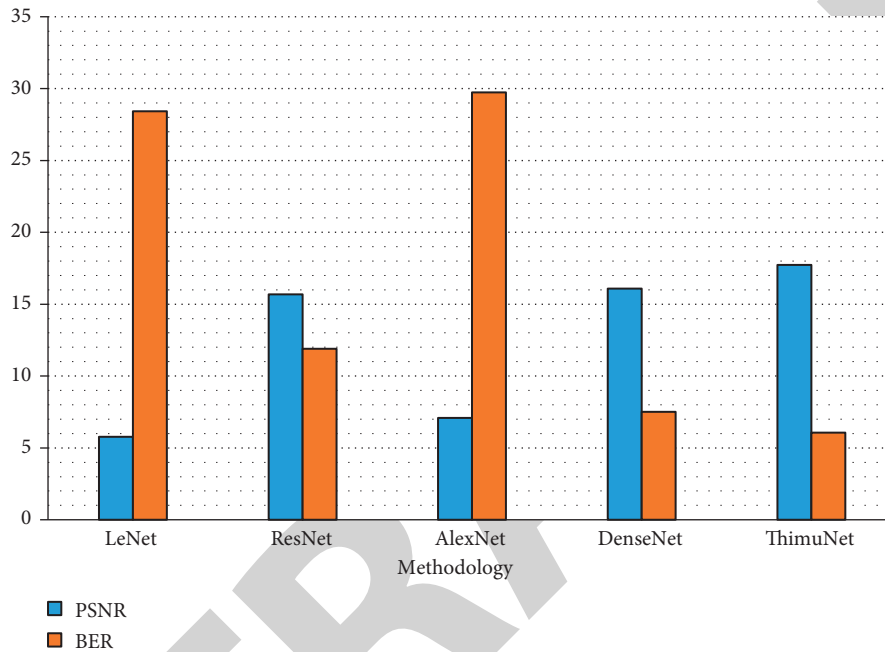


FIGURE 11: Graphical representation of proposed ThimuNet in terms of PSNR and BER.

TABLE 3: Experimental results of proposed ThimuNet with existing techniques.

Network type	Precision (%)	Recall (%)	Sensitivity (%)	Specificity (%)	F-measure (%)
LeNet	63	65	65	79	63.87
ResNet	91	79	79	99	85.01
AlexNet	84	47	47	93	57.45
DenseNet	94	85	86	99	90.14
ThimuNet	92	89	89	99	90.51

AlexNet achieved 66% accuracy, DenseNet achieved 92% accuracy, and the proposed ThimuNet achieved 94% accuracy. Finally, the prediction time of proposed ThimuNet for classifying the cursive Tamil Palm Leaves Manuscript is given in Table 5 and Figure 14.

The existing LeNet, AlexNet, and ResNet prediction time achieved nearly 1.08 seconds, while the DenseNet achieved

0.95 seconds for the classification. However, the proposed ThimuNet achieved less prediction time (0.80 seconds) than existing techniques. The proposed ThimuNet achieved better performance from all experimental analysis and classified wrong cursive letters due to irrelevant features. This problem can be solved by incorporating efficient feature selection techniques with proposed ThimuNet.

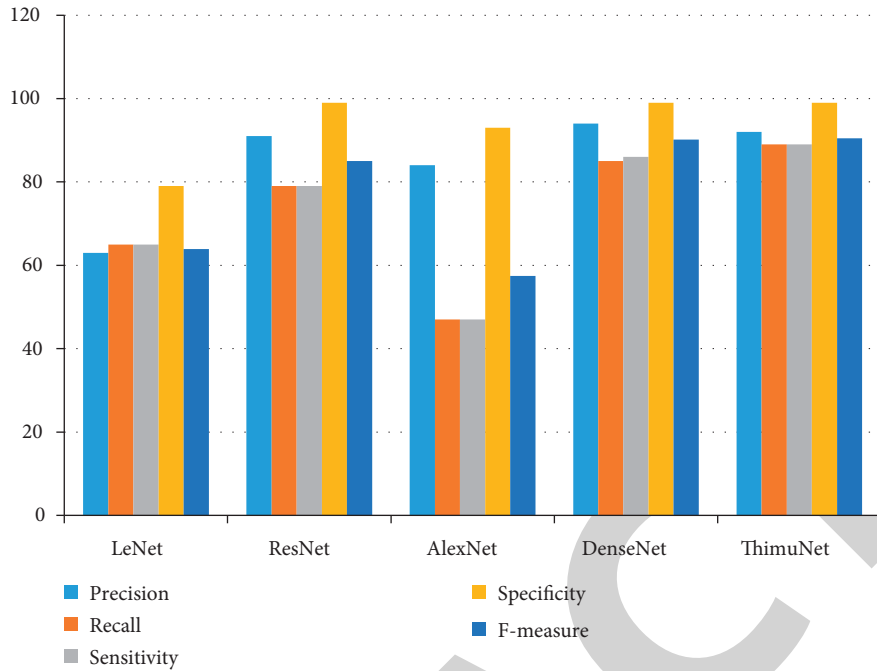


FIGURE 12: Performance analysis of proposed ThimuNet with existing techniques.

TABLE 4: Comparative analysis of proposed ThimuNet on the basis of accuracy.

Network type	Accuracy (%)
LeNet	71
ResNet	88
AlexNet	66
DenseNet	92
ThimuNet	94

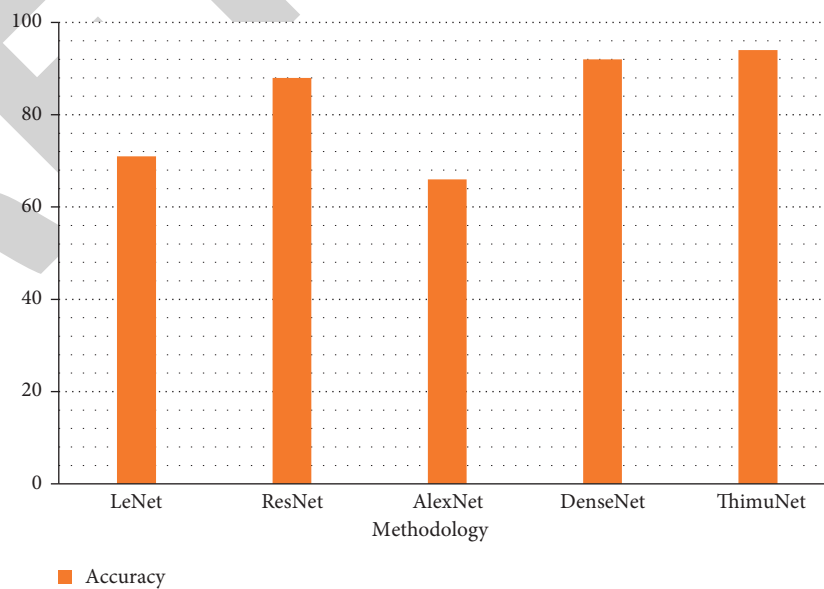


FIGURE 13: Graphical representation of proposed ThimuNet in terms of accuracy.

TABLE 5: Prediction time of proposed ThimuNet for classification of cursive Tamil palm leaves manuscript.

Network type	Prediction time (s)
LeNet	1.06
ResNet	1.09
AlexNet	1.04
DenseNet	0.95
ThimuNet	0.80

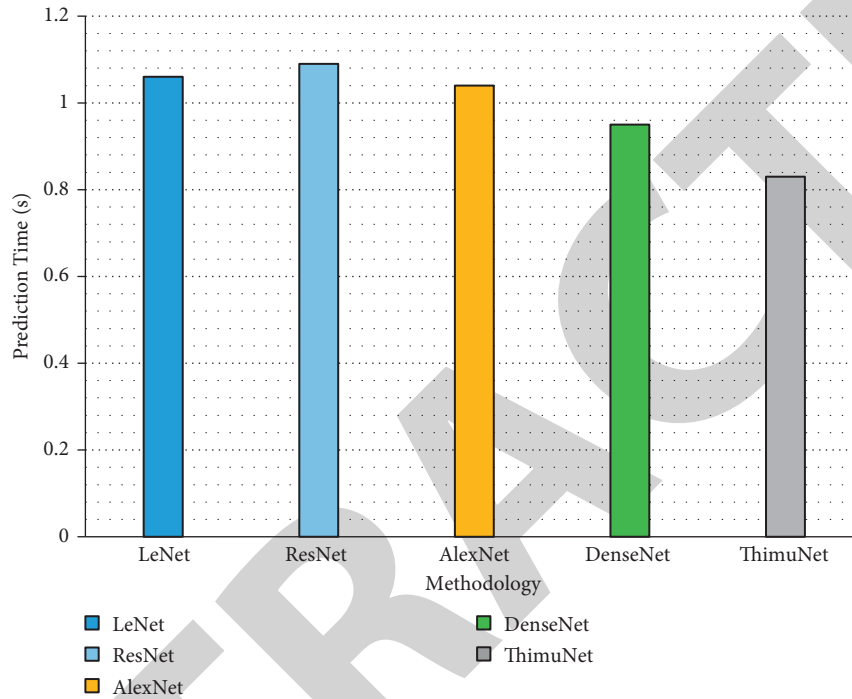


FIGURE 14: Comparative analysis of proposed ThimuNet in terms of prediction time.

5. Conclusion

Tamil is the world's oldest languages and has attainment popularity among scholars throughout the world due to its historical relevance and capacity to persist for centuries. Individual interests and ideologies of academics have shaped study on palm leaves manuscripts in Tamil literature over the periods and in a variety of ways. The most prevalent topic of research in the actual world is cursive character recognition of ancient Tamil characteristics, and it focuses on the documentation of Tamil characters to guarantee that a large amount of data is gathered. However, due to improper care, palm leaf manuscripts have been destroyed, as numerous people held them by default. As a result, a rigorous examination of Tamil palm leaf texts, translation, and cataloguing of their publishing were necessary. Thousands of palm leaf manuscripts exist in the hands of institutes and separate healers, but they must be digitized, and proper catalogues are developed for the future. Similarly, numerous researches on the cursive handwriting recognition of the Tamil language are included in the survey. There is, however, no universal technique for identifying cursive Tamil letters (with sufficient accuracy) in their whole. As a result, various

approaches have been used in each stage of the recognition process.

This work is one of the first attempts to build cursive training datasets by manual and automated segregation of Tamil palm leaf scripts. Future researchers could use the data set to develop expert systems that could then be used for a range of purposes, such as a character classification based on the century they have evolved and the identification of extinguished characters, and the identification of characters whose shape could have changed. The research attempted to segregate ancient Tamil palm leaf scripts to produce a massive amount of Tamil cursive character information. There have been no previous investigations on the identification of cursive characters in Tamil palm leaf scripts. The ThimuNet of CNN is used in this research project to classify cursive Tamil characters, and its performance is confirmed. In order to improve the performance of proposed model, the weight of the input parameters must be optimized, which is considered as future scope of this research model. In addition, preprocessing plays a major role in removing unwanted cursive writings; therefore, an effective preprocessing technique is implemented along with the developed model for better performance.

Retraction

Retracted: Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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. Gupta, S. Kumar, S. L. Bangare, S. Nuhmani, A. C. Alguno, and I. A. Samori, "Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3490860, 9 pages, 2022.

Research Article

Homogeneous Decision Community Extraction Based on End-User Mental Behavior on Social Media

Suneet Gupta ¹, Sumit Kumar ², Sunil L. Bangare ³, Shibili Nuhmani ⁴,
Arnold C. Alguno ⁵ and Issah Abubakari Samori ⁶

¹Department of CSE, School of Engineering and Technology, Mody University, Lakshmanagarh, Rajasthan, India

²Indian Institute of Management, Kozhikode, India

³Department of Information Technology, Sinhgad Academy of Engineering, Savitribai Phule Pune University, Pune, India

⁴Department of Physical Therapy, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

⁵Department of Physics, Mindanao State University, Iligan Institute of Technology, Tibanga Highway, Iligan, Philippines

⁶School of Engineering Sciences, University of Ghana, Accra, Ghana

Correspondence should be addressed to Issah Abubakari Samori; iasamori@st.ug.edu.gh

Received 30 January 2022; Revised 9 February 2022; Accepted 21 February 2022; Published 8 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Suneet Gupta 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.

Aiming at the inadequacy of the group decision-making method with the current attribute value as interval language information, an interval binary semantic decision-making method is proposed, which considers the decision maker's psychological behavior. The scope of this research is that this paper is based on localized amplification method. The localized amplification method used in this research may amplify physiological movement after removing unwanted noise, allowing the movement trend to be seen with the naked eye, improving the CNN network's mental identification accuracy. These two algorithms analyze the input picture from various perspectives, allowing the CNN network to extract more information and enhance identification accuracy. A new distance formula with interval binary semantics closer to decision-makers thinking habits is defined; time degree is introduced. An optimization model is established to solve the time series weights by considering the comprehensive consistency of expert evaluation. Based on prospect theory, a prospect deviation value is constructed and minimized weight optimization model, using the interactive multiple attribute decision community making (TODIM) method based on the new distance measure to calculate the total overall dominance of the schemes to rank the schemes. Taking the selection and evaluation of supply chain collaboration partners as an example, the effectiveness and rationality of the proposed method are verified.

1. Introduction

With the increasing complexity of decision-making problems in various fields, individual decision-making has long been unable to meet the requirements of scientific decision-making, and group decision-making has attracted more and more attention and attention from experts and scholars [1, 2]. For multiattribute group decision-making, most existing research is based on the classical expected utility theory, which assumes that the decision-maker is entirely rational, but this is not in line with reality. Prospect theory [3, 4] believes that decision-makers have systematic perception bias in the decision-making process. Decision-makers do

not continuously pursue the maximum utility in behavior but show reference dependence and loss aversion. Therefore, it is necessary to consider the psychological behavior of decision community makers in the decision-making process.

In addition, in the entire group decision-making process, due to the limitations of gender and decision makers' cognition, decision-makers are more inclined to give evaluations in natural language. [5] proposed a binary semantic information evaluation model, which attracted the attention of scholars at home and abroad. At present, the research based on binary semantics mainly focuses on two aspects: on the one hand, the research on binary semantics set counters [6–10]. However, in the actual decision-making process, due

to the ambiguity of decision-making information and the limitations of decision-makers cognition, decision-makers are often more willing to give evaluation information in the form of interval language to reduce decision-making pressure. In response to such problems, [11] showed the definition of interval binary semantics and several set counters and used the interval binary semantic possibility formula to sort the solutions; reference [12], based on the maximum dispersion [13], defined a new interval binary semantic Bonferroni average operator and its corresponding weighting method; reference [14] combined interval binary semantics and VIKOR method, and proposed an outsourcing supplier selection method; reference [15] proposed a subway door failure risk assessment method based on interval binary semantics and failure mode. The method proposed in the literature [11–15] has the following shortcomings: first, the evaluation information is uniformly distributed in the interval by default, and the distribution of information in the gap that is more in line with the psychology of decision-makers is not considered; second, most of the existing research in the single static stage, [16] it is not suitable for the situation that requires dynamic multistage analysis; thirdly, the above methods all assume that the decision-maker is entirely rational, [17] but in the actual decision-making, the decision-makers psychological behavior is often bounded rationality. The impact of social media on mental health is stress, anxiety, and sadness which have all been related to the social network use. Individuals who often use digital platforms, according to latest studies cited by The Young Mind Centre and The National Centre for Health Research, are more sad and unhappy with life than others who stay longer on nonscreen-related activities. To sum up, based on existing research, this paper proposes an interval binary semantic dynamic group decision-making that considers the decision maker's psychological behavior for the multiattribute group decision-making problem in which the attribute value is interval language, and the attribute weight and time series weight are entirely unknown method. Considering the thinking habits of human beings, it is believed that the density of evaluation information in the interval is more similar to the normal distribution, and based on this, a new interval binary semantic distance formula is proposed. Solve the time series weights based on the time degree and entropy; [18, 19] by determining the positive ideal scheme under each time series, establish a prospect theoretical profit and loss matrix, and build a linear programming model to minimize the sum of the squares of the foreground deviation values, and determine the attribute weights under each time series; then, Construct an ITL-TODIM method based on interval binary semantic distance measure to calculate the complete overall dominance of each candidate scheme, and determine the pros and cons of the system according to the total general authority; the evaluation is used as an example to verify the effectiveness and rationality of the proposed method. The ITL-TODIM approach which is based on intermediate binary semantic distance measure to compute the total overall dominance of each candidate scheme and identify the system's advantages and disadvantages in terms of total general authority. The

evaluation is used as an example to check the suggested method's efficacy and reasonableness. The TODIM technique completely analyses managements' risk aversion perspective and may represent key stakeholders' risk preference by altering the dimensions, which is more in accordance with the actual decision-making needs. Any two possibilities, however, must be evaluated using the TODIM approach, which has a significant computational cost. Psychological behavior is one of the main ways people express their emotions. Research on various expression recognitions has made significant progress [1–6]. In recent years, the recognition of spontaneous expressions has become a new research hotspot [7, 8], and psychological expressions are often generated when people want to suppress their feelings, which can neither be faked nor suppressed [7–9]. The complete expression usually lasts 0.5~4 s [10], which is relatively easy to be recognized by people. However, psychology believes that when a person tries to hide his genuine emotions, occasionally, there are emotions that leak out. Psychological behaviors were first discovered in 1966 [20]. Three years later, [21] used the term psychological behaviors when analyzing a video interview of a patient who attempted suicide. Psychological behaviors usually change uncontrollably between 1/25 and 1/2 s [22], and the frequency of occurrence is low, and untrained individuals do not have high recognition ability [23]. Therefore, the results reported by different researchers also vary considerably [11, 24]. After this, Ekman and Friesen proposed the Brief Affect Recognition Test (BART) in 1979 [12]. In subsequent experiments, they found that the subjects' psychological-expression recognition ability was positively correlated with lie recognition ability [13]. Afterwards, the Japanese and Caucasian Brief Affect Recognition Test (JACBART) [14, 15] was conducted, verifying that the subjects' psychological-expression recognition ability was positively correlated with lie recognition ability [20], it can be proved that psychological-expressions can effectively help people identify lies. Face psychological expression recognition involves image processing and analysis, computer vision, artificial intelligence, psychology, biology, and other directions.

2. Literature Survey

In 2002, Authors [24] developed a psychological-expression recognition tool, METT (Psychological behavior Training Tool). Studies have shown that METT tools can improve an individual's ability to recognize psychological expressions by 30% to 40% on average. In addition, an Action Coding System (FACS) [21] is also designed, according to the anatomical characteristics of the face, it is divided into several independent and interconnected motion units (Action Unit, AU), The motion characteristics of these motor units and the main areas they control and the expressions associated with them are analyzed, and a lot of photo descriptions are given. Although human emotions are complex and diverse, they can still be divided into 6 basic emotional categories [22, 23]. Therefore, researchers combine different motor units to form FACS codes to correspond to different expressions, mainly divided into happy, angry, fearful, sad, surprised, and

others. When people express their inner state and psychological needs, they will produce many psychological behaviors [24]. Still, because psychological expressions exist for too short a time and are not easily detected by the human eye, computers can be used to solve this problem.

3. Psychological-Expression Recognition Technology

Psychological expression is the tiny movement change of the human face, including texture change. The movements of psychological expressions are too small and of short duration to be easily captured by the human eye. The physical character of all substances is dictated by their real physical composition, which is referred to as texture. Textures can elicit behavioral processes in people. This mental reaction enables us to feel something without really touching it. The blood circulation in our faces alters as we experience various emotions. This causes tiny colour shifts that other persons might see. People can properly determine someone's sentiments from these visual shifts up to 75% of the time, according to recent research. Therefore, the psychological expressions of the face can be studied through machine vision. According to the above characteristics, several mainstream methods for psychological-expression recognition are Convolution Neural Networks (CNN) [11], Optical Flow (Optical Flow) method, and Local Binary Pattern (LBP).

3.1. Convolution Neural Networks. Convolution Neural Networks (CNN) [11] has been widely used in various fields such as machine vision and speech recognition since their birth. Usually, CNN is used as feature extraction and depth feature extraction for image class input, and the desired output results can be obtained after analyzing the extracted features. Relatively small movements characterize Psychological-expressions. If convolution neural networks are used, it is usually necessary to use other auxiliary methods to change the input of the web, or to change and optimize the network structure so that the network can extract more valuable features, thereby improving the recognition of psychological-expressions accuracy.

3.2. Local Binary Patterns and Improvement Methods. Local Binary Patterns (LBP) [12] can effectively deal with illumination changes and are widely used in texture analysis, texture recognition, and other fields, with grey and rotation invariance degeneration and other significant advantages. The LBP value of the central pixel reflects the texture information of the surrounding area of the pixel, as shown in Figure 1.

This feature is widely used because of its simplicity and ease of computation. However, the traditional LBP algorithm cannot be applied to video signals for video images, so some improvements to the LBP algorithm are needed. Among them, [25] proposed a robust dynamic texture descriptor that performs local binary patterns from three orthogonal planes (LBP-TOP), widely used for psychological behavior. To consider both

the spatial and temporal information of the video, LBP-TOP extends LBP. Compared to LBP, this method finds three types (XY , XT , YT) instead of one plane (spatial XY). Given a video sequence, it can be viewed as a stack of XY , XT , and YT planes along the temporal T -axis, spatial Y -axis, and spatial X -axis, respectively. The three histograms come from three planes, respectively, and are concatenated into one histogram as a dynamic video texture descriptor, as shown in Figure 2. LBP-TOP extends the application of the LBP algorithm to a higher dimension and can identify textures in time series. The information between the frames before and after is correlated through this method, making the LBP algorithm more widely used. Local Binary Patterns (LBP) are frequently employed in texture analysis, texture identification, and other domains and have major benefits such as grayscale and rotation invariance degradation. The texture knowledge of the pixel's surrounding area is reflected in the central pixel's LBP value. Any radius and number of neighborhood pixels may be achieved by using a circular neighborhood and subsection linear interpolation data at noninteger pixel locations. The complimentary comparison measure might be the grey scale variation of the immediate region.

They generate fairly extensive descriptive statistics, which slow down identification speed, particularly on huge face databases. (2) They overlook the spatial patterns in some circumstances because they do not examine the influence of the central pixel.

The properties of LBP can be found in LBP-TOP, and vice versa. LBP-TOP is unaffected if the pixel value increases or lowers by the same amount. This method can be employed in real life despite the interference to the image created by the natural parallel light environment, when in, the robustness is very high. However, because it is pixel-based, this method requires more expertise in Computer Science and Engineering. The first use of LBP is to compare the centre pixel to P pixels in the vicinity of radius R . When R is 1, the centre pixel and the surrounding 8 pixels have a size of $2P$. The final size is 28×256 . R is no longer a single number, and the size of the neighborhood grows exponentially. The LBP operator's mode type is utilized for dimensionality reduction in [26]'s proposal to employ a "equivalent pattern" (Uniform Pattern) as a solution to this problem. There are two transitions from 0 to 1 or 1 to 0, according to Ojala et al., in natural images. As a result, the "equivalent mode" is defined as having at most two transitions from 0 to 1 or 1 to 0 in a cyclic binary integer. An analogous Pattern class refers to the binary that corresponds to the LBP. It takes 256 to get down to 58 using this strategy. In other words, the values are divided into 59 categories, with the 58 uniform patterns constituting one and the remaining values constituting the 59th. Consequently, the histogram's 256-dimensional dimension is reduced to 59 dimensions using this method, rather than the original $2P$'s $P(P-1)+2$. This decreases the influence of high-frequency noise on the eigenvectors by making them less dimensional. Using LBP-TOP, an image is divided into 59×3 blocks, with each block generating an array with a size of 3.59. This is because LBP-TOP applies LBP in the time dimension. The finished feature's dimensions are $4 \times 4 \times 59 \times 3 = 2.832$. LBP-TOP is a high-

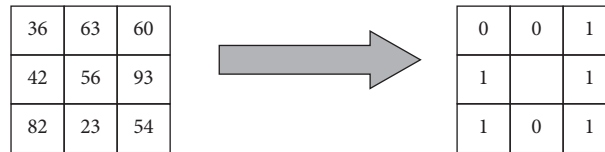


FIGURE 1: LBP.

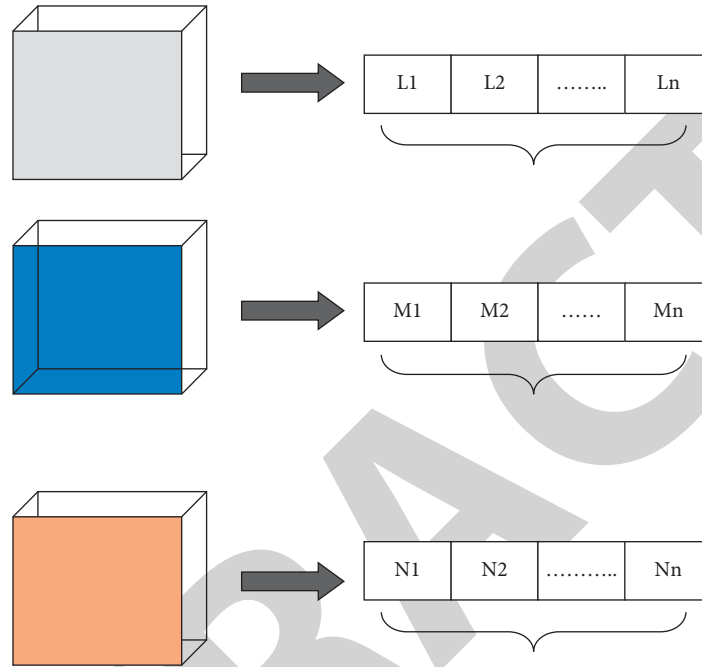


FIGURE 2: LBP-TOP.

dimensional characteristic that has a significant impact on calculation speed and accuracy.

4. Mental Behavior for Psychological Change

Psychological-expression recognition has made some progress at present, but there is still a lot of room for improvement in technology. The following is an analysis and prospect of the problems existing in different methods of psychological expression:

- (a) Convolution neural network used as a Traditional method is widely used, and the current mainstream idea is to use the optical flow method as input and combine it with LSTM to achieve better results. However, there is a fatal problem in using convolution not high. For convolution neural networks, it is a small sample problem. Therefore, data enhancement is required during calculation, or Cross-domain experiments, while improving robustness. In the subsequent process, you can try to use transfer learning, which can alleviate the problem of overfitting to a certain extent. In addition, you can also use methods such as adversarial neural networks to generate some samples for training, which can alleviate the problem of small sample data.

- (b) There are three main methods of feature screening: filter, wrapper, and embedded. The filtering feature selection method assigns weights to the features of each dimension, and then sorts the features according to the weight; the encapsulating feature selection method views subset selection as an optimization problem, generates different combinations, evaluates the combinations, and then compares them with others. By comparing the two approaches, we can determine which features are most critical for a specific model to be trained with. There are three basic screening procedures, and each is designed for a certain scenario. The statistical performance of all training samples is directly used to evaluate the relevance of each feature in the filtering feature selection technique. Although its performance is excellent, this algorithm has a lot of advantages, including the ability to remove a large number of unnecessary attributes, its universality, and the ability to prescreen features. It is necessary to combine the feature selection approach with the subsequent classification algorithm, evaluate each feature's significance based on the classifier's accuracy, and select the ideal feature subset for a certain classification algorithm. It is similar to the filtering approach, but the integrated feature screening

method uses machine learning training rather than relying solely on the statistical indications of a feature to evaluate its advantages and disadvantages. Machine learning is also used in the embedding method, as opposed to the packaging method. With this approach, all of the features are used for training, rather than just a subset of them. By utilizing the embedded feature selection method, both the feature selection and learner training are done in the same optimization process. Each of the three ways has its own pros and weaknesses; therefore they can all be utilized together. An ideal subset of features is picked out based on preprocessing and classifier approach advantages in order to increase the accuracy of psychological-expression recognition.

5. Psychological-Expression Recognition Method

5.1. Method for Psychological-Expression Recognition Based on CNN. CNN is a commonly used method when processing image signals. Still, because the psychological-expression changes are relatively small, the effect of only using CNN is not very good, so some literature choose to process the input image. The traditional psychological-expression feature extraction needs to consider the spatial image features and the temporal sequence. The amount of pixels used to generate a digital image is referred to as spatial resolution. Higher spatial resolution images include more pixels than lower spatial resolution images. Object (matrix) and picture (graph) modes are available for representing spatial objects. Each spatial object can be specified as a spot, line, or polygon in object mode. Each spatial object can be specified in picture mode as a collection of adjacent cells called regions. Generally, feature extraction is performed on all frames of the psychological expression from the beginning to the end, while proposing a psychological expression. Stage classifier: First, the psychological expressions are divided into three stages: onset, peak, and offset using temporal information, and then spatial information is used to detect intensity changes. Compared with the traditional psychological-expression feature extraction method, this method only uses a psychological-expression (Apex) frame and an initial (Onset) frame to extract features, increasing the number of valid frames reducing the amount of computation and outperforms other traditional methods. In addition, [30] proposed a video magnification method based on eye interference elimination and used convolution neural network to realize the task of psychological-expression recognition. First, zoom in on the data to extract the eye position coordinates. After that, the original eye video will be replaced with the enlarged video for image fusion to eliminate eye interference. Finally, the convolution neural network model network is designed using the idea of VGG16 to realize emotion recognition. The local amplification technology applied in this paper can amplify the psychological-expression movement after reducing the noise interference so that the movement trend can be seen by the naked eye, which can increase the accuracy of the CNN network for psychological-expression

recognition. These two methods process the input image from different angles so that the CNN network can better extract features and improve the accuracy of recognition. They generate fairly extensive descriptive statistics, which slow down identification speed, particularly on huge face databases. (2) They overlook the spatial patterns in some circumstances because they do not examine the influence of the central pixel.

At present, most of the psychological-expression feature extraction based on CNN is combined with the optical flow method extraction, which is also one of the advantages of the CNN framework. However, if CNN is used for deep learning, it is still unrealistic for the current data set. The amount of data in the existing data sets cannot meet the needs of the deep learning network, and it is easy to cause overfitting, which leads to the huge size of CNN. The advantage cannot be played.

5.2. Psychological-Expression Recognition Method Based on Optical Flow Method. FACS requires motion records of various positions such as eyebrows and mouth corners as an essential tool for recognizing psychological expressions. This method uses ROIS and HOOF features together, and the obtained results correspond to AU, which can identify psychological expressions more accurately. Psychological-expression detection is often independent, and a long-video expression automatic recognition method consisting of macros and psychological expressions (time segmentation). This method utilizes the stress generated by nonrigid motion on the skin during the expression process. It uses the central difference method to calculate the solid and dense optical flow field observed in several areas of each subject's face (chin, mouth, cheeks, forehead). Strain level: This method can successfully detect and distinguish psychological expressions and fast local psychological expressions, which is of great help to the detection of psychological expressions in complex scenes.

5.3. Method of Psychological-Expression Recognition. According to the problems of LBP-TOP, STCLQP adds information such as amplitude and direction components, extracts more helpful information, and fuses the extracted data into a feature vector, which improves the utilization of image features. Compared with LBPTOP, the STCLQP method extracts more information and thus results in higher dimensionality. In this paper, the LQP technique is used. Compared with the LBP, whose dimension increases exponentially with the neighborhood radius R , LQP maps the extracted features to the lookup table. When R increases, the dimension no longer shows exponential growth. Therefore, this method is more suitable when the R -value needs to be improved, and because the LQP technology alleviates some of the previously extracted more dimensional information.

LBP-SIP reduces redundancy in LBP-TOP mode, provides a more compact and lightweight representation, improves accuracy, and reduces computational complexity. Unlike LBP-TOP, this method only uses 4 pixels at the top, bottom, left, and right of a pixel at time t and 6 pixels at the

same position at time $t-1$ and $t+1$. Because the extracted features are few, it is more suitable when there are many images and a long time, and the operation is faster than LBP-TOP. Its speed is 2.8 times that of LBPTOP. The feature extraction time of this method is 15.88 s, which is about 2.4 s faster than that of the LBPTOP method, and the recognition time is 0.208 s, which is 0.3 s more quickly than the LBP-TOP method. Compared with the LBP-TOP method, LBP-SIP reduces a considerable amount of time, but this time is still too long for subsequent practical applications to achieve real-time detection. But it provides an idea to reduce the processing time. In addition, to reduce the dimension of the features extracted by LBP-TOP and then use SVM for classification. This method first uses the LBP-TOP operator to extract psychological-expression features. It then proposes a feature selection algorithm based on Relief combined with a Locally Linear embedded (LLE) manifold learning algorithm, which can effectively reduce the number of postextraction features. Feature dimension: Finally, the SVM classifier with Radial Basis Function (RBF) kernel is used for classification, and good results are obtained. This method is based on the LBP-TOP method for feature vector extraction, so the application environment is consistent with LBP-TOP. The filter method is used to exclude irrelevant features, and then the wrapper method is used to filter out the features with greater influence, which avoids the dimensional disaster and reduces the amount of calculation accordingly. This method gives a direction for future development. It is a revelation that for the pixel-by-pixel feature extraction method of LBP-TOP, the extracted features can be screened to improve the recognition accuracy during classification.

For the redundant features of LBP-TOP, the above methods adopt different ideas for dimensionality reduction, but it is far from enough for practical application in real life. The method of pixel-by-pixel feature extraction is easy to cause dimensional disasters. The selection of features and feature extraction methods is particularly important before. It is only when the processing speed is reduced that there is an opportunity to apply such methods to social life.

5.3.1. CNN (Convolution Neural Network). CNN is an excellent method for dealing with image problems, but the network structure needs to be adjusted for different issues. When using CNN, feature extraction is usually performed first, and then further processing is performed after the features are extracted. However, the optical flow method and the improved local binary mode method use different removal methods. For the feature map, CNN can continue to extract deep features, so the use of CNN and optical flow method and local binary mode improvement method does not conflict with itself. CNNs are often used in conjunction with a popular form of recurrent neural network called the Long Short-Term Memory (LSTM) module for video-type files.

5.3.2. Applicable Scene Analysis of CNN. Optical flow method and improved local binary mode of the existing psychological-expression recognition methods are all based

on datasets. The generalization ability of psychological-expression recognition, in reality, is not high and cannot be applied. However, the required environment and application scenarios can be estimated according to the characteristics of existing methods. As far as the current development is concerned, if the multiple existing networks are integrated into the detection of psychological-expression recognition, complex scenes can be detected, such as shopping malls, streets with more critical locations, prisons, etc. Existing psychological-expression recognition methods can be combined with other networks. It can automatically detect various features of pedestrians in the crowd to ensure the accuracy is improved.

Based on the assumption of the consistent grey level of the optical flow method, the visual flow method cannot detect scenes or objects with changing brightness and has stricter requirements on the environment. The camera cannot rotate too fast. Therefore, psychological-expression recognition can be performed in interrogation rooms and negotiation rooms where the brightness is relatively fixed, and the target does not need to move. Psychological-expressions are based on the premise that the target suppresses their expressions, but when faced with a specific scene, the mark may have the problem of combining psychological-expressions and psychological-expressions. The work proposed by Shreve et al. [36] proposed a solution for this type of problem by using the optical flow method, making psychological-expression recognition more suitable for more stringent places such as interrogation rooms and optical flow rooms. Compared with other methods, the extracted feature dimension of the optical flow method is smaller, and it is the most viable method for real-time detection. For local binary mode improvement methods, such as pixel-by-pixel analysis methods such as LBP-TOP, the overall brightness change of the environment has little effect on it, so it can be used as a complement to the optical flow method, but its obvious disadvantage is that in the extracted feature dimension, the number is too high and the computational burden is too large for real-time detection, so it can be used as a behind-the-scenes tool for psychological-expression calibration and recognition. With the current research, after the feature dimension extracted by such methods is reduced to real time, it may be necessary to extract specific parts of the target's face, and it is necessary to minimize the occlusion of the key parts of the target to be detected. Therefore, such methods are used in it that may be more suitable for areas where the behavior and clothing of the target are strictly managed, such as detention centers or the military.

LSTM for video-type inputs has a better effect. Rather than using CNN alone, it is more appropriate to use CNN as a framework, and CNN can be combined with various methods to make the improved network recognition more effective. An LSTM differs from a CNN in that it is often used to attempt to predict performance, and computational CNN, on the other hand, is meant to identify "spatial patterns" in information and performs well on pictures and sounds. For example, the TV-L1 method extracts the optical flow image and superimposes the original image as input while using the CNN + LSTM method and then sends the extracted features to the LSTM network for psychological-expression recognition. After using the ROI method, it

TABLE 1: Performance comparison of CNN methods.

Method	Accuracy	F1-score
CNN + LSTM	60.98	65.85
ELRCN-SE	47.15	49.52
ELRCN-TE	52.44	55.63

TABLE 2: Communities comparison of CNN methods.

Method	Modularity's	NMI
CNN + LSTM	56.63	85.63
ELRCN-SE	46.52	75.62
ELRCN-TE	53.23	80.56

TABLE 3: Performance comparison of improved local binary mode methods.

Method	Accuracy	F1-score
LBP-TOP	57.16	59.63
STLBP-IP	59.51	62.45
STLBP-IIP	62.75	66.48

TABLE 4: Performance of community extraction by local binary mode methods.

Method	Modularity's	NMI
LBP-TOP	66.63	89.24
STLBP-IP	52.63	80.54
STLBP-IIP	56.62	82.26

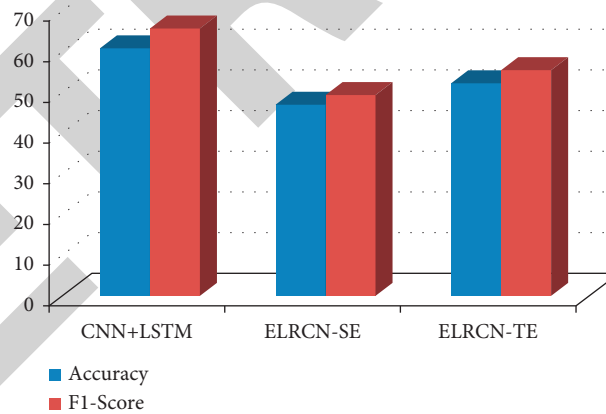


FIGURE 3: Performance comparison of CNN methods.

cooperates with the Flow Net 2.0 optical flow method to identify the visual flow of a specific area and finally enables the ROI + Revised HOOF to recognize different expressions with FACS. These two methods use other optical flow methods. Still, the main body of the TVL1+LSTM method is CNN, and the visual flow method is only used as the standard input with the original image to increase the recognition accuracy, from Tables 1–4. It can be seen that CNN, as a framework, is very inclusive and can be used in combination with a variety of methods. Therefore, it is a good tool for solving image problems. The performance

comparison of CNN methods is shown graphically in Figure 3.

Figure 4 shows the community features of the proposed binary method. The improved local binary mode methods performance is shown in Figure 5 and the community features for local binary mode methods are shown in Figure 6.

But the optical flow method is essential for the environment. The requirements of the database are more stringent, and it needs to be modified on this basis before it can be applied to parts outside the database.

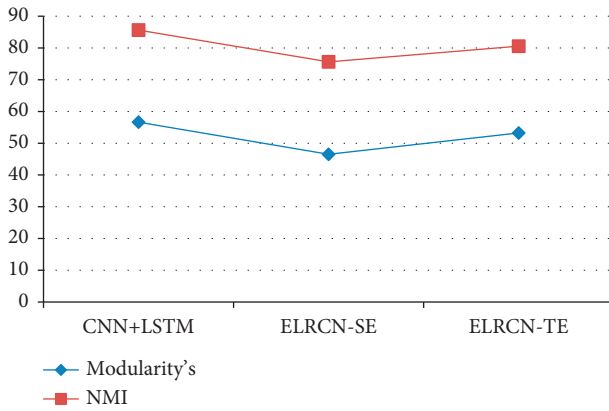


FIGURE 4: Community feature of proposed methods.

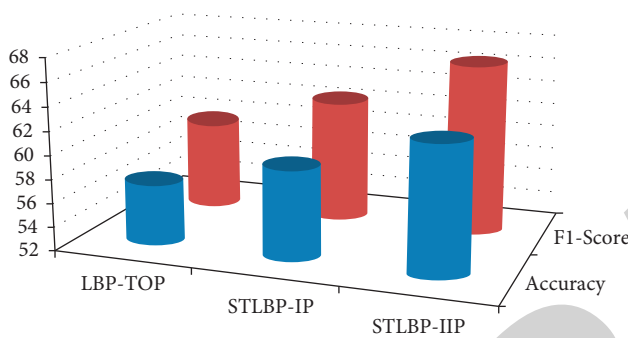


FIGURE 5: Performance comparisons of improved local binary mode methods.

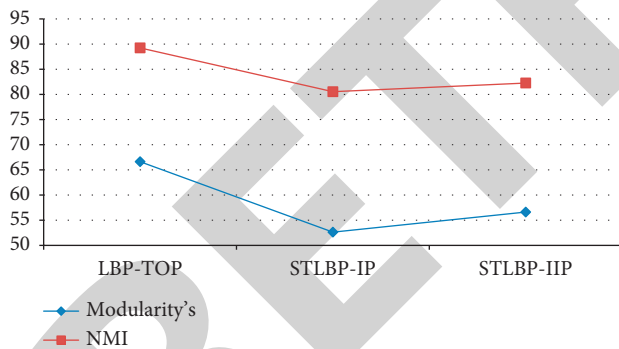


FIGURE 6: Community Feature of local binary mode methods.

6. Conclusion

Convolution neural network and its improvement, optical flow method and its improvement, and local binary pattern and its improvement are discussed in this study. The recognition of psychological-expression information can also be improved through improvements to the convolution neural network itself. Convolution neural networks can be useful for psychological-expression recognition when they have been trained to extract visual features. After decreasing random noise, the local amplification method used in this study may magnify the psychological-expression movement such that the development trend can be perceived with the

unaided eye. The optical flow method is a popular method for psychological-expression identification because it can identify small moving targets, which is ideal for detecting love. However, there are also evident issues. It is necessary to refine and optimize the algorithm in order to ensure that the brightness of the identified target remains consistent, which is sometimes difficult to do in actual applications. Optical flow, on the other hand, is a reliable method for detecting face psychological expressions. Using three planes, LBP-TOP can accurately detect videos. The update of LBP-TOP minimizes the number of redundant parameters, improves the accuracy of recognizing texture features, and can do all three at once. So, using the upgraded method of LBP for psychological-expression detection would be a good idea because it improves texture recognition.

Data Availability

The data shall be made available upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] H.-H. Shuai, "A comprehensive study on social network mental disorders detection via online social media mining," *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 7, pp. 1212–1225, 2018.
- [2] H. Yazdavar, M. S. Mahdavinjad, G. Bajaj, K. Thirunarayan, J. Pathak, and A. Sheth, "Mental health analysis via social media data," in *Proceedings of the 2018 IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 459–460, New York, NY, USA, June 2018.
- [3] Y. Chai, F. Wu, R. Sun et al., "Predicting future alleviation of mental illness in social media: an empathy-based social network perspective," in *Proceedings of the 2019 IEEE International Conference on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, pp. 1564–1571, Xiamen, China, December 2019.
- [4] J. Bai, Q. Kong, L. Li, L. Wang, and D. Zeng, "Exploring cognitive dissonance on social media," in *Proceedings of the 2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 143–145, Shenzhen, China, July 2019.
- [5] M. E. Aragon, A. P. Lopez-Monroy, L.-C. G. Gonzalez-Gurrola, and M. Montes, "Detecting mental disorders in social media through emotional patterns—the case of anorexia and depression," *IEEE Transactions on Affective Computing*, 2021.
- [6] M. Hu and X. Wu, "Research on children's network privacy protection in mobile social media," in *Proceedings of the 2018 IEEE 20th International Conference on High Performance Computing and Communications; IEEE 16th International Conference on Smart City; IEEE 4th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pp. 1135–1138, Exeter, UK, June 2018.
- [7] U. Iqbal and A. Hussain Mir, "Secure and practical access control mechanism for WSN with node privacy," *Journal of*

Retraction

Retracted: A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] K. Zeberga, M. Attique, B. Shah, F. Ali, Y. Z. Jembre, and T. Chung, "A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 7893775, 18 pages, 2022.

Research Article

A Novel Text Mining Approach for Mental Health Prediction Using Bi-LSTM and BERT Model

Kamil Zeberga ¹, Muhammad Attique ², Babar Shah ³, Farman Ali ²,
Yalew Zelalem Jembre ⁴ and Tae-Sun Chung ¹

¹Department of Artificial Intelligence, Ajou University, Suwon, Republic of Korea

²Department of Software, Sejong University, Republic of Korea

³College of Technological Innovation, Zayed University, Abu Dhabi, UAE

⁴Department of Electronic Engineering, Keimyung University, Daegu, Republic of Korea

Correspondence should be addressed to Tae-Sun Chung; tchung@ajou.ac.kr

Received 4 November 2021; Accepted 30 December 2021; Published 3 March 2022

Academic Editor: Syed Ahmad Chan Bukhari

Copyright © 2022 Kamil Zeberga 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 current advancement in the Internet, there has been a growing demand for building intelligent and smart systems that can efficiently address the detection of health-related problems on social media, such as the detection of depression and anxiety. These types of systems, which are mainly dependent on machine learning techniques, must be able to deal with obtaining the semantic and syntactic meaning of texts posted by users on social media. The data generated by users on social media contains unstructured and unpredictable content. Several systems based on machine learning and social media platforms have recently been introduced to identify health-related problems. However, the text representation and deep learning techniques employed provide only limited information and knowledge about the different texts posted by users. This is owing to a lack of long-term dependencies between each word in the entire text and a lack of proper exploitation of recent deep learning schemes. In this paper, we propose a novel framework to efficiently and effectively identify depression and anxiety-related posts while maintaining the contextual and semantic meaning of the words used in the whole corpus when applying bidirectional encoder representations from transformers (BERT). In addition, we propose a knowledge distillation technique, which is a recent technique for transferring knowledge from a large pretrained model (BERT) to a smaller model to boost performance and accuracy. We also devised our own data collection framework from Reddit and Twitter, which are the most common social media sites. Finally, we employed word2vec and BERT with Bi-LSTM to effectively analyze and detect depression and anxiety signs from social media posts. Our system surpasses other state-of-the-art methods and achieves an accuracy of 98% using the knowledge distillation technique.

1. Introduction

The automatic detection of mental health conditions is one of the most important and complex health concerns in the real world. Mental health affects the behavior, thinking, and mood of individuals interacting with the world around them. In addition, mental health problems are becoming a leading disability, contributing largely to the universal burden of disease. In general, the number of people with signs of depression in 2015 was projected to be 4.4% (more than 332 million people) [1]. As reported in a WHO study, depression is a shared universal mental disorder that affects a large number of people regardless of their age. There are

many limitations in depression recognition and treatment, including a lack of professionals in the health sector, social shaming, or an inappropriate diagnosis. Prolonged depression and anxiety can lead to suicide if the affected individual is not provided good care and immediate help. In addition, depressive disorders have been categorized as one of the largest contributors to nonfatal health loss. Suicide has become a major reason for death among young people with a general suicide rate of 10.5 per 100,000 people, which is nearly 800,000 cases every year in absolute measure [2, 3].

The rapid integration of smart sensors in hand phones and wearable devices has increased access to intelligent mental healthcare, permitting the gathering of measurable

signs in a clear and modest way, giving a reasonable estimate of the physical and psychological states of users [4, 5]. Many types of sensors can be incorporated into the mental healthcare process and can provide accurate, momentary, and continuous patient data [6]. Until recent days, many systems in the medical sector were designed to supervise the mental health of users using smartphones and wearable sensors as a data source [4, 7]. However, such systems are not positioned to gather valuable information on demand keeping the freshness of data. It is also known that digital devices yield a large portion of mental-health-problem-related data, which are not enough to efficiently supervise patients. Moreover, obtaining meaningful information from these data and efficiently examining them has become extremely challenging for traditional mental health problem detection systems.

Social media has recently become a persuasive tool to inspect the mental health and mental state of the users, particularly the youth. It also provides anonymous contributions in numerous online platforms to leave room for open dialog regarding socially defamed topics and motivate users to fight against mental health problems [8]. In addition, patients can share their ideas about the recent common health problems. Healthcare monitoring systems for the detection of depression and anxiety can apply social media data to recognize the mental state of users based on their posts and comments. However, the data published on social networking sites on mental-health-related problems contain unstructured, unpredictable data composed of idioms, jargon, and dynamic topics. Thus, it is becoming very difficult for an application to fetch the desired data about patients and evaluate them to confirm that they receive the proper treatment they need as early as possible. As a result, there should be a smart approach that is capable of retrieving the most valuable data features with a minimal dimensionality that maximizes the accuracy of systems in the healthcare environment on mental health.

Machine learning algorithms such as decision trees, support vector machines (SVMs), logistic regression, Ada-Boost, and multilayer perceptron (MLP) have been used to support doctors in diagnosing symptoms of depression and anxiety [9–13]. However, unremitting patient monitoring gives a large portion of health-related data containing voice patterns, textual data, sensor data, and emojis, all of which have increased significantly [14]. The existing systems that apply machine learning are not capable of dealing with these types of data to precisely get meaningful evidence and cannot capture semantic significance in text posts disseminated in social media. Moreover, such data will remain useless for healthcare businesses until they are managed and classified wisely in time-sensitive manner [15, 16]. This demands a smart approach that can accurately classify the textual data related to mental health problems [3, 9, 17, 18].

The concept of deep learning addressed many key issues in natural language processing tasks such as sentiment analysis, and the current sentiment analysis on social media content is increasingly focused on the use of new deep learning architectures and models for mental problem identification and classification [19–21]. However, the new

development in health sector systems poses a huge burden for researchers because they require collaboration with interdisciplinary steps, new technologies, and changes in society. As a result, these previous healthcare systems are inefficient and ineffective in coping with these new trends.

This study proposes a smart and context-aware deep learning framework based on bidirectional encoder representations from transformers (BERT) to effectively identify mental-health-related problems from user posts on social media with improved classification accuracy. This study combines different sources of information for an efficient analysis of mental-problem-related data. We have adopted a knowledge distillation scheme to transfer knowledge from large pretrained BERT to a smaller model and used bidirectional long short-term memory (Bi-LSTM) to examine depression- and anxiety-related data. The results indicate that our presented system precisely handles the mixed data and enhances the performance of mental health classification. In general, our key contributions to this research work are fourfold:

- (i) A new framework is presented to extract a huge size of highly appropriate depression- and anxiety-related data from Twitter and Reddit. In addition, we implemented a combined cyber-community-group-based labeling and keyword-based data crawling technique based on the circumplex model of emotion to identify the desired mental health problem data.
- (ii) A deep neural network-based bidirectional text representation model, that is, BERT, is used to embody mental health problem textual data maintaining contextual and semantic connotations. In addition, we proposed a sequence processing model called bidirectional long short-term memory (Bi-LSTM) as a classifier, which effectively maximizes the amount of information accessible to the network, improving the content available to the algorithm in knowing what words immediately follow and come before a given word in a sentence.
- (iii) We propose a knowledge distillation technique, which is a means of transferring knowledge from a large pretrained model (BERT) to a smaller model to maximize performance and accuracy. We filtered the large network (BERT) into another much smaller network (Distiled_BERT) for mental health-related problem identification, and it performs very well by transferring the required domain knowledge and applying it to a specific healthcare environment.
- (iv) We conducted extensive experiments using a principal component analysis (PCA) and different deep learning/ML models, the results of which are compared with other related models. This evaluation plays a key role in regulating the shortcomings of the already applied methods and classification models. The experimental results show that our model performs considerably well over the

compared methods, which, after many hyperparameter optimizations, provides an accuracy of 98%.

The remaining sections are organized as follows: Section 2 contains a description of mental health monitoring system using wearable devices and a deep learning approach. Section 3 presents the whole framework proposed in this research. Section 4 explains the results obtained in our experiment. Finally, in Section 5, we provide a conclusion of our study.

2. Related Studies

Machine learning (ML) models and big data play important roles in building a smart healthcare monitoring system for patients. Recently intelligent devices such as cellphones and many wearable sensors have been converged to generate the maximum evidence-based mental health data possible. Furthermore, the rapid growth of social media platforms and their application has increased at an unprecedented pace. This section discusses the detection of common mental illnesses based on wearable sensors using ML practices, social media data, and deep learning approaches using large sets of data.

2.1. Analysis of Mental-Health-Related Problems Using Wearable Sensors. The growing capacity of intelligent devices like cell phones makes them a prospective way for ecological momentary assessment measurements and the monitoring, treatment, and interventions of mental illness. This will reduce costs and help expand the mental health service for the larger societal group. The most commonly used smart devices in our day-to-day life such as cell phones and fitness bands contain sensors [22]. This extends the chance of many applications that consume information generated from sensors in the healthcare domain [23]. We can merge the data generated by these sensors to produce contextual information about patients regarding their mental status and social relationships [24]. Moreover, these types of multiple sensor integration have been used to get better results in different application areas compared to single sensors [25].

However, accumulated wearable sensor-based information exists in large volumes and is not well-structured. The previous research works did not apply innovative data processing techniques to get the desired latent information in them. Different studies, particularly on sensors, have been conducted on existing techniques to collect data, applying ML techniques to process the collected data [22]. Nevertheless, the previously proposed research works were not very successful in many different scenarios. Furthermore, mental health diagnosis is difficult to realize on a large scale because of the old data collection techniques such as interviews and questionnaires [26]. These approaches are unscalable to reach larger societal groups within a certain community. In addition, it is difficult to use them for knowledge extraction without designing an efficient and scalable system that can operate on a large set of data.

Therefore, health organizations have moved away from conventional connections and now allow online group meetings for sharing information and seeking advice, thereby helping scale their approach to a certain extent. Recent studies have indicated that many social groups are willing to contribute and share ideas about health-related issues [9].

2.2. Analysis of Mental-Health-Related Problems Using Social Media Data with Machine Learning. Social media platforms recently are considered as backbones for the detection of mental health conditions. However, detecting depression through online social media is extremely challenging because it demands a well-designed and robust system that can deal with the complex nature of the data. It is very difficult to get an important and appropriate quantity of data related to mental illness. In recent years, the practice of utilizing social media data has boomed, and people started revealing their concerns without hesitation. This actually motivates researchers to conduct more research on the detection of health-related problems such as depression and anxiety as early as possible. In addition, data found in social media is replete with vague information, which mainly includes related metadata such as location, age, and other factors [27]. Kowsari et al. [28] proposed a deep learning model to analyze the mental status of patients based on Reddit posts. Many of the previous systems have applied insufficient data sets for depression and anxiety, which may produce imprecise results and thus perhaps mislead healthcare workers [29]. Artificial intelligence (AI) based techniques have been proposed to examine patients' posts on social media platforms and recognize serious problems in those patients [27]. An anxiety-related dictionary is constructed, and a given text is evaluated on the basis of the dictionary whether related to anxiety or not. However, the nature of texts posted on these platforms is always unorganized, and it is impractical to efficiently process them without using deep learning algorithms.

Tadesse et al. developed the idea of processing a given post to check as to whether it contains the idea of suicide through deep learning and ML-based classification methods targeting Reddit data. The authors make use of an LSTM-CNN merged to compare it to different working approaches. There are commonly two steps to analyze social media data. First, data are collected from different sources such as networking sites, and the second task is to process the available data using statistical models. We proposed a combined cyber community-group-based labeling and keyword-based data crawling technique based on the circumplex model of emotion to identify the desired mental health problem data, as depicted in Figure 1 [30].

2.3. Deep-Learning-Based Text Embedding and Classification. Text representation is one of the underlying difficulties in sentiment analysis. The purpose of a text representation is to numerically denote unstructured text documents to make them mathematically commutable by maintaining the semantic and circumstantial meaning of the words in the text.

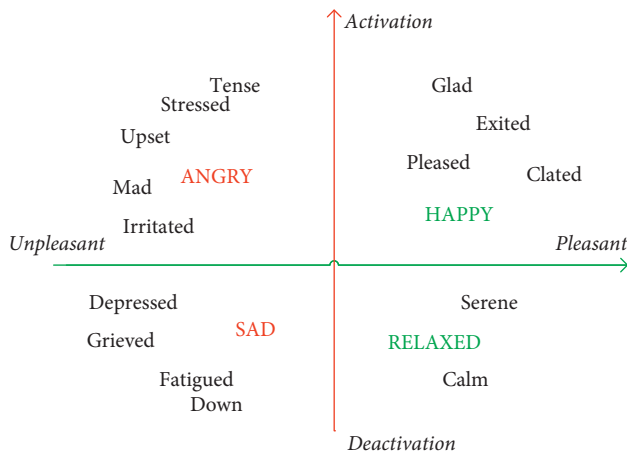


FIGURE 1: Graphical representation of the circumplex model of affect with the horizontal axis representing the valence or pleasant dimension and the vertical axis representing the arousal or activation dimension.

Different ways of representing a given text corpus in a real-valued vector were studied and proposed by researchers [31]. Any given word in a text should be converted into a single vector so that it will be fed to the artificial neural network. As we can see from Figure 2, many studies have been conducted on word embedding techniques based on statistical methods for learning an individual word embedding from a text corpus such as word2vec, fastText, and GloVe. In addition, recent transformer-based deep learning techniques for representing words such as BERT, GPT, and GPT2 are extremely common [32, 33]. These models bring a major change to NLP downstream tasks by practicing predicting missing words in the text, and because they analyze every sentence with no specific direction, they do a better job at understanding the meaning of homonyms than previous ML approaches such as word embedding methods. Deep learning methods contain concurrent processing layers to represent data hierarchically and have exhibited good results in many domains. Young et al. reviewed many deep-learning-based practices that are used for various natural language processing (NLP) tasks [34]. Many recent studies have used deep learning to solve problems related to understanding the sole meaning of a given post, such as sentiment. However, all of these deep learning-based text representation techniques lack adequately labeled data sets for mental health problems. In addition, applying the best-fit word embedding approach to vectorize a very large data set has not been adequately addressed. We proposed a BERT-based text representation technique that efficiently and effectively captures the semantic meaning of words in a given text based on the attention mechanism [35].

The other key challenge in dealing with sentiment analysis is the implementation of appropriate deep-learning-based classifier models. Researchers have implemented different classifier models, such as CNN and XGBoost, along with their own text-preprocessing techniques. The major challenge in NLP, in general, is capturing the semantic and syntactic meanings of a word in a largely given text corpus,

which is generally termed as maintaining long-range dependency. Many studies have used a classical feature analysis, such as a CNN, an LSTM, and an LSTM-CNN merged to detect the idea of suicide in online forums [36, 37]. In this regard, the authors did not incorporate many recent deep learning techniques that deal with the representation of words along with their context. Researchers have also proposed a GRU and an LSTM-RNN, which are suitable for processing long textual data and applying them to the task of sentiment analysis [18, 38, 39]. Another interesting approach proposed by Burdisso et al. is a new way of dealing with texts, called smoothness, significance, and sanction (SS3), aiming to offer assistance for sentiment analysis in an integrated, easy, and efficient way. The classification process in most previous techniques was not self-explaining, and human beings are unable to naturally understand the motivations behind the classification [40]. The authors developed SS3 to address the incremental classification of chronological data and help with early classification and explainability. Nonetheless, most such research methods follow the same pattern when training the model. They first train their model on emotions using extracted textual data and then predict whether a given unobserved text is related to depression. Because the semantic and contextual relationships (long-term dependency) between words are properly captured, the models provide insufficient information and an inaccurate prediction. Therefore, their system attains minimal accuracy for sentiment classification.

More recently, a range of model compression techniques has been established. We are using any model with the primary objective of making a very good generalization on unseen data after training it with a meaningful and sufficient data set. We need deeper models to train a large data set. However, we only need a lighter model that works well on any unobserved data set during testing. As a result, researchers have proposed knowledge distillation that effectively incorporates a smaller and light model from a large and complex model [34, 41]. As far as our research work is concerned, many research studies were done on mental health problems based on user posts. These studies do not leverage deep learning techniques and knowledge distillation effectively for representing collected texts [42, 43].

Finally, the attention mechanism has been one of the most significant developments in deep learning research during the last decade [44]. It has initiated the growth of many recent discoveries in NLP, including the transformer architecture [45] and Google's BERT [19]. Another comparable study associated with depression detection was conducted using a classification of news headlines, which used a manually annotated corpus [38, 45]. Xu et al. also presented long short-term memory (LSTM) neural network to obtain an appropriate meaning from large text [46]. From this literature, we see that researchers integrate classifiers differently with text embedding techniques and data sources to achieve a better result. However, we still need a more robust technique to automatically obtain sensible information from large text data posted on social media platforms keeping better accuracy.

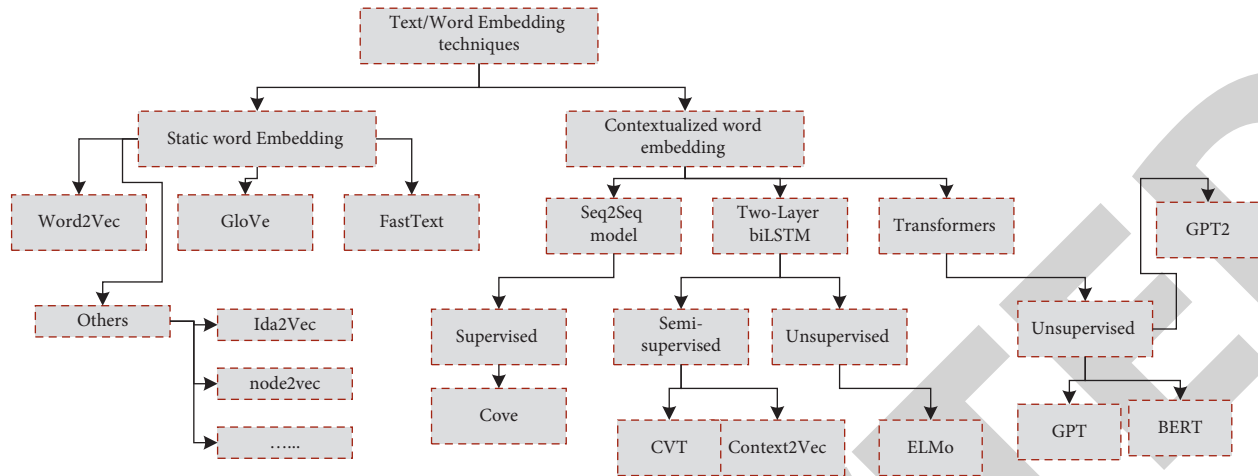


FIGURE 2: Different real-valued vector representation techniques for a predefined large-sized vocabulary from a corpus of text.

This work differs from existing studies in several aspects. First, many of the research works studied were not able to exploit the capability of context-aware deep learning framework based on bidirectional encoder representations from transformers (BERT) to effectively identify mental health-related problems from user's sentiment. Second, the application of knowledge distillation on top of deep learning enhances the effectiveness of our proposed model compared with many of the research works in this specific domain. Third, we have used multiple combinations of vectorization techniques that are listed in Figure 2 to make sure that our proposed approach performs better in maintaining long-term dependencies among a bag of words in a long text corpus that are posts from Reddit and Twitter in our case.

3. Proposed Framework

In this section, we present the framework intended to retrieve, process, evaluate, and classify social networking data about mental health such as depression and anxiety. Figure 3 illustrates a general architecture that is proposed in this research work. Monitoring comments and posts on social media platforms can provide insights into how individuals self-reveal and talk about mental health issues. This type of information disseminated and shared across such platforms can be utilized as a data source for mental health problem identification. However, these types of data collected from social media are usually unstructured and include informal expressions, vague, and contentiously changing topics. It is extremely difficult to generate meaningful information from social media platforms that could be consumed in the domain of mental health evaluation and depression detection. Therefore, the proposed practical framework contains various modules such as data collection, data preprocessing, labeling techniques, word embedding, and classification. The primary objective of this research is to develop an automated system capable of discovering and evaluating mental health conditions using deep learning techniques such as BERT and Bi-LSTM. First, realtime data are extracted based on keyword-based queries from Twitter and Reddit using APIs

(Tweedy and PRAW, respectively), as shown in Figure 3. After the desired number of data are gathered, different NLP techniques are applied to analyze the data (as depicted in task 2 of Figure 3. In addition, we examine the users' thoughts regarding depression and anxiety to categorize their mental health conditions as depression-related (positive) or standard posts (negative), as shown in task 3 in Figure 3. Text representation models, fastText, word2vec, and BERT, are used to embody the data with an appropriate vector. Finally, Bi-LSTM with softmax is trained to classify mental health problems. By analyzing mental-health-related data, as shown in Figure 3, the system helps healthcare monitoring systems automatically identify depression and anxiety such that patients will receive the social support they need.

3.1. Data Collection. This section discusses the data collection procedure from two different sources. We collected our own data sets from Reddit and Twitter using common API wrappers, PRAW and Tweepy. Figure 3 of task 1 shows the entire flow of data collection for the proposed framework. For Twitter, we selected appropriate words preceded by the hashtags symbol, which represents the main theme of content for specific topics. For the Reddit data, we focused on specific subreddits that were suitable for our targeted topics and then performed a search query on such topics. These platforms offer application programming interfaces (APIs) such as PRAW and Tweepy, which allow us to access the data.

3.2. Preprocessing. Data preprocessing is a method of cleaning and filtering noisy and vague data so that they can be easily used for feature extraction, as shown in Figure 4. In real world, people exchange information on social media in an informal way, with texts that contain hashtags, special characters, and needless words. We have to apply the concept of machine learning and data mining in order to extract some sense out of such text corpus before feeding them to any classification model. In addition, substituting

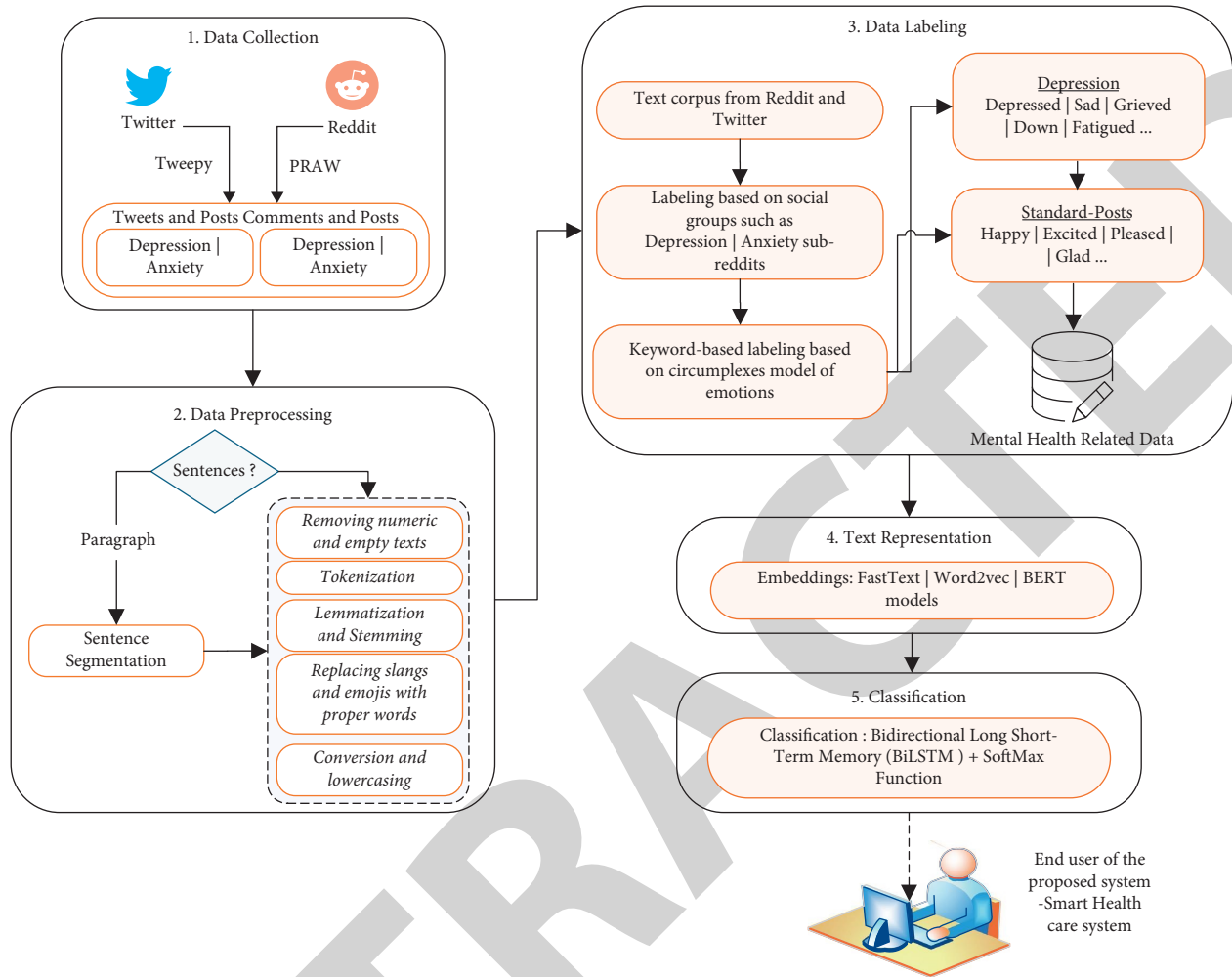


FIGURE 3: These layered schematic diagrams show an organized view of automatic detection of depression and anxiety from user posts. The data collection, preprocessing, and labeling layer deals with generating meaningful data from data sources such as Reddit and Twitter. The text vectorization and classification section deals with the technique of extracting meaningful insights from the labeled data using deep learning.

jargon and emojis with the corresponding informative text using Emojipedia is the main area of our focus. It is well known that the Internet is becoming a communication medium where people frequently use colloquial speech and emojis to convey their opinions and thoughts. Extracting meaningful text from posts in social media helps understand the context and intensify the emotions associated with it. We applied different preprocessing methods to our collected text, which contain many jargons and informal words in a way it will help us identify depression and anxiety, as depicted in Figure 4.

3.2.1. Tokenization. Tokenization is a method of splitting a sizable amount of text into smaller portions, commonly known as tokens. These tokens are utilized to discover some patterns and are taken as an input for the next common steps in the NLP pipeline, such as stemming and lemmatization, as shown in 1 of task 2. In general, a large text is composed of hash signs, punctuation, and characters that are not even

texts. In our proposed system, this process is conducted using the TreebankWordTokenizer that is found in the natural language toolkit (NLTK) to purify the words called tokens. Tokenization is used to reduce nonalphanumeric characters and break down sentences into words. Finally, all the text in the entire given file is represented by a bag of words for further analysis.

3.2.2. Removing Stop Words. In this subsection, we discuss the removal of words that have no significant information on mental-health-related problems in general. The most frequent words considered irrelevant are pronouns, prepositions, symbols (e.g., dates, #), conjunctions, and articles (a, an, and the). In addition, universal resource locators (URLs) in any given text data should be filtered because they have no useful information in text processing. To remove stop words from a given sentence, we first split our text into words and then eliminate the word if it occurs in the list of stop words provided by the NLTK, as described above. Substituting

jargon and emojis with their factual text using Emojipedia is an extremely critical step because they contain helpful information regarding mental health.

3.2.3. Conversion to Lowercase. People usually present their opinions about how they feel using emojis. Some words may be written differently such as “b4,” that is, “before”; “);,” that is, “happy”; and “2moro,” that is, “tomorrow.” Each word in our scenario is converted into its original and generic form, followed by conversion into lowercase, which maintains consistency and avoids confusion during a text analysis.

3.2.4. Part-of-Speech (POS) Tagging. The practice of categorizing words into their parts of speech and marking them accordingly is called part-of-speech tagging (POS). These labels can be nouns, verbs, adjectives, adverbs, determiners, and conjunctions, among others. The main reason behind this step is to discard any POS that has no contribution to the identification of depression and anxiety. In addition, POS tagging recognizes nouns and adjectives, which are considered invaluable indicators for sentiment analysis. We used the NLTK POS-tagger for our proposed system, as shown in Figure 4.

3.3. Text Representation. Word embedding is a technique used to represent words in a given text corpus with a real-valued vector that capture the meaning of the word such that the words that remain in a close neighborhood are anticipated to be identical in meaning. We implemented many word-embedding techniques to represent our collected text corpus (bag of words) into the equivalent vectors. Word vectors are much better ways to represent words than the older techniques, such as a one-hot encoded vector, where the index assigned to each word does not hold any semantic meaning. In addition, word vectors consume much less space than one-hot encoded vectors and also maintain a semantic representation of words. We have summarized many real-valued vector representation techniques, as shown in Figure 2, and most of them are implemented using Python 3.7. Figure 2 shows the word embedding techniques from the old context-free techniques such as word2vec and GloVe to bidirectional contextual representations such as BERT, ELMo, and OpenAI GPT [19, 20]. We propose the recent and smart method, BERT, to obtain the context of a given word in our collected text corpus.

3.3.1. Word2vec. The word2vec algorithm accepts a text corpus and generates an equivalent vector associated with the given input. The generated vector will have high dimensions, and each word included in the text will be given an associated vector in vector space. The vectors are placed in space such that words with similar contextual meaning will stay in closer proximity to one another. However, the word2vec algorithm is unable to represent a new word with a vector if it is not included in the training data set. As shown in Figure 5, the word2vec model has very

famous architectures, that is, continuous bag-of-words and skip-gram, in which the training phase uses one hidden layer. However, the purpose of training each of them is different. CBOW is a model that finds a target word using neighboring context words, whereas the skip-gram model finds context words given a center word as an input. In our proposed system, the skip-gram model was trained to use its improved performance with consistent words. To find word representations that can infer adjacent words within a context c with high accuracy, the skip-gram model for a given series of words $\{w_1, w_2, w_3, \dots, w_n\}$ increases the objective of the average log probability over all N target words and their respective contexts.

$$\frac{1}{N} = \sum_{n=1}^N \sum_{-c \leq j \leq c, j \neq 0} \log P(w_{n+j} | w_n). \quad (1)$$

The probability predicted for a given center word is highly dependent on the inner product of the vectors used to represent the input and output candidates $\{w_I$ and $\{w_O\}$, respectively, normalized to the requirements of a probability distribution over all words in the vocabulary of size N using the softmax function as follows:

$$P(w_o | w_i) = \frac{\exp(v_{w_o}^T v_{w_i})}{\sum_{w=1}^W \exp(v_w^T v_{w_i})}. \quad (2)$$

However, the complexity of determining these probabilities and correlated gradients for all words becomes expensive as the size of the words increase. In most of the previous word embedding techniques depicted in Figure 2, the major drawback is extremely small context, and no global occurrence is used. For the entire data set collected from Reddit and Twitter, we passed each pair into the neural network and trained our model to represent the collected texts.

3.3.2. FastText Model. As the working principle of fastText, the morphological structure of a word contains key information about the meaning of the word. Such a structure was not exploited much by conventional word representation techniques such as word2vec, which trains a distinctive word representation for every single word in the list. This is more useful for morphologically rich languages where a given word may have a larger number of morphological forms, making it difficult to train good word embeddings. FastText tries to resolve this by considering each word as a collection of subwords. For simplification and language independence, subwords are considered as the character n -grams of the word. The vector for a word is simply considered as the sum of all vectors of its component char n -grams. Based on a comprehensive comparison of word2vec and fastText in our experiment, fastText performs considerably better on syntactical tasks as compared to the original word2vec, particularly when there is a smaller number of training data. word2vec marginally outperforms fastText on semantic tasks.

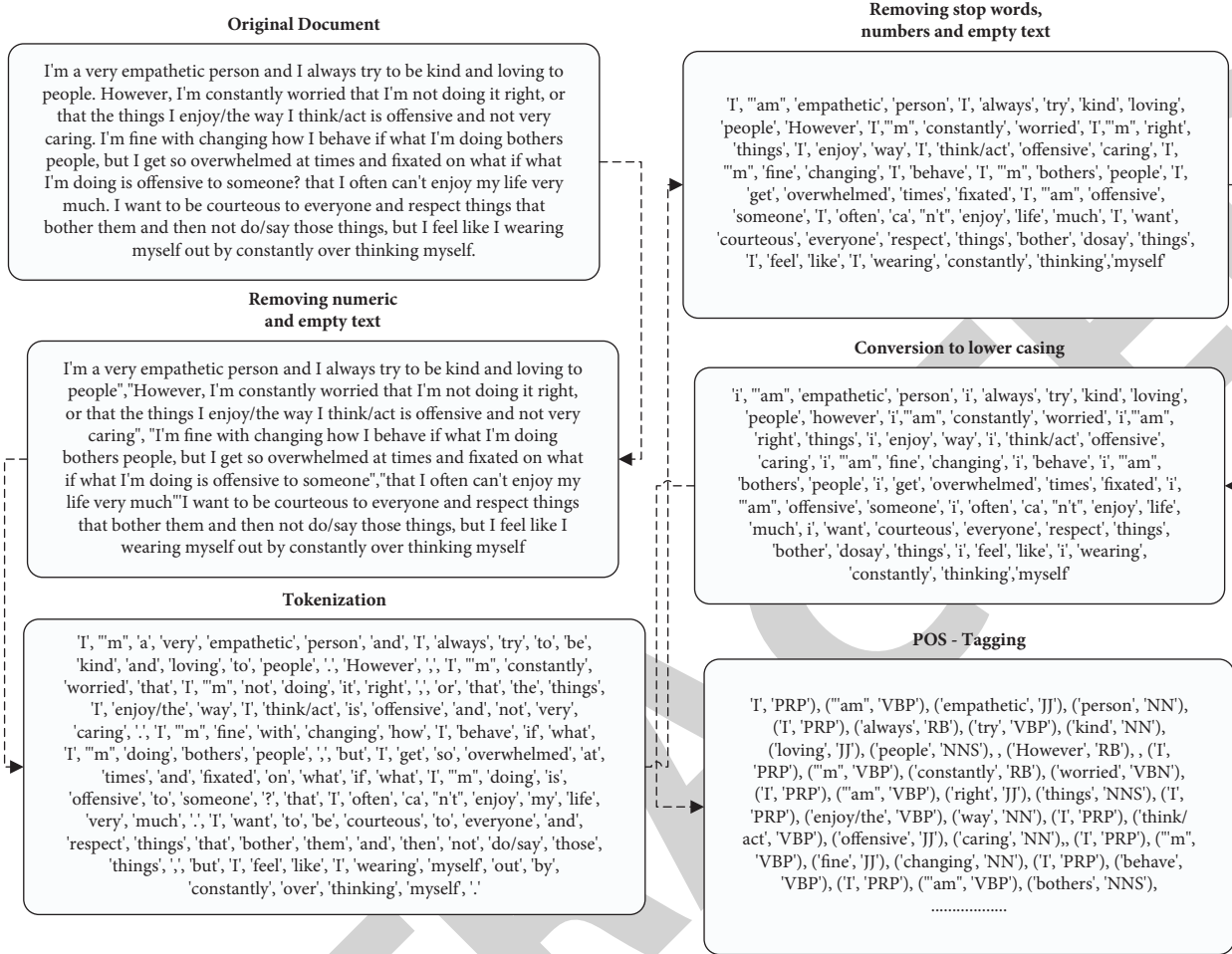


FIGURE 4: The basic natural language processing pipelines used while preparing data. We used these techniques to convert the collected text into small chunks so that they will be consumed by our machine learning model.

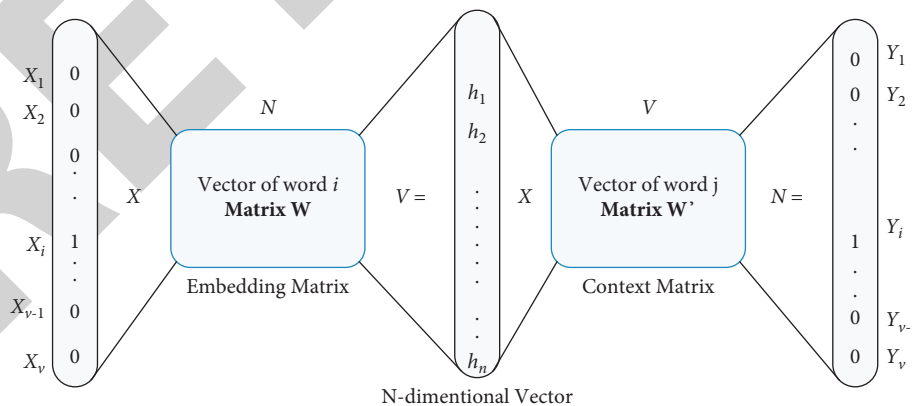


FIGURE 5: Word2vec embedding architecture with the skip-gram model. Both the input vector x and the output y are one-hot encoded word representations. The hidden layer is a word embedding of size N .

3.3.3. *BERT*. BERT is a recent approach proposed by researchers using the Google AI language. BERT uses the transformer architecture, an attention mechanism that discovers contextual contacts between words [45]. The

transformer contains two separate components: an encoder that reads the text input and a decoder that generates a prediction for the task. Because BERT’s main target is to produce a language model, only the encoder mechanism is

necessary. During training, BERT utilizes two schemes to capture the contextual meaning of words in both the right and left directions.

(1) *Masked Language Model (MLM)*. Before passing the input sequences into the BERT model, 15% of the words in each sequence were substituted with a [MASK] token. BERT then tries to infer the initial value of the masked words based on the context provided by the other words in the sequence. A prediction is conducted based on three main steps, as depicted in Figure 6:

- (i) Applying a classification layer on top of the encoder output
- (ii) Multiplying the output vectors by the embedding matrix and transforming them into the vocabulary dimensions
- (iii) Computing probability distribution for all the words in the vocabulary using softmax

(2) *Next Sentence Prediction (NSP)*. During the training process, the BERT model takes two sentences and finds some correlation to predict if the second sentence is the follow-up (next sentence) of the first sentence in the original document. During the training phase, half of the input is a pairing in which the second sentence is the consequent sentence in the original document, and the other half of the input is a second sentence randomly selected from the text corpus.

BERT uses an encoder from transformer architecture that is an encoder-decoder network based on self-attention on the encoder side and attention on the decoder side [45]. BERT is of two sizes: a base BERT (BERTBASE) and a large BERT (BERTLARGE). BERTBASE has 12 layers in the encoder stack, whereas BERTLARGE has 24 layers in the encoder stack. The BERT architectures (BASE and LARGE) also have larger feedforward networks (768 and 1,024 hidden units, respectively), and more attention heads (12 and 16, respectively) than the transformer architecture suggested in the original study [45]. It contains 512 hidden units and 8 attention heads. BERTBASE contains 110M parameters, whereas BERTLARGE has 340M parameters. This model takes the classification token (CLS) as the first input, followed by a sequence of words that are being forwarded to the next layer. Each layer applies a self-attention and passes the result to the next encoder through a feedforward neural network. We propose BERT-based word embedding and fine-tuning to our specific health-related problem task based on the hyperparameters described in Table 1.

3.4. Classifier Models. We implemented many recent state-of-the-art classifier models to create the last sentiment classification on our data set. In our experiment, we implemented an SVM, a logistic regression, a random forest, and AdaBoost. With deep learning techniques, words are represented as a dense vector, and a machine-learning-/deep-learning-based sentiment classification will be applied by passing dense vectors to the classification models. Recurrent neural networks (RNNs), which are the most

common sequence modeling techniques, are capable of obtaining appropriate features from data and can be a leading choice for capturing the semantics of long texts. However, an RNN is somehow an inclined model because it allocates a high priority to recently appearing words in a sequence, which might lessen its efficiency when capturing the semantics of an entire document. As a result, LSTM was instituted to surmount the flaws in maintaining long-term dependencies in RNN models [36]. The LSTM model along with a convolutional neural network (CNN) for sentence classification provides accurate results and has been recently utilized in a variety of NLP tasks. CNN models use convolutional and maximum pooling layers to extract the most relevant features, whereas LSTM models maintain the relationship between words for a longer period using memory cells. Hence, they are better used for text classification [36, 47]. A bidirectional LSTM consisting of two such memories was developed to tackle sequence classification problems. As an example, it was applied in traffic event analysis along with OLDA using social networking data and achieved an extremely high accuracy [48]. Based on the results, We propose a bidirectional LSTM (Bi-LSTM), which uses two LSTM units, that works in both left and right directions to combine past and future context information from our collected social media data. Bi-LSTM also retains a long-term relationship between words along with duplicate context information, as depicted in Figure 7 [49].

In each LSTM unit shown in Figure 7, there are hidden layers with the capability of keeping the previous information for a reasonably longer period. The LSTM architecture contains the main component called the memory cell $\{c_t\}$ that is being updated by using the input gate $\{i_t\}$ and forget gate $\{f_t\}$, as shown in (4). The input gate $\{i_t\}$ is responsible for deciding which information should be stored in the memory cell. The forget gate $\{f_t\}$ is used to decide which information needs to be discarded from the memory cell. At each time step, $\{c_t\}$ for the forward LSTM can be updated using (4) as follows:

$$\begin{aligned} f_t &= \sigma(\omega_f \cdot (h_{t-1}, x_t) + b_f), \\ i_t &= \sigma(\omega_i \cdot (h_{t-1}, x_t) + b_i), \end{aligned} \quad (3)$$

$$C_{t-2} = \text{tanh}(w_c \cdot h_{t-1}, x_t) + b_c.$$

$$C_t = f_t * c_{t-1} + i_t * C_{t-2}. \quad (4)$$

$$\begin{aligned} o_t &= \sigma(\omega_o \cdot (h_{t-1}, x_t) + b_o), \\ h_t &= o_t * \tanh(C_t). \end{aligned} \quad (5)$$

In our proposed framework, we used a principal component analysis (PCA) before feeding the embedding vectors to the classifiers. PCA is a technique with the primary purpose of reducing the dimension of a data set consisting of many related variables, either heavily or lightly, while preserving the variation available in the data set. As it can be seen in Figure 8, we conducted intensive experimentation to exploit and compare the key advantages of each word embedding and classification technique, which have recently

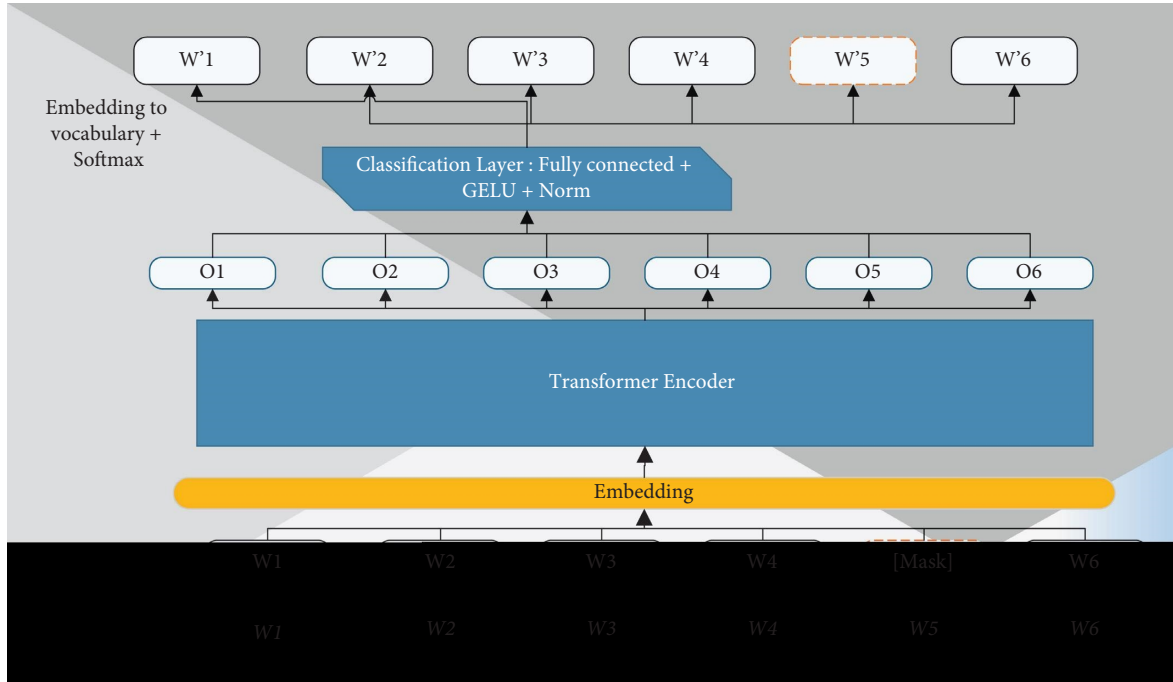


FIGURE 6: Masked language modeling used to train BERT, where a model uses the context words surrounding a mask token to try to predict what the masked word should be.

TABLE 1: Hyperparameters used in the deep learning techniques in our experimental environment.

Hyperparameters	Deep learning model				
	CNN	LSTM	Bi-LSTM	BERT fine-tuning	Proposed KD-based classification
Batch size	64	64	64	4	4
Epochs	20	20	20	3	3
Optimizer	Adam	Adam	Adam	Adam	Adam
Number of hidden layers	128	128	128	768	768

become extremely popular. Our research was conducted on Reddit and Twitter data, which are classifications of depression/anxiety-related posts (positive/negative) using deep-learning-based text analysis techniques.

We propose a two-way implementation approach, as it is depicted in Figure 8, the first of which is the implementation of state-of-the-art machine/deep learning techniques for depression/anxiety identification along with dimensionality reduction to maximize the accuracy. The second most important part is applying fine-tuning and model optimization (knowledge distillation) to build a lighter and smarter depression/anxiety detection model. Knowledge distillation here is building a compressed model by teaching it exactly what to do in a sequential manner using a larger pretrained BERT model [19]. Our proposed knowledge distillation-based sentiment analysis model is further described in Figure 9, which shows how pretrained weights from BERT are used are fine-tuned to our depression/anxiety detection downstream task to build a smarter and lighter model.

The teacher-student architecture used throughout this research is a universal structure for knowledge transfer [41]. In other words, the quality of knowledge gain and extraction

from teacher to student is also dependent on how the teacher-student network is designed. The guidance message from the larger model, which is known as “knowledge,” supports the smaller model to simulate the behavior of the larger complex model. In the sentiment analysis task, logits (e.g., the output of the last layer in a deep neural network) are used as a means to propagate knowledge from the teacher model, which is not clearly given by the training data sets. Given a vector of logits z as the output of the last fully connected layer of a deep model, such that $\{z_i\}$ is the logit for the i -th class, the probability $\{p_i\}$ that the input belongs to the i -th class can then be estimated using the softmax function as follows:

$$p_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)}. \quad (6)$$

Therefore, the predictions of the soft targets obtained by the teacher model include dark knowledge and can be utilized as a supervisor to migrate knowledge from the complex model to a simple one. In our proposed framework, we built a smaller student model from a pretrained

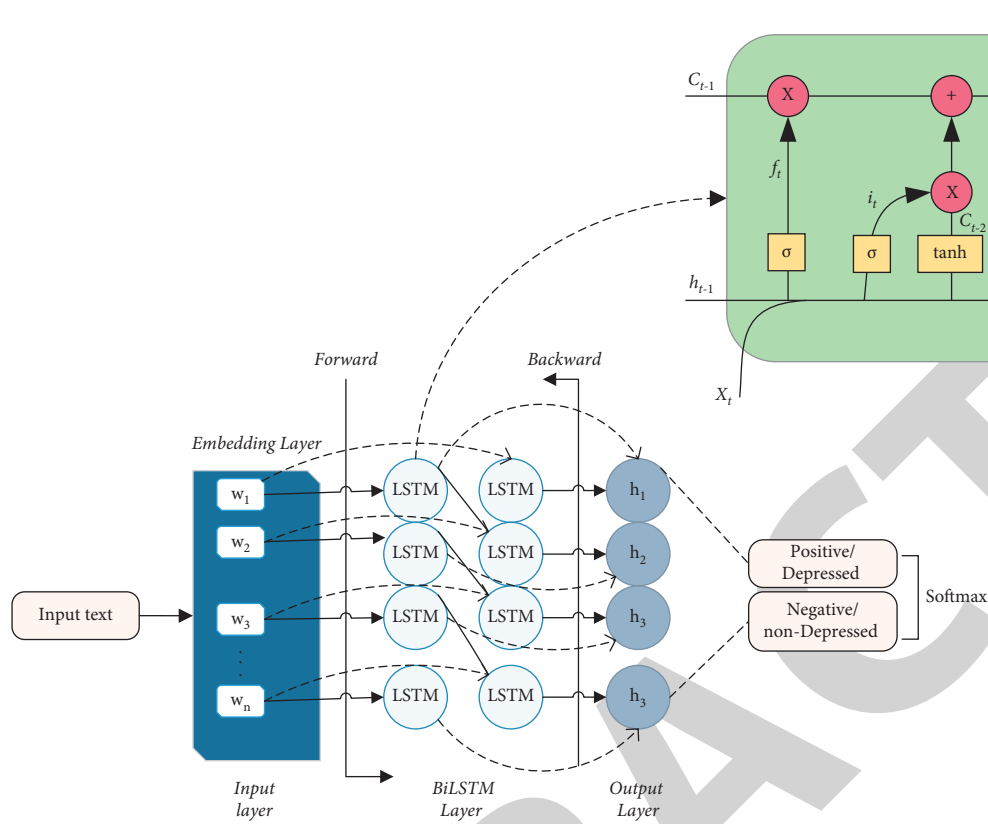


FIGURE 7: Bi-LSTM process used to capture sequential features.

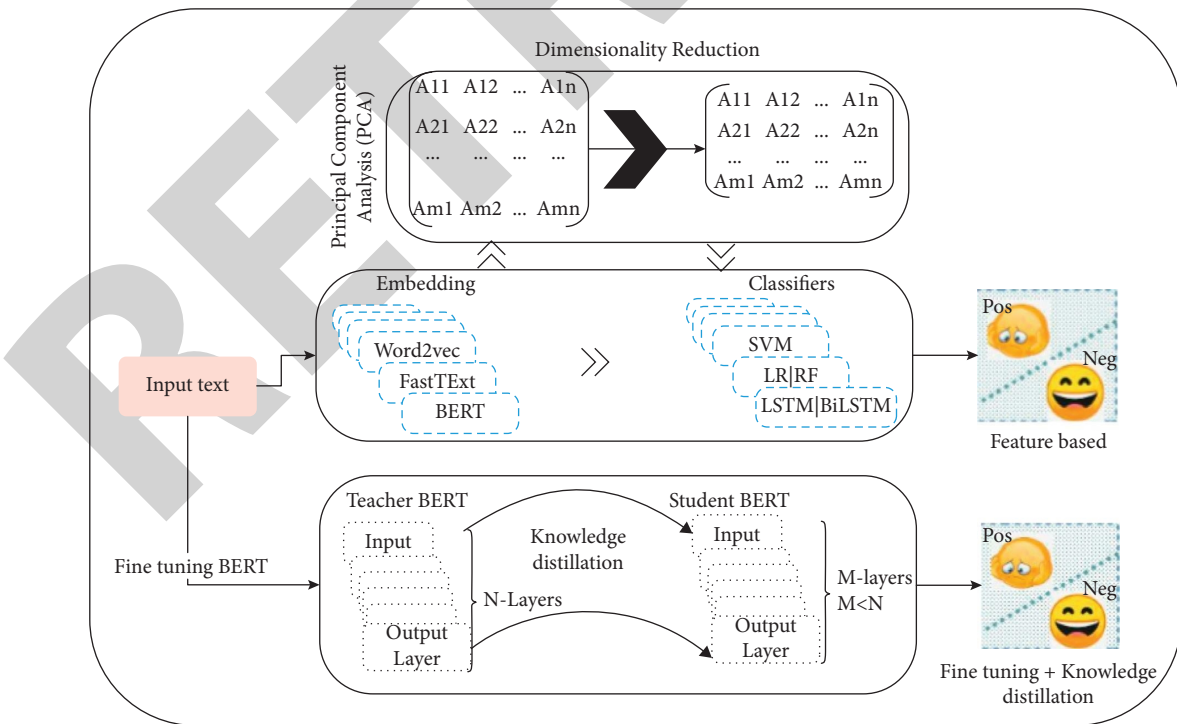


FIGURE 8: Our proposed general framework for text representation, dimensionality reduction, and depression/anxiety detection based on fine-tuning and knowledge distillation techniques along with the common feature-based classification.

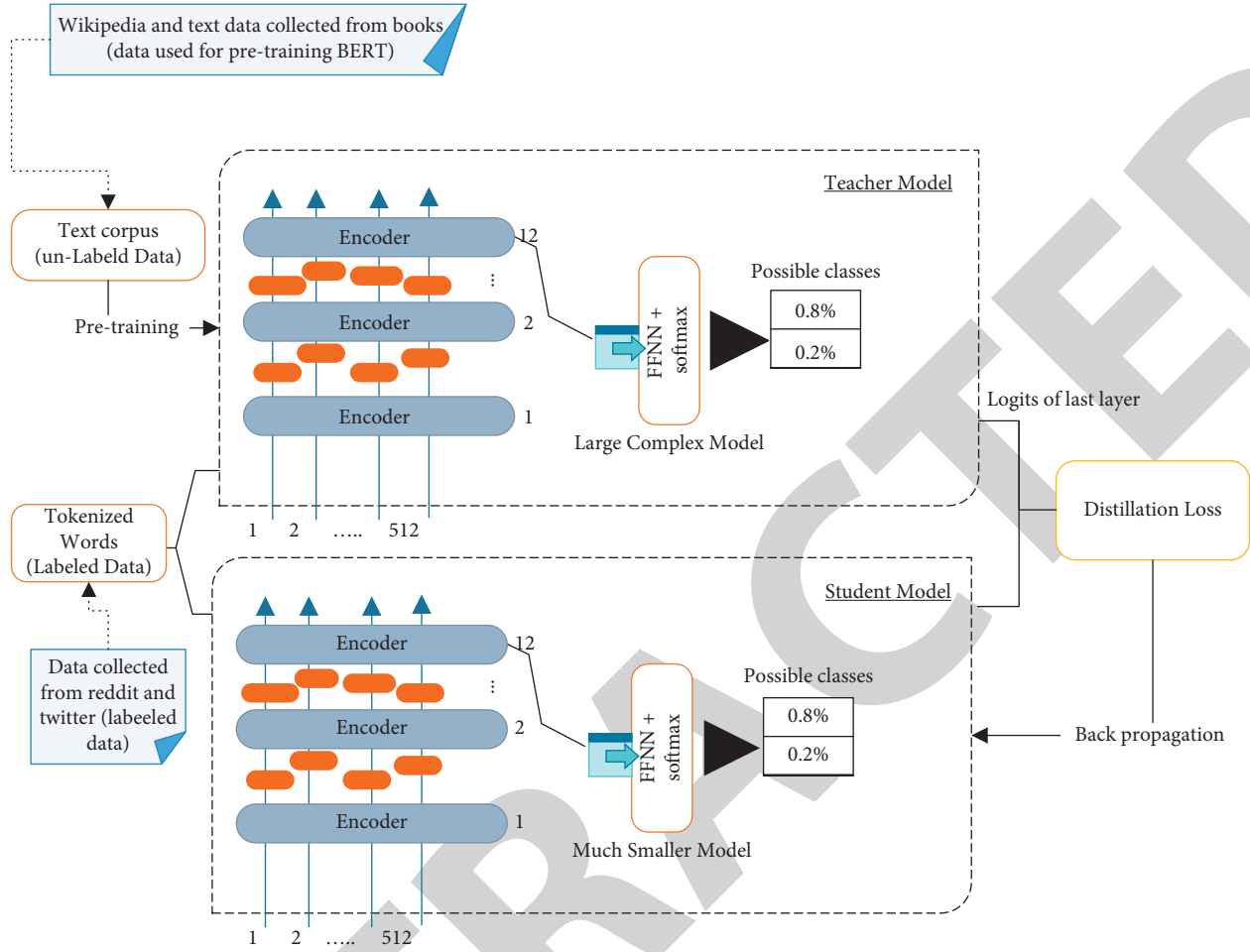


FIGURE 9: Detailed architecture of applying knowledge distillation from pretrained BERT to build a smaller task-specific model on health-related problem identification.

(unsupervised) large BERT model using our collected data set (labeled data). Our proposed framework has the following key steps:

- (i) Training the teacher model: the complex teacher network (BERT) is initially trained on a large generic data set (Wikipedia and BookCorpus). This phase demands high-performance computational environment such as high-performance GPUs.
- (ii) Establish correspondence: in constructing a student network (distilled BERT), there should be a connection between the immediate outputs the student and teacher models. This relationship can be done either by directly forwarding the output of a layer in the teacher network to the student network or doing data augmentation before passing it to the student network.
- (iii) Forward pass through the larger model: the data are forwarded through the teacher network to get all the immediate results and then apply data augmentation.
- (iv) Applying backpropagation: the result from the teacher network and the correlation is used to backpropagate the error in the student network.

We used response-based knowledge, which usually refers to the result of the last output layer of the teacher model (last logits). Response-based knowledge distillation is known to be simple and effective for compressing large models and was commonly used in many NLP tasks recently. As described in Figure 9, the distillation loss of response-based knowledge can be formulated as follows:

$$L_D(p(z_t), p(z_s)) = L_{KL}(p(z_s), p(z_t)), \quad (7)$$

where LKL is the Kullback–Leibler divergence loss. The Kullback–Leibler divergence score, or KL divergence score, quantifies the extent to which one probability distribution differs from another probability distribution. In our case, the KL divergence between two distributions Q (teacher model) and P (student model) is often stated using the following notation: $KL(P - Q)$, where the $-$ operator indicates a “divergence,” or the divergence of P from Q . KL divergence can be calculated as the negative sum of probability of each

event in P multiplied by the log of the probability of the event in Q over the probability of the event in P as follows:

$$KL(P \parallel Q) = -\sum_{x \in X} P(x) * \log\left(\frac{Q(x)}{P(x)}\right). \quad (8)$$

Based on this loss calculation, backpropagation takes place such that the student network can learn to replicate the behavior of the teacher network.

4. Experimental Results

The results of our proposed scheme are discussed in this section. The next remaining subsection describes the data set utilized to train and test the classification models. In Subsection 4.2, we introduce the methods used to effectively measure the performance of our classifier model, and the results will be explained in subsection 4.3. Finally, Subsection 4 describes the various experimental results achieved through our study.

4.1. Data Sets. We designed a data collection framework using Twitter and Reddit APIs (Tweepy and PRAW, respectively) to collect data on depression and anxiety, studied different text-based emotion detection techniques, and used the circumplex model of affect to use keywords belonging to different emotions, as shown in Figure 1 [50]. A common way to define a set of emotions can be using a list of words expressing each emotional trait. We used a keyword-based approach that is based on predetermining a set of terms to classify the text into emotion categories such as happy, angry, and sad, which were used to gather the most significant data on depression and anxiety from social media platforms. In our data set, we collected 100,000 tweets and 95,000 posts from Twitter and Reddit, respectively. One common problem in data preparation is using an imbalanced data set where data belonging to one class are significantly higher or lower in number than those belonging to other classes. As a result, the learning algorithm might choose to ignore those underrepresented classes. Therefore, we tried to balance the number of positive and negative classes (52% and 48%, respectively). We collected posts and comments from Reddit using two commonly applied subreddits called *r/Depression* and *r/Anxiety* and labeled them as positive, whereas the standard posts filtered by keywords based on the circumplex model shown in Figure 1 [30] are labeled as negative.

Table 2 shows a summary of the data sets collected based on our proposed framework. We merged them together and prepared training and test data sets. We used word cloud to visualize frequent words in a text where the size of the words represents their frequency in the positive (depression-related posts) and negative (standard posts) directions collected using keywords related to happiness and excitement, as depicted in Figure 10.

4.2. Performance Matrices. To evaluate our model, we used the commonly used metrics, such as precision, recall, and

accuracy. A confusion matrix is a matrix used for evaluating the classification performance, which is also called an error matrix because it shows the number of incorrect predictions versus the number of correct predictions in a tabulated manner. Based on the confusion matrix, we can compute the accuracy, precision, and recall as follows:

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP}, \\ \text{recall} &= \frac{TP}{TP + FN}, \\ \text{accuracy} &= \frac{TP + TN}{TP + N + FP + FN}. \end{aligned} \quad (9)$$

The most commonly used terminologies that are used in calculating the confusion matrix are as follows:

- (i) P: an actual positive case, which is the depression-/anxiety-related class in our model
- (ii) N: an actual negative case, which is nondepression-/nonanxiety-related class in our model
- (iii) True positive (TP): a case in which the actual class of the data point (collected text data) is true (1), and the class predicted by our model is also true (1)
- (iv) True negative (TN): a case in which the actual class of the data point is false, and the predicted class is also false
- (v) False positive (FP): a case in which the actual class of the data point is 0 (false), and the predicted class is 1 (true)
- (vi) False negative (FN): a case in which the actual class of the data point is true, and the predicted class is false

4.3. Results. In the experiment we have conducted, we studied our model attributes along with numerical performance knowledge distillation-based model and state-of-the-art machine learning algorithms. We used two data sets, namely, depression- and anxiety-related text posts collected from social media specifically from Reddit and Twitter users.

4.3.1. Our Proposed Approach versus State-of-the-Art Algorithms. In our experiment, the proposed Bi-LSTM and BERT-based knowledge distillation scheme were compared with RF, LG, SVM, CNN, AdaBoost, MLP, and LSTM algorithms for identifying depression-/anxiety-related sentiments using our own data as shown in Tables 3 and 4. We used an RF with 150 iterations, 100 estimators, and SVM with a training parameter ridge estimator and a radial basis function (kernel = rbf), respectively. In baseline algorithms, we use three famous word embedding techniques, such as TF-IDF, word2vec, and fastText that are applied on SVM, NB, RF, and AdaBoost models as shown in Table 3. Our main target here was to examine which feature extraction techniques best favor the performance for depression and anxiety detection based on our collected data. We applied

TABLE 5: Deep-learning-based text representation and classifier models applied to our generated data set including our proposed scheme.

Embedding methods	Classification models						
	MLP Accuracy	SVM Accuracy	LSTM Accuracy	Bi-LSTM Accuracy	1-D CNN Accuracy	BERT_fine-tuning Accuracy	Proposed KD distilled_BERT Accuracy
Data sets used-depression-related data (depressed vs. normal)							
GloVe	82%	78%	87%	88%	84.50%	—	—
BERT	90.32	90.35	89%	91.34	87%	90%	97%
Data sets used-anxiety-related data (anxiety related vs. normal)							
Word2vec	90%	88%	94.20%	96%	94.30%	—	—
BERT	91%	85%	95.80%	96%	94.30%	96%	98%

Bayes (NB), and logistic regression (LR) when using the word2vec vector. It is because word2vec is unable to handle new words, and it influences these models to be inefficient in terms of accuracy.

Table 5 presents the results obtained from the proposed Bi-LSTM, knowledge distillation, and five other classification algorithms. These classification models were used to predict mental-health-related problems using three types of text vectorization techniques such as word2vec, GloVe, and BERT, and the results of all baseline approaches were compared to assess the performance of the word vectorization models. Bi-LSTM and distilled BERT achieved higher classification accuracies of 96% and 98%, respectively. The main reason is that the student (distilled BERT) model mimics the teacher model that initially was trained on general text corpus such as Wikipedia and BookCorpus. Therefore, distilled BERT obtains a competitive or even a superior performance when fine-tuned to our depression- and anxiety-related data domain. The learning of this small model from the bigger pretrained model in our proposed framework is termed knowledge distillation. In addition, BERT outperforms word2vec and GloVe for both anxiety and depression prediction tasks because BERT is able to distinguish and capture two different semantic meanings by producing two different vectors for the same word in a given text corpus. We can see that Bi-LSTM with BERT can precisely infer depression and anxiety from a given large text corpus as shown in Table 5.

Figure 11 presents the comparison of accuracy and loss with respect to the training and testing phase of our proposed Bi-LSTM. We compared the training and test accuracy of the Bi-LSTM model with the BERT-based text representation as shown in Figure 11. We get training accuracy after applying the model on the training data, while test accuracy is the accuracy on the test data. We have run our experiment for 20 epochs both for model accuracy (a) and model loss (b) in Figure 11. We can see that they follow the same trend under different parameter settings (batch_size = 64, epochs = 20, verbose = 1, and validation_split = 0.2). This small difference between our training and test accuracy shows that we have a proper setting of the regularization of all parameters of the network and a well representative data batch. Finally, Table 1 presents the hyperparameters used in our repeated experiment. We have set optimal

hyperparameters and optimizer algorithms used to train our models as shown in Table 1. The parameters used were batch size, epochs, optimizer, and number of hidden layers and Adam as an optimizer performed well in our experiment. We applied hyperparameter tuning by conducting a repeated experiment and applied these parameters to control the learning process. We have achieved the best accuracy under these parameter settings listed in Table 1.

4.3.2. Comparison with the Existing System. We compared our proposed system with the state-of-the-art systems, which were designed to analyze and predict mental-health-related problems from social media data, as presented in Table 6. Kim et al. utilized the word2vec model with XGBoost and a CNN for depression- and anxiety-related text classification and achieved an accuracy of 75.13% on depression-related data and 77.81% on anxiety-related data [28]. Mickael et al. presented depression detection techniques using combined features (LIWC + LDA + bigram), classifying them using SVM and MLP, and obtained accuracies of 90% and 91%, respectively [37]. Hatoun et al. applied linguistic inquiry and word count (LIWC) with a linear SVM to predict depression from Twitter data and obtained an accuracy of 82.50% [43]. Zogan et al. employed multimodalities + word embedding (word2vec) with a bidirectional gated recurrent unit-convolutional neural network (BiGRU-CNN) for depression detection from Twitter data and achieved an accuracy of 85% [26]. In 2019, Kumar et al. utilized a feature matrix with ensemble vote classification using RF, NB, and gradient boosting and obtained an accuracy of 85.09% [27]. The accuracy of our system is shown in the last row of Table 6. During this experiment, we adopted the concept of knowledge distillation using BERT and fine-tuned our task of depression and anxiety detection, obtaining an accuracy of 98% for anxiety-related data and 97% for depression-related data. In addition, we employed an attention mechanism to hold the long-term relationship among words using transformer architecture and applied Bi-LSTM to predict anxiety, obtaining an accuracy of 96%. The results obtained show that the proposed techniques might help in the development of smart and efficient systems for the detection of depression, anxiety, and other health-related problems from social media textual contents created by users.

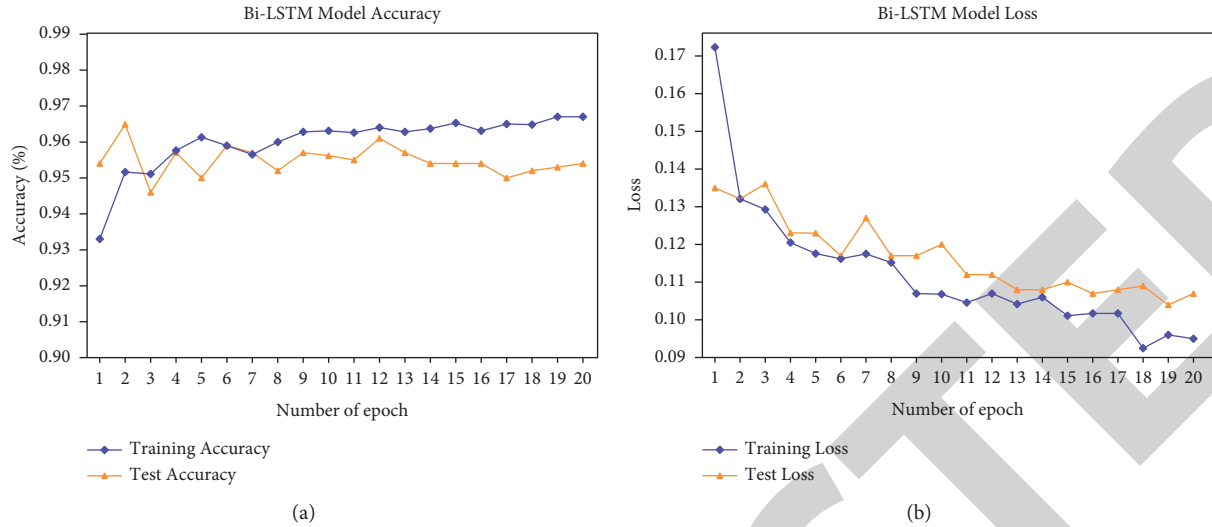


FIGURE 11: Comparison of training and test accuracy (a) versus training and test loss (b) for proposed Bi-LSTM model. The parameters we have used for the experiment were `batch_size = 64`, `epochs = 20`, `verbose = 1`, and `validation_split = 0.2`.

TABLE 6: Detailed comparison of existing methods for the detection of depression-/anxiety-related data on social media.

Author (year)	Data source	Word embedding model	Classifiers	Overall accuracy
Jina Kim (2020)	Reddit	Word2vec (form gensim)	XGBoost and CNN	75.13% (depression) 77.81% (anxiety)
Michael (2019)	Reddit	LIWC + LDA + bigram (combined feature)	SVM and CNN	90% and 91%, respectively
Hatoon (2020)	Twitter	LIWC Sentiment analysis	Linear SVM	82.50%
Tadesse (2020a)	Reddit	Word2vec	LSTM-CNN	93.80%
Akshi Kumar (2019)	Twitter	Feature matrix	Ensemble vote, classification (RF, NB, and B)	85.09%
Hamad (2020)	Twitter	Multimodalities + word embedding (word2vec)	BiGRU-CNN	85%
Our proposed approach	Reddit and Twitter	Word2vec, fastText, and BERT	LSTM, Bi-LSTM, and BERT + knowledge distillation	94%, 96%, and 98%, respectively

5. Conclusion

In this study, we developed a strongly constructed framework for the detection of mental health problems using deep learning techniques such as BERT, Bi-LSTM, and a knowledge distillation based on social media content created by users. The proposed framework enhances the accuracy of smart healthcare systems to detect mental-health-related problems mainly depression and anxiety. This research work can be utilized to build a real-time system for early mental-health-related problem detection mainly based on user posts on Reddit and Twitter. We discussed various key features, including mental-health-related textual data collection from social networks using application programming interfaces and preprocessing module that focuses on the conversion of unstructured data into a meaningful form using various data filtering techniques. We also have employed a keyword and circumplex model-based text labeling technique on the collected text corpus to extract useful features related to mental health. Furthermore, our mental health problem

detection framework applies the most recent text embedding technique based on deep learning that ensures capturing the semantic and contextual meaning of words included in user posts.

The proposed BERT-based text representation model transforms collected words into vectors capturing the semantic meaning in the collected text corpus to improve the accuracy of the classification task using an attention mechanism. We also proposed response-based knowledge distillation, which is based on the neural response of the last output layer of the teacher model (last logits) in BERT for building a smaller and smarter model for depression/anxiety detection and classification. Moreover, we developed our own data collection module to prepare the data set from Twitter and Reddit by mining the most relevant textual data that can be used to build an intelligent model for smart healthcare systems. Finally, we conducted an intensive experiment, based upon which the proposed BERT-Bi-LSTM model improves the accuracy of text sentiment classification from user posts. The main reason the model outperforms

other machine learning classification models is that it combines the strengths of both BERT and Bi-LSTM models to comprehend the syntactic and contextual information of each word. In addition, we implemented a response-based knowledge transfer using BERT, fine-tuned the task of depression/anxiety detection, and achieved extremely high accuracy. In future work, a multimodel depression detection system can be developed to utilize more diverse data such as text, image, and behavioral features to achieve effective results.

Data Availability

The data sets used for mental health prediction are available from the corresponding author upon request.

Disclosure

Kamil Zeberga and Muhammad Attique are considered the co-first authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Kamil Zeberga and Muhammad Attique contributed equally to this work.

Acknowledgments

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-0-02051) supervised by the IITP (Institute for Information and Communications Technology Planning and Evaluation) and the BK21 FOUR program of the National Research Foundation of Korea funded by the Ministry of Education (NRF5199991014091). This research was also supported by the National Research Foundation of Korea funded by the Ministry of Education (2020R1G1A1013221) and supported by the Research Incentive Fund R20129 of Zayed University, UAE.

References

- [1] L. Neuhauser and G. Kreps, "Rethinking communication in the e-health era," *Journal of Health Psychology*, vol. 8, no. 7–23, 01 2003.
- [2] K. Shrestha, "Machine learning for depression diagnosis using twitter data," *International Journal of Computer Engineering in Research Trends*, vol. 5, no. 2, 2018.
- [3] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Detection of suicide ideation in social media forums using deep learning," *Algorithms*, vol. 13, 2020.
- [4] J. Seppälä, I. Vita, T. Jämsä et al., "Smartphone and wearable sensors-based m-health approach for psychiatric disorders and symptoms – a systematic review and link to m-resist project (preprint)," *JMIR Mental Health*, vol. 6, 2018.
- [5] L. S. Radloff, "The CES-D scale," *Applied Psychological Measurement*, vol. 1, no. 3, pp. 385–401, 1977.
- [6] N. Drissi, S. Ouhbi, M. Idrissi, M. Koutbi, and M. Ghogho, "On the use of sensors in mental healthcare," vol. 26, pp. 307–316, 2019.
- [7] A. Gaggioli and G. Riva, "From mobile mental health to mobile wellbeing: opportunities and challenges," *Studies in Health Technology and Informatics*, vol. 184, pp. 141–147, 2013.
- [8] N. C. Dang, M. N. Moreno-García, and F. de la Prieta, "Sentiment analysis based on deep learning: a comparative study," 2020, <https://arxiv.org/abs/2006.03541>.
- [9] A. Fiallos and K. Jimenes, "Using reddit data for multi-label text classification of twitter users interests," in *Proceedings of the 2019 6th International Conference on eDemocracy and eGovernment, ICEDEG 2019*, pp. 324–327, IEEE, Quito, Ecuador, 24 April 2019.
- [10] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," *Processing*, pp. 1–6, 2009.
- [11] A M. Q, "Project. Machine learning for mental health detection," *Psychol.Col*.vol. 49, 2019.
- [12] A Sau and I Bhakta, "Informatics in Medicine Unlocked Screening of anxiety and depression among seafarers using machine learning technology," *Informatics in Medicine Unlocked*, vol. 16, no. August, Article ID 100228, 2019.
- [13] S Neha, P H C. Shekar, K S. Kumar, and A Vg, "Emotion recognition and depression detection using deep learning," pp. 3031–3036, 2020.
- [14] F. Ali, S. El-sappagh, S. M. R. Islam et al., "An intelligent healthcare monitoring framework using wearable sensors and social networking data," *Future Generation Computer Systems*, vol. 114, pp. 23–43, 2021.
- [15] M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," in *Proceedings of the International Conference on Intelligent Sustainable Systems, ICISS 2017*, no. Iciss, pp. 858–862, 2018.
- [16] R. Wald, T. Khoshgoftaar, and C. Sumner, "Machine prediction of personality from Facebook profiles," vol. 2, pp. 109–115, in *Proceedings of the 2012 IEEE 13th International Conference on Information Reuse and Integration, IRI 2012*, vol. 2, pp. 109–115, IEEE, Las Vegas, NV, USA, 8 August 2012.
- [17] A. Yadav and D. K. Vishwakarma, "Sentiment analysis using deep learning architectures: a review," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.
- [18] S. Biswas, E. Chadda, and F. Ahmad, "Sentiment analysis with gated recurrent units," *Advances in Computer Science and Information Technology (ACSIT)*, vol. 2, no. 11, pp. 59–63, 2015.
- [19] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1, no. M1m, pp. 4171–4186, North American, 2 June 2019.
- [20] A. Radfort, K. Narasimhan, T. Salimans, and I. Sutskever.
- [21] K. Kowsari, K. J. Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: a survey," *Information*, vol. 10, no. 4, pp. 1–68, 2019.
- [22] E. Garcia-Ceja, M. Riegler, T. Nordgreen, P. Jakobsen, K. J. Oedegaard, and J. Tørresen, "Mental health monitoring with multimodal sensing and machine learning: a survey," *Pervasive and Mobile Computing*, vol. 51, pp. 1–26, 2018.

Retraction

Retracted: Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers

Computational Intelligence and Neuroscience

Received 10 October 2023; Accepted 10 October 2023; Published 11 October 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] A. Goswami, M. M. Krishna, J. Vankara et al., "Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9194031, 18 pages, 2022.

Research Article

Sentiment Analysis of Statements on Social Media and Electronic Media Using Machine and Deep Learning Classifiers

Anjali Goswami,¹ Muddada Murali Krishna,² Jayavani Vankara,² Syam Machinathu Parambil Gangadharan ,³ Chandra Shekhar Yadav,⁴ Manoj Kumar ,⁵ and Mohammad Monirujjaman Khan ⁶

¹Department of Mathematics and Statistics, College of Science and Theoretical Studies, Saudi Electronic University, Riyadh-13323, Saudi Arabia

²Department of Computer Science & Engineering, Dr Lankapalli Bullayya College of Engineering, Visakhapatnam, Andhra Pradesh, India

³Big Data Engineer, General Mills, Minneapolis, USA

⁴Ministry of Electronics and Information Technology, Delhi, India

⁵School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand-248007, India

⁶Electrical and Computer Engineering, North South University, Bashundhara, Dhaka-1229, Bangladesh

Correspondence should be addressed to Manoj Kumar; wss.manojkumar@gmail.com and Mohammad Monirujjaman Khan; monirujjaman.khan@northsouth.edu

Received 23 November 2021; Revised 30 December 2021; Accepted 17 January 2022; Published 2 March 2022

Academic Editor: Anastasios D. Doulamis

Copyright © 2022 Anjali Goswami 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.

When it comes to our everyday life, emotions have a critical role to play. It goes without saying that it is critical in the context of mobile-computer interaction. In social and mobile communication, it is vital to understand the influence of emotions on the way people interact with one another and with the material they access. This study tried to investigate the relationship between the expressive state of mind and the efficacy of the human-mobile interaction while accessing a variety of different sorts of material over the course of learning. In addition, the difficulty of the feeling of many individuals is taken into account in this research. Human hardness is an important factor in determining a person's personality characteristics, and the material that they can access will alter depending on how they engage with a mobile device. It analyzes the link between the human-mobile interaction and the person's mental toughness to provide excellent suggestion material in the appropriate manner. In this study, an explicit feedback selection method is used to gather information on the emotional state of the mind of the participants. It has also been shown that the emotional state of a person's mind influences the human-mobile connection, with persons with varying levels of hardness accessing a variety of various sorts of material. It is hoped that this research will assist content producers in identifying engaging material that will encourage mobile users to promote good content by studying their personality features.

1. Introduction

Experiencing emotions is a necessary component of our everyday lives. They are there in everything we do, wherever we are, and everywhere we travel, and they do so without our even realizing they are there. Because we can see the other person's emotional expression on their face and because we do not receive an emotional reaction when we anticipate

one, when it comes to mobile engagement, whether it is with people, technology, or humans through technology, all of us are suddenly more mindful of emotion. It is tough to distinguish between different emotions.

Emotions are complicated processes that are created by interactions between intentions, beliefs, perceptions, and circumstances, among other factors. To comprehend sentiments, we must, firstly, comprehend the mental state and

situations of the person, as well as the sorts of processes that may be elicited by these conditions. The environment has an important part in determining the condition of the human mind. The human-computer/human-mobile interaction investigates the interaction between a person and a machine. It depends on supporting information from both the device and the human side [1]. A large number of scholars from throughout the globe have dedicated their careers to understanding the function of emotion in our interactions with technology.

In the field of affective computing [2–4], the research focuses on four main areas: (1) reducing user frustration, (2) enabling comfortable communication of user emotion, (3) developing infrastructure and applications to handle adequate information, and (4) developing tools that aid in the development of social and emotional skills.

This research seeks to determine the relationship between human emotional states and interaction, as well as to demonstrate that emotion has a greater impact on interaction than other factors. One of the primary goals of this research is to determine the relationship between human emotional states and interactions. As a result, it will aid the researcher in comprehending and inventing new techniques and technologies to realize the goal of affective computing as mentioned above. Our research team performed trials with real individuals who were interacting with different computer systems to evaluate this hypothesis, and they discovered the importance of the emotional state in human-computer interaction. In this experiment, the following are the goals that will be pursued.

The research begins by examining the many sorts of emotional states that people experience as a result of external events, challenges, and circumstances, as well as the impact these emotions have on their own life. Following up on the findings of the preceding research, we assume/fix other emotional scales for several additional events/problems/environments.

Secondly, we have divided events/situations into two categories: those that elicit good feelings and those that elicit negative emotions. The influence of good and negative emotions on human-computer interaction is also explored in more depth in this research.

The third part of our research looked at the influence of emotional mindset on human-computer interaction across a variety of interaction activities.

On a fourth level, we looked at how people react to certain emotional or mental states by looking into their own personal “hardness factor” and how that component affects them.

The fifth point is that the research examines the complete influence of stressful/festive events on human-computer interaction by comparing the average interaction impacts between the two groups of participants.

To conclude, to get fantastic content suggestions in understanding their mind-state and personality to modify their emotional level of a person is essential.

It is proposed to use the strategy of combining two (or more) procedures as a way of absorbing the benefits of both, and as a result, filling in some of the weaknesses of individual

methods. Alfrjani and colleagues used a combination of machine learning and a semantic knowledge base to increase the accuracy of sentiment analysis on online reviews (an improvement from 1 percent to 6 percent). In another instance, a hybrid approach for sentiment analysis on tweets that incorporate both lexical and machine learning techniques (improvement: 2 percent to 6 percent). As a result, a hybrid system with collaborative features is better prepared to resolve any issues connected with a single system, if any such pitfalls exist. The efficiency of the integrated models may fluctuate depending on the job being performed. CNNs that have been upgraded by SVM, CNNs that have been enhanced by RNN, and Lexicon-based analysis with machine learning produced improved results. When doing sentiment analysis on diverse domains and kinds of datasets, the combination of CNN, LSTM, and SVM attempts to make use of the two deep network architectural models and SVM algorithms. There are also numerous sorts of input data that may be gathered via social networks, such as tweets and reviews, among other things. Within and across these forms of data, there are variations in the input data, such as the distribution of tweet and review durations, the variety of themes in each dataset, the sample size, and the higher or lower inclusion of explicit emotions and irrelevant information in each dataset. Some techniques may be unable to perform effectively over a range of domains, with poor accuracy and performance in sentiment analysis, for example. Specific procedures may be ill-suited and difficult to apply to certain kinds of input data as a consequence of this phenomenon.

One issue posed by our research is whether hybrid models outperform single models independent of the peculiarities of the datasets. Our findings suggest that they do. As a result, our research investigates the behavior of the chosen hybrid models when applied to various kinds of datasets from various domains. Using three different models, CNN, LSTM, and SVM, we investigated and verified their combination in this study. A link between the models and their increased capabilities to extract characteristics, retain previous information and nodes, and categorize text was studied. Firstly, in the early phases of the model, two potential variants in the sequence of CNN and LSTM are added, one for each of the two possible variations. Finally, two additional possibilities are presented for each of these options: the usage of CNN with ReLU function or SVM, and they are described in detail below. On eight datasets, including tweets and reviews, we tested these models with word embedding and found them to be effective. Using our research, we discovered that combining several models boosted the accuracy of sentiment analysis significantly.

Aside from highlighting four hybrid deep learning models for sentiment analysis that produce improved accuracy regardless of the types of social network datasets used, this paper also provides an experimental study to evaluate the performance of hybrid deep learning models and details a performance comparison of sentiment analysis methods with some state-of-the-art methods, all of which are significant contributions to the literature.

In the wake of our investigation, we have discovered that emotional mentality plays an important role in human-computer/mobile-interface interaction. Apart from that, the toughness factor of individual users might be either good or bad, depending on their experience. This element has an influence on the interaction between humans and computers or mobile devices. Whenever the hardness factor is positive, the emotional mentality has only a little influence on human-computer/mobile interaction activities. At the same time, when the emotional aspect has a stronger impact on the interaction task, the hardness factor has a negative effect. Consequently, according to their hardness level, the user receives varied material suggestions to replicate happy feelings in their minds, which will cause them to be more optimistic and help them to operate more productively in their settings.

The remainder of the paper is organized according to the arrangement outlined subsequently. Section 2 examines the most important techniques in human-emotion-based literature, as well as the most typical works in this area of study. In the third page, we talk about the significance of human emotion and the many emotional mindsets that a typical human being has under different situations. It categorizes both good and negative emotions, as well as events or circumstances that elicit either positive or negative emotions in the participant. Section 4 goes to great lengths on the emotional difficulties that are investigated experimentally. Section 5 delves into the relationship between mobile interaction suggestions and other variables in more depth than the previous sections. A summary of the lessons we have gained is provided in Section 6. The aspects of system development that are influenced by emotion are discussed in Section 7, as is the inclusion of emotional information into system development. Section 8 describes the process of suggestion for transforming negative emotional states into good emotional states. Finally, Section 9 summarizes our findings and our plans for the future.

2. Literature Survey

Consider the significance of employing current technology tools in everyday life and how much effort has been done to provide answers to practical computing throughout the globe. Because of the arrival of Android technology, microcomputer system developers gained a fresh perspective, and a slew of new apps were built. This section examines the many research projects that have been undertaken in this field, as well as the creation of successful computing-based solutions.

Michigan Technological University's Myoungsoon Jeon published a chapter in the book "Foundations of Affective Sciences" in 2017, titled "Emotions and Affect in Human Factors and Human-Computer Interaction: Taxonomy, Theories, Approaches, and Methods," in which he surveyed several approaches of Affective Science in Human Factors/Human-Computer Interaction. Their backgrounds, objectives, and approaches are all distinct. However, all of them clearly demonstrate that emotions and affect are not ancillary, but rather one of the fundamental aspects in Human

Factors/Human-Computer Interaction design, and they all attempt to integrate all of the techniques to achieve effective results. By bridging the gap between conventional Human Factors/Human-Computer Interaction research and the new subject of Affective Sciences research, this handbook helps each discipline advance its own knowledge and understanding.

Researchers from Michigan Technological University published a chapter in the book "Fundamentals of Affective Sciences" in 2017 called "Neural Mechanisms of Emotions and Affect," in which they reviewed many capabilities and components of the brain's functioning in depth. 3. Jacob Aday comes to the conclusion that emotion is not a single process and that the emotional brain is not a specific area of the brain. Emotions are multistage, complicated events, and the brain representation of them is distributed and connected to the stage of emotional processing at which they occur.

According to Hillary A. Efenbein et al., this paper explores the social perception of emotional intelligence and extends the data to assess its reliability and cross-judge agreement, as well as its convergent, divergent, and predictive validity, among other factors. Although several parameters were taken into consideration, the observation research nevertheless showed a substantial correlation with interdependent task performance. Notably, the predictive validity of observer-rated emotional intelligence was higher than that of self-rated or ability-tested emotional intelligence.

The fourth point to mention is that, in the year 2019, Sam Freed's chapter "Prevailing Prejudices About Artificial Intelligence" in the book "AI and Human Thought and Emotion" shows that some outdated ideas, opinions, attitudes, and metaphors have had an impact on cognitive science and Artificial Intelligence (AI), despite the fact that there is insufficient evidence to support their claims. It delineates some of the underlying assumptions that support much of our culture, and in particular, much of artificial intelligence research and development.

"Artificial intelligence for affective computing: an emotion recognition case study," by Pablo Arnau-González and colleagues, provides an introduction to the benefits of artificial intelligence techniques for the field of affective computing through a case study about emotion recognition via brain (electroencephalography-EEG) signals. Pablo Arnau-González and colleagues' chapter "Artificial intelligence for affective computing: an emotion recognition case study" is available online. It was discovered that artificial intelligence-based systems for emotion identification have great potential, as shown by the data collected. There is great potential for this application to make a significant contribution to the fields of human-computer interaction (HCI) and quality of experience (QoE).

In 2008, Brian R. Duffy, a professor at the Massachusetts Institute of Technology, stressed the need of automating emotion to regulate systems at various levels.

In a 2004 paper, Antonella De Angeli and Graham I. Johnson pointed out that computational reductionism has led to an inaccurate perspective of emotions, which thinks

that they are only physical and cognitive, however, they are really deeply anchored in social contexts.

In 2005, Jackeline Spinola de Freitas emphasized the need for emotion design to accurately express emotions. Nevertheless, no comparison was made between emotional and nonemotional systems. Furthermore, it is vital to establish techniques for recording emotional experiences to better understand them.

The multicomponent approach to emotions proposed by Scherer [5] was highlighted in a 2006 study from the Berlin University of Technology, in which different aspects of emotions in an interactive context were investigated: subjective feelings, physiological activation, motor expression, cognitive appraisals, and behavioral tendencies [6–8]. It was discovered that the quality of use differed between two versions of an interactive system that used different emotional states in different ways. Researchers discovered that systems with high usability provide more positive sentiments than systems with usability problems, according to the results of the research. There are a range of ways that may be used to discover differential variations, including rating scales for subjective feelings, electromyography (EMG) of muscle activity in the face, heart rate, ECG (EDA), performance data, and questionnaires for cognitive evaluations. Additionally, it was hypothesized that this combination would provide an extensive framework to comprehend emotional expressions in human-computer interaction (HCI).

In 2018, Jackeline Spinola de Freitas and colleagues aimed to examine the recording of emotions using a number of approaches, including audios, videos, and explicit procedures, to better understand the phenomenon. Each design makes no effort to evaluate or analyze the human harness or the mental state of the subject. Therefore, new approaches for capturing emotional states, as well as fundamental sentiments, are necessary, according to the findings of the study. Our mission is to create a new Android application that will assist scientists in gaining access to emotional mind states, human hardiness, and situation-specific emotions, among other things.

Kövecses analyzes the eleventh in the year 2018. According to the author, conceptual metaphors may be employed to frame media discourses both intratextually and intertextually, and they can be utilized both ways. However, they may also be able to explain what seems to be incoherent media speech, which appears to be the norm in the majority of instances. The idea of conceptual metaphor contributes to our understanding of why people instinctively blend metaphors. As a result of the priming effect, the context in which media discourse is presented has a significant impact on how effective it is. A variety of practical considerations make headlines, especially vulnerable to the contextual impact in their content in particular. Finding noteworthy patterns in media discourse may be accomplished by recording the structure of this cognitive process and then analyzing it. In this essay, the author argues that metaphors should be used to express the meaning of the activities that are being done.

According to Kövecses, in his book titled, “Where Metaphors Come from: Reconsidering Context in

Metaphor,” he describes the desire for happiness as a metaphor and how to apply it in different information processing settings in the chapter headed, “Happiness in Context”.

When it comes to understanding human well-being, the authors of “Emotion in the Digital Age: Technologies, Data, and Psychosocial Life” (Emotion in the Digital Age: Technologies, Data, and Psychosocial Life, published in 2021) stress the need for dynamic modeling and knowledge. Also emphasized is the idea that this constitutes a psycho-social style of being, in which digital technology and emotion function as significant parts of the method in which we simultaneously connect to ourselves as individual subjects and to others as the members of collectives.

The relevance of synthetic cognitive emotions in the design of robot-based systems is highlighted by Jordi Vallverdi (2016) in his article. According to him, it is one of the most significant elements in the design of complex systems.

According to F. Gregory Ashby of the University of California, Santa Barbara’s positive effect hypothesis, which was created in 1999, pleasant emotion enhances performance on a broad variety of cognitive tasks. Several contemporary neuropsychological hypotheses propose that pleasurable feeling is associated with increased dopamine levels in the brain, which, according to the theory, would account for many of these benefits. It predicts or accounts for the impacts of positive emotion on olfaction, long-term (i.e., episodic) memories, working memory, and the capacity to solve creative issues, to name a few aspects.

Dr. Barbara L. Fredrickson’s research in 2018 is focused on positive emotions and psychophysiological laboratory to develop a data-driven model from the bottom up. It describes how individual heterogeneity in immune, parasympathetic, and oxytocin profiles manifests itself in different situations. Furthermore, it serves as a foundation for individual variations and good feelings. She utilizes the word “purpose” as an umbrella phrase to embrace these diverse types of self-transcendence because it is both accessible and alliterative with good feelings and because it is both accessible and alliterative with positive emotions. A major emphasis of the author’s research is on possible biological processes via which pleasant mental states, such as happy emotions and a sense of purpose, might be connected to physical health over time. The broaden-and-build hypothesis serves as a theoretical framework for investigating the biological bases of positive feelings and intentions, as well as their consequences. Then, pleasant emotions and a sense of purpose are dynamically connected, and each seems to play a crucial role in regulating the leukocyte gene expression, which appears to be a reasonable biological substrate of living a fulfilling life.

Denis Rangulov and Muhammad Fahim created a method to detect the emotions shown in video clips in the year 2020. When predicting dimensional emotions from video data, the system model design model blends a convolutional neural network (CNN) with a recurrent neural network (RNN). CNN begins by extracting feature vectors from video frames in the first stage. In the second phase, we

input these feature vectors into an RNN that was being trained to exploit the temporal dynamics of video. In addition, we looked at how each neural network contributed to the overall performance of the system in question.

It is crucial, according to Tianyi Zhang et al., 2021, to try to discern user emotions when viewing short-form videos at any time and from any location. It will allow for video content customization and personalization. However, the majority of the studies either categorize a single emotion per video stimulus or confine it to static, desktop surroundings. It made use of wearable gadgets to track the emotions of participants.

In a study published in 2020, Shihan Ran and colleagues sought to detect emotions using electroencephalography (EEG)-based emotion identification. Their work has improved the area of affective computing and allowed applications in human-computer interactions. The decoding of emotion using supervised machine-learning algorithms has advanced significantly in recent years, while other research have used data-driven, unsupervised methodologies to investigate the underlying EEG patterns during an emotion experiment. It investigates the relationship between such dynamics and subjective assessments of emotion.

Researchers S. Hsu et al. observed in 2018 that a better knowledge of how neurocognitive processes affect hypnotherapy may enhance therapeutic results. When continuous electroencephalographic (EEG) data captured under hypnosis is used to decode cortical state changes, it presents a significant difficulty. We use adaptive mixture independent component analysis (AMICA) to predict changes in the brain state dynamics during hypnosis. AMICA is an unsupervised technique that trains many independent component analysis models for defining nonstationary, unlabeled data. The application of the AMICA model to EEG data from six hypnosis sessions revealed changes in system-wide brain activity that corresponded to transitions between hypnosis levels. The findings also revealed that AMICA-based models were constant across sessions and participants and that they mirrored diverse patterns of source activity in various hypnosis stages as a consequence. Changes in the theta, alpha, and other spectral properties of source activity were identified during the treatment sessions via the analysis of independent component clusters associated with separate classes of model probability patterns.

Banzige (2009) from the Geneva University indicated that emotion detection capacity is a critical component of emotional competence and is a core component of emotional intelligence. Aim: to develop a tool that objectively assesses this ability using actor portrayals of dynamic expressions of 10 emotions (2 variants each for five emotion families), operationalized as recognition accuracy in four presentation modes that combine the visual and auditory sense modalities (audio/video, audio-only, video-only, and still picture). The many experiments and the usefulness of a test design are used to assess coarse and fine-grained emotion discrimination and modality-specific abilities in children. The results of the factor analysis indicate that there are two distinct talents, namely visual recognition and aural

recognition, that seem to be substantially independent of personality dispositions.

The preceding works demonstrate that all of these studies instantaneously capture the emotions using a variety of approaches, including audios, movies, and explicit methods, among others. None of these studies took into account the consumers' attitude or level of difficulty. As a result, this work takes into account the consumers' point of view as well as their genuine circumstances and basic feelings. It will provide more effective outcomes, as seen by this work in progress. It also aids in the recommendation of more interesting information to consumers.

As a field, affective computing exists at the intersection of artificial intelligence and human-computer interaction design. "To understand human emotions and give robots feelings, "emotional modules" will be added to rule-based artificial intelligence systems," the researchers said. The field is seeing unprecedented growth in popularity. A new generation of wearable computers equipped with physiological sensors and pattern recognition is being developed to collect and analyze signals that are intended to communicate the user's emotional state [1, 9].

Scientists are concentrating their efforts on emotion-based computational architectures [8], whose models physiologically inspire the emotion process explored by neuroscience, allowing them to include emotions into machines' functions control or generate an "artificial emotion." According to widespread expectations, adding an emotion model in the computational system would increase machine performance in areas, such as decision-making competence, action selection and control, and autonomous and reliable system response. Human emotion understanding, learning, and measurement collect information about a user's emotional state during an encounter with the user, allowing for the provision of appropriate help or services that are tailored to the user's requirements. A user's emotional requirements are recognized, and providing those needs is likely to result in a satisfactory connection with that user. It is the motivation for human emotion learning. Although it is difficult to comprehend and model human emotions, it is possible to organize and quantify the intensity of emotion experienced by a person when engaging with a system [10–14].

It is necessary to identify significant patterns from the collected data to recognize emotional information [15–18]. Various algorithms, such as voice recognition, natural language processing, and facial expression detection, are used to parse the data, all of which are based on the human component rather than programming, and hence, they need human intervention [19]. We study emotional learning in the context of a mixed-initiative interaction [19], in which the system alters its behavior or function in response to the user's mood. The most typical technique to developing an emotional system is to learn about the user's emotional state and then operate the system in a variety of different modes. We aimed to investigate the link between human emotion and computer interaction in this research. Table 1 represents the parameter of the input data set.

TABLE 1: Parameters of the recommender system.

User Id	Movie Id	Tag	Timestamp
2	60756	Funny	1.45E + 09
2	60756	Highly quotable	1.45E + 09
2	60756	Will Ferrell	1.45E + 09
2	89774	Boxing story	1.45E + 09
2	89774	MMA	1.45E + 09
2	89774	Tom Hardy	1.45E + 09
2	106782	Drugs	1.45E + 09
2	106782	Leonardo DiCaprio	1.45E + 09
2	106782	Martin Scorsese	1.45E + 09
7	48516	Way too long	1.17E + 09
18	431	Al Pacino	1.46E + 09
18	431	Gangster	1.46E + 09
18	431	Mafia	1.46E + 09
18	1221	Al Pacino	1.46E + 09

In accordance with their emotional state, each individual works and interacts with the system in a distinct manner [20–22]. Consequently, it is reasonable to assume that each user will engage with the system in a personalized manner. As a result, the kind of support a user desires and the activities a user expects from the system alter, depending on the user’s emotional state. In the case of a particular user who is under a lot of stress, he or she may choose to relax for a few hours or do anything monotonous. As a result of the system’s improved emotional awareness of the user, the system provides new means of engagement in response [23]. When the user is under extreme stress, the system may instruct him or her to work in safe mode. In the same manner, we must recognize and respond to the emotional states of various users to adjust the system’s engagement with them appropriately [24].

Briefly, the significance of emotional learning rests in its ability to enable the user to provide free input while engaging with the system. The next section explains the range of human emotional states and their implications when dealing with a system [25], which should be taken into account while designing new tactics for engaging with it. It is also necessary to find new technologies and procedures for determining the emotional condition of people [26].

3. Model System Design for Sentiment Analysis while Video Playing

There are two possible outcomes for the data:

It may be regarded as expressing which movies the users have seen (and rated) and which films they have not watched. In this case, users’ timepieces serve as a type of tacit feedback, letting us know which items they want to view and which they would rather not see.

It may also be seen as a way of expressing how much the users enjoyed the films that they did view. It is an example of explicit feedback: assuming that a person has viewed a movie, we can infer how much they enjoyed it by looking at the rating they have given to the film.

An android mobile app is designed and developed to play the videos in it. In this system, the mobile app is designed to recommend different videos to the users. Based on their experience, the user instructs the other video contents to other users. The system design asks a different question when watching the video [27], and the user is giving the answers. By considering this feedback, the users are classified into different categories. The factors taken for the classification of users are as follows: 1) mindset, 2) human hardness, 3) positive emotion, and 4) negative emotion. Now, it will make a recommendation using the following KNN algorithm. The complete details of the KNN algorithm are presented in the latter part of this section. In this work, we have done a further analysis of different personalities and their recommendations for consideration. The effectiveness of these recommendations is analyzed to understand the relativity between the personality types and their recommendations [28]. Figure 1 depicts how the human mobile video content accessing architecture shows how video content accesses from live video streaming. Figure 2 shows a snapshot of a video mobile video app.

We divided the dataset into two groups: training sets and test sets. The following function offers two split modes, one of which is random and the other one is seq-aware. If one chooses the random mode, the function divides the 100k interactions randomly without taking into account the timestamp, and by default, it utilises 90 percent of the data as training samples and the remaining 10 percent of the data as test samples. In the seq-aware mode, we exclude the item that a user has rated the most recently from the test, as well as the users’ previous interactions from the training set. The historical interactions of users are arranged from the oldest to the newest, depending on the timestamp of their interactions. It will be utilized in the sequence-aware recommendation phase of the Algorithm 1.

Eight textual tweets and review datasets from diverse fields are used to train and evaluate hybrid deep sentiment analysis learning models that integrate long short-term memory (LSTM) networks, convolutional neural networks (CNN), and support vector machines (SVM).

Each user profile is a combined group of emotional attributes that will be modeled as $U_x = \{Ex_1, Ex_2, Ex_3, Ex_4, Ex_5\}$, where “Ex” represents the emotion status of the user, i.e., Ex1 is representing the hardness factor, Ex2 is representing the stress factor, Ex3 is representing the positive emotion factor, Ex4 is representing the negative emotion factor, and Ex5 is representing positive thinking/negative thinking factor [29]. The contents are selected based on the user’s relation with the user and the movies.

In this experimentation, the IMDb Dataset is taken for experimentation. It is a dataset available for access to customers for personal and noncommercial use. The IMDb Movie Reviews dataset is a binary sentiment analysis dataset consisting of 50,000 reviews from the Internet Movie Database (IMDb) labeled as positive or negative. The dataset contains an even number of positive and negative reviews. Only highly polarizing reviews are considered. It is further tuned to meet our requirements in this work.

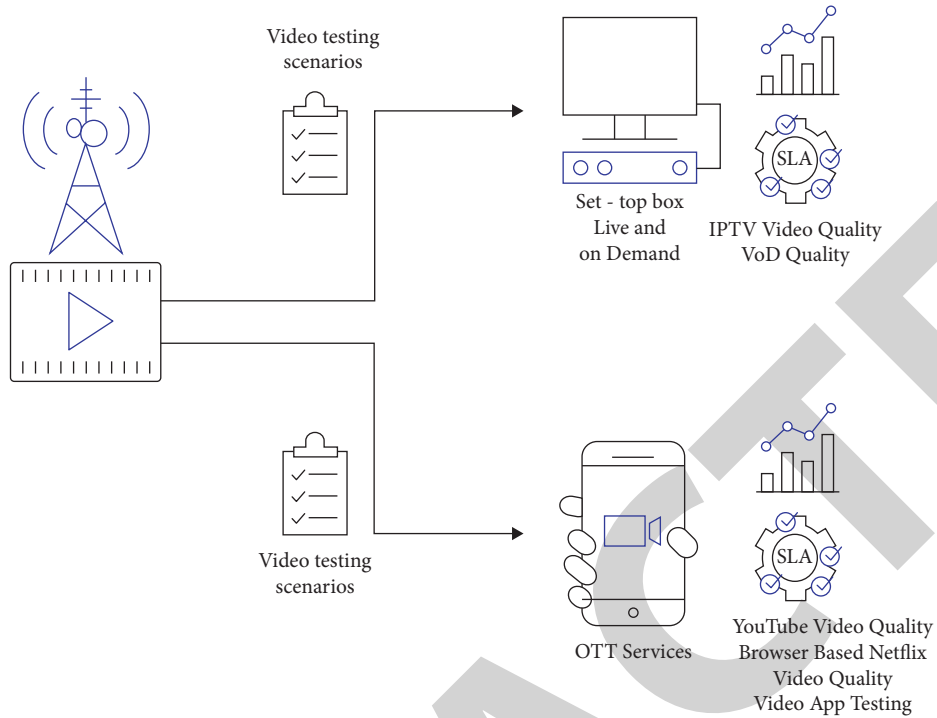


FIGURE 1: Human mobile interaction system in accessing video contents.

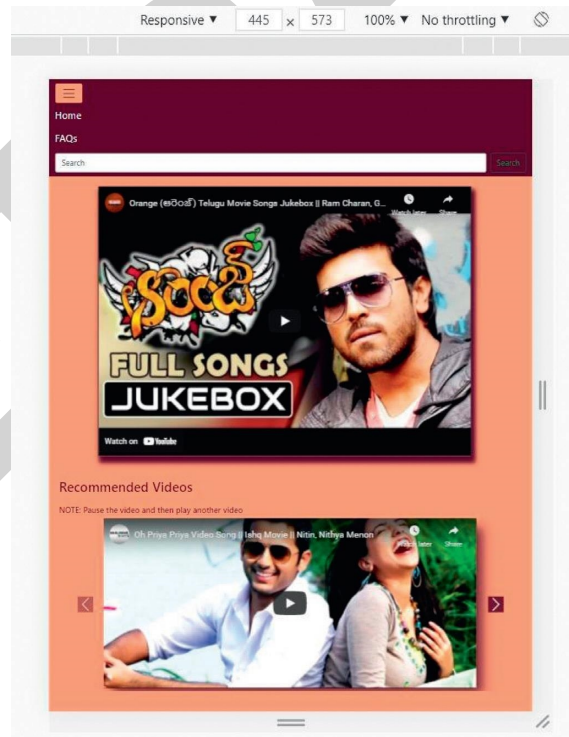


FIGURE 2: A snapshot of Video Mobile App.

The performance of the proposed recommendation scheme is evaluated based on the evaluation metrics of accuracy, precision, recall, and F-measure [30]. The accuracy metrics are

defined as the potential in predicting labels associated with a categorical class. In other words, it is computed as the proportion of correctly predicted instances as specified in (1)₂

- (1) Formulate the table column with user emotional categories and column with the movie names.
- (2) Load the training data.
- (3) Prepare data by scaling, missing value treatment, and dimensionality reduction as required.
- (4) Find the optimal value for K .
- (5) Predict and recommend a class value for new data.
- (6) Calculate distance (X, X_i) from $i = 1, 2, 3, \dots, n$.
- (7) where X = new data point, X_i = training data, distance as per your chosen distance metric.
- (8) Sort these distances in an increasing order with the corresponding train data.
- (9) From this sorted list, select the top " K " rows.
- (10) Find the most frequent class from these chosen " K " rows. It will be the predicted class for recommendation.

ALGORITHM 1: KNN Algorithm for grouping the similar set of users.

$$Accuracy = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}}. \quad (1)$$

Precision refers to the closeness measures of instances with one another, and it is computed based on (2).

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}}. \quad (2)$$

Recall is defined as the exact positive instance in the dataset that has been accurately determined as positive by the utilized classifier [31], as calculated based on (3).

$$Recall = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}. \quad (3)$$

In addition, the F -Measure is computed based on the weighted harmonic mean of the recall and precision as mentioned in (4).

$$F\text{-Measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (4)$$

TP (true positive): if the instance is positive and the outcome of the predicted recommendation is also positive. TN (true negative): if the instance is negative and the outcome of the predicted recommendation is also negative [32]. FN (false negative): if the record is positive and the outcome of the predicted recommendation is also negative. FP (false positive): if the data record is negative and the outcome of the predicted recommendation is also positive.

4. User-Interaction Emotional Issues: Empirical Study

The goal of our experiment was to study different aspects of the recommendations between a user and the mobile under various mindsets. Specifically, we learned classified different types of events/situations that evoked positive and negative emotions. Furthermore, the study extends the impact of positive and negative feelings in human-mobile interaction. Besides, we studied various critical aspects of human hardness to estimate the average interacting capacity [33]. This section is organized as follows: Section 4.1 describes our experiment participants' characteristics and background. Section 4.2 describes our experiment

procedure. Finally, Section 4.3 presents the results we have obtained.

4.1. Experiment Participants. Thirty users ranging in age from 19 to 45 years—14 male and 16 female—participated in the experiment. Five different categories of participants, such as 1. Sports Videos Watchers (23%), 2. Movie Videos Watchers (13%), 3. Social Video Watchers (20%), 4. Student/Learner Videos Watchers (17%), and 5. Medical/Professional/Business Man Content Watchers (27%). These video watchers have specifically watched a set of videos in Table 2, and Figure 1 shows the percentage.

We categorized the first group of participants as hard personality (37%), the second group as moderately hard personality (43%), and the last group of participants as low-hard personality (20%). Figure 2 shows the personality traits survey done using one-time questionnaire by a professor from the University of Chicago.

4.2. Experiment Procedure. All participants were given a survey with questions regarding their mindset at the beginning of playing the videos and the performance after watching the videos. The users were asked to answer the questionnaire providing as much information as possible. We consider all kinds of answers to be equally important and valuable for our study since we want to know how the different users feel about watching videos. The user is permitted to recommend a set of videos to the other users. The effectiveness of the user recommendation is measured and studied in this work.

4.3. Experiment Results. The following sections describe in detail the results we have obtained from our experiment. How does human hardness impact human-mobile interaction? Each section addresses a particular issue regarding human hardness and emotional states in user interaction. Based on their hardness level, different videos were recommended in understanding the personality.

TABLE 2: Different types of participants and the Videos used by them.

S.No.	Type of participants	Type of contents	Number of users		
			Male	Female	Total
1	Sport videos watchers	Swimming, badminton, wrestling, Olympic shooting, cricket, football, tennis, hockey, ice hockey, kabaddi, gymnasium, weight lifting, volleyball, table tennis, baseball, Formula, MotoGP, chess, boxing, fencing, and basketball	5	2	7
2	Movie videos watchers	Action, comedy, drama, fantasy, horror, mystery, romance, thriller, and western	2	2	4
3	Social video watchers	Facebook, Instagram, Twitter, tutorials and how-to videos, product demo videos, user-generated videos, announcements/reveals, interview, and Q&A videos, event videos, behind-the-scenes, videos that promote exciting offers and deals, tell relatable stories, and final thoughts	2	4	6
4	Student/Learner videos watchers	Teaser videos and course videos	4	1	5
5	Medical/Professional/Business man content watchers	Surgery videos, Gynaecology and STDs, health and fitness, orthopaedics, cardiology, plastic surgery, medical examination, clinical skills, product videos, explainer videos, onboarding videos, internal training videos, testimonial videos, promotional videos, company culture videos, video voicemails, aerospace engineering videos, chemical engineering videos, electrical and electronics engineering videos, petroleum engineering videos, telecommunication engineering videos, machine learning and artificial intelligence videos, robotics engineering videos, and biochemical engineering videos.	6	2	8

4.3.1. Does Human-Hardness Impact the Human-Computer Interaction? Each person has a different personality and interacts with the system in a certain way. The hardness of each person differs because of other reasons. Other methods are available for measuring the user's mental status [11, 12, 14, 34]—the technique of hardness assessment is given in Annexure-A. For example, some people are with high hardness, while some people have low hardness. The average interaction score is fixed for each type of interaction task, and such a score converts to an average score of 10. Then, based on the hardness score, the users classify into three categories, namely the high-hardness person, moderate hardness person, and low-hardness person. Their recommendation considers attributes, such as accuracy, precision, recall, and F1-measures, plotted in the following Figures 3–7. It is observed that the high-hardness persons have a good level of recommendations compared with those of moderate hardness and low hardness. It established that the hardness aspect of the system plays a vital role in the mobile/computer-human interaction. Therefore, human hardness increases, and there is a high interaction performance. Table 3 includes the average mobile interaction score.

4.3.2. Does Stress Impact Human-Computer Interaction? Each person has different encounter situations in everyday life, and the conditions and events change the stress rate of the person. Annexure B gives the list of stress events and their rating, and each person has a different stress rate than their computer/mobile-interaction efficiency on two additional days. The stress rate readjusts to the given hardness and interaction values. Then, based on the stress score, the user classification is carried out into two categories. One is the person with high stress and the person with low-stress,

and their recommendation considers attributes, such as accuracy, precision, recall, and F1-measures plots, in the following Figures 8–11. It observes that the low-stress persons are having a good level of recommendations compared with high-stress persons.

4.3.3. Does Positive Emotion Impact Human-Computer Interaction? Each person is encountering different situations in day-to-day life. The conditions and events change the emotional status of the person. Each person with different positive emotional levels is asked to present their emotional status twice or thrice a day for a particular period. Its average compares with their recommendations efficiency. Based on the positive emotion score, the users are classified into two categories. One is the person with high-positive emotion and the other is the person with low positive emotion. Their recommendation considers that the high-positive feelings persons have a good level of requests compared with high-stress persons. Appendix C presents the list of positive emotional events, and their ratings as shown in Figures 12–15.

4.3.4. Does Negative Emotion Impact Human-Computer Interaction? Each person has seen a different situation in everyday life. The situations and events change the emotional status of the person. Appendix-D presents the list of negative emotional events and their rating. Each person with different harmful emotional levels was asked to explain their emotional situation twice or thrice a day for a particular period. Its average compares with their recommendations efficiency. Based on the negative emotion score, the users are classified into two categories. One is the person with high-negative emotion. The person with low negative emotion

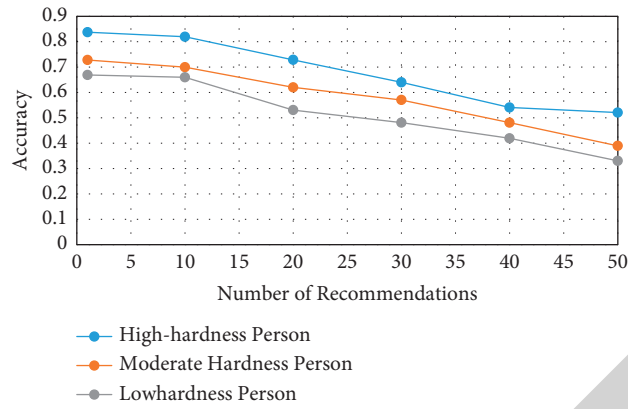


FIGURE 3: Accuracy measure of different human hardness of recommendations.

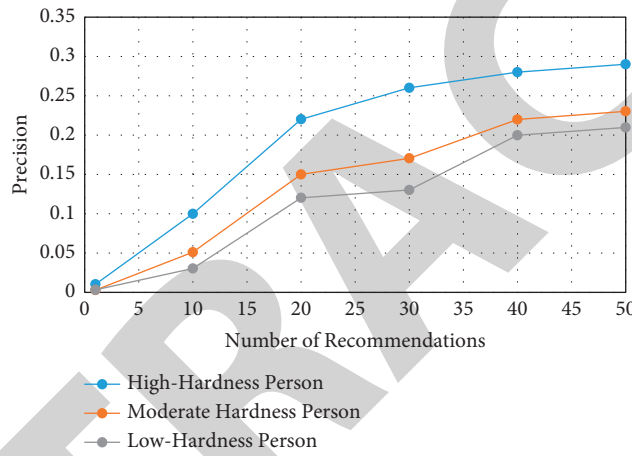


FIGURE 4: Precision measure of different human hardness of recommendations.

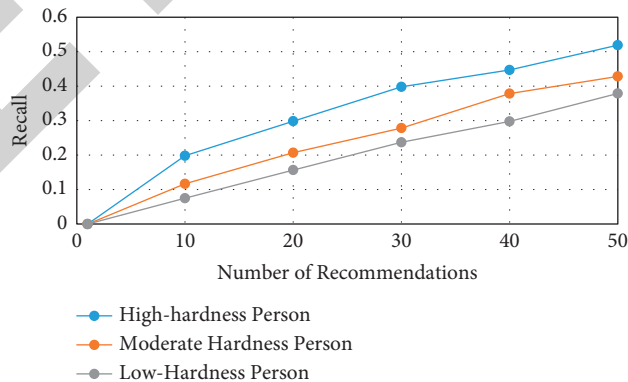


FIGURE 5: Recall measure of different human hardness of recommendations.

and their recommendations consider attributes, such as accuracy, precision, recall, and F1-measures plots, in the following Figures 16–19. It observes that the low-negative emotion persons have a good level of recommendations compared with those of high-negative emotional people.

4.3.5. *Does Positive Thinking Impact Human-Computer Interaction?* Each person has a different personality and interacts with the system in a certain way. The positive attitude of each person differs because of various reasons. Annexure-E presents the method of attitude assessment.

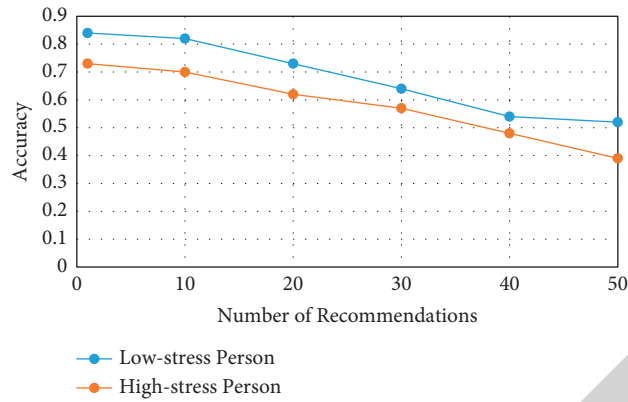


FIGURE 6: Accuracy measure of different stress level of recommendations.

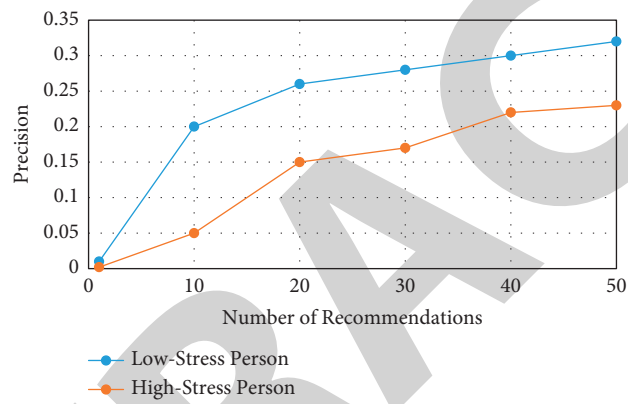


FIGURE 7: Precision measure of different stress level of recommendations.

TABLE 3: Average mobile interaction score.

S. No.	Type of watchers	Type of contents watched	Interaction score									
1	Sport videos watchers	Volleyball, cricket, football, tennis	45	50	55	60	65	70	75	80	85	90
			1	2	3	4	5	6	7	8	9	10
2	Movie videos watchers	Action, comedy, drama, fantasy, romance	10	20	30	40	50	60	70	80	90	100
			1	2	3	4	5	6	7	8	9	10
3	Social video watchers	Facebook, tutorials videos, Product demo videos, interview, and Q&A videos, event videos	6	8	10	12	14	16	18	20	22	24
			1	2	3	4	5	6	7	8	9	10
4	Student/Learner videos watchers	Teaser video, course videos	1-3			4-7			8-12			
			4			8			10			
5	Medical/Professional/Business man content watchers	Surgery videos, health and fitness, medical examination, clinical skills	1			2			3			
			4			8			10			

For example, some people have a high positive thinking attitude, and some people have a low positive thinking attitude. The score is fixed for the average interaction for each interaction task, and such a score converts to an average score of 10. Then, the average actual observed interaction performance is compared to the positive thinking attitude score and is presented in the following Figure 7.

The following table presents the simple correlation between the interaction and between the attributes compared. Table 4 represents the correlation details.

Let AC, HD, ST, PE, NE, and PT be represented by $X_1, X_2, X_3, X_4, X_5,$ and $X_6,$ respectively, since X_1 is dependent on $X_2, X_3, X_4, X_5,$ and $X_6.$ We take the measurements of these variables from their expected values. Consequently,

$$E(X_1) = 0 = E(X_2) = E(X_3) = E(X_4) = E(X_5) = E(X_6). \tag{5}$$

Let the standard deviation be $\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5,$ and $\sigma_6.$ Let the correlation between these variables be $r_{12}, r_{13}, r_{14}, r_{15}, r_{16}, r_{23}, r_{24}, r_{25}, r_{26}, r_{34}, r_{35}, r_{36}, r_{45}, r_{46},$ and $r_{56}.$

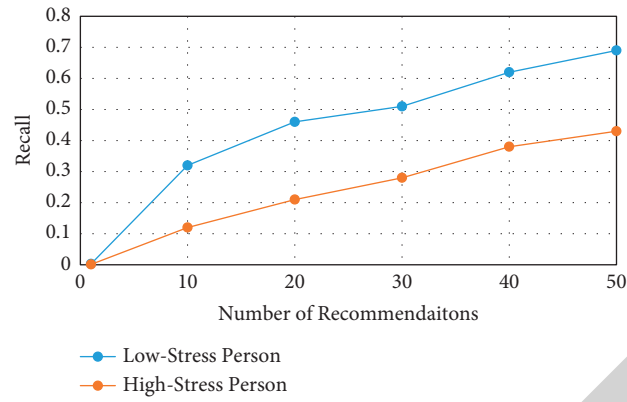


FIGURE 8: Recall measure of different stress level of recommendations.

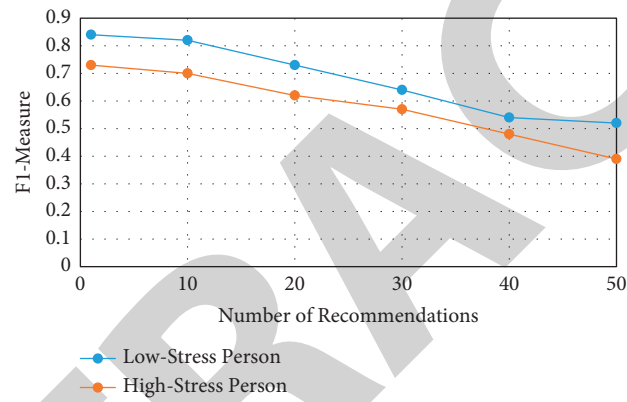


FIGURE 9: F1 measure of different stress level of recommendations.

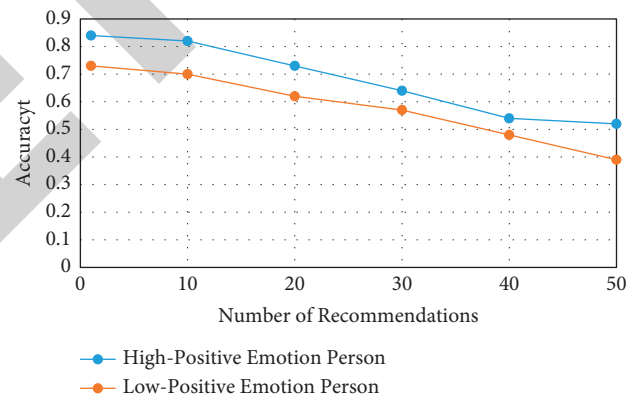


FIGURE 10: Accuracy measure of different positive emotion person of recommendations.

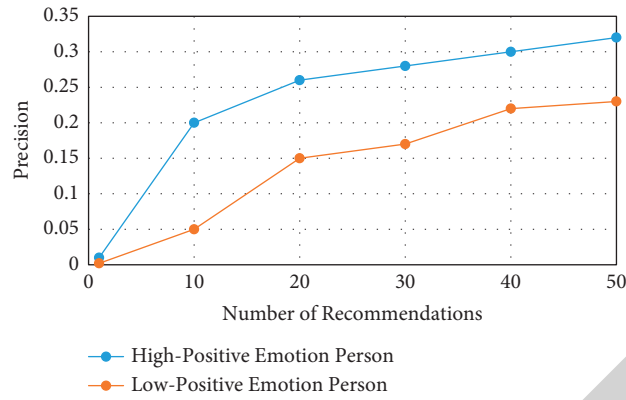


FIGURE 11: Precision measure of different positive emotion person of recommendations.

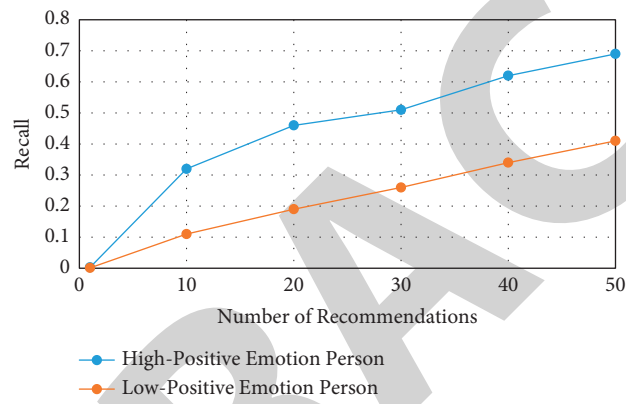


FIGURE 12: Recall measure of different positive emotion person of recommendations.

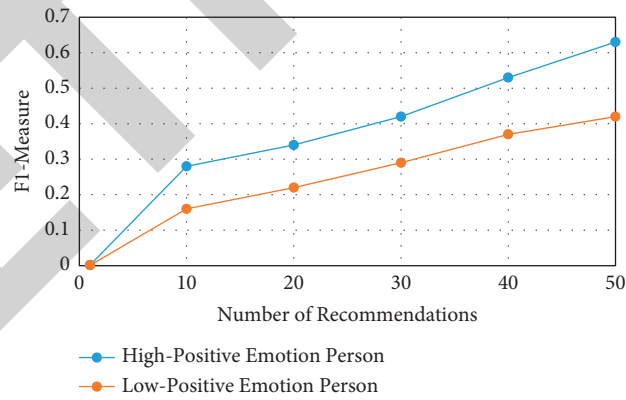


FIGURE 13: F1-measure of different positive emotion person of recommendations.

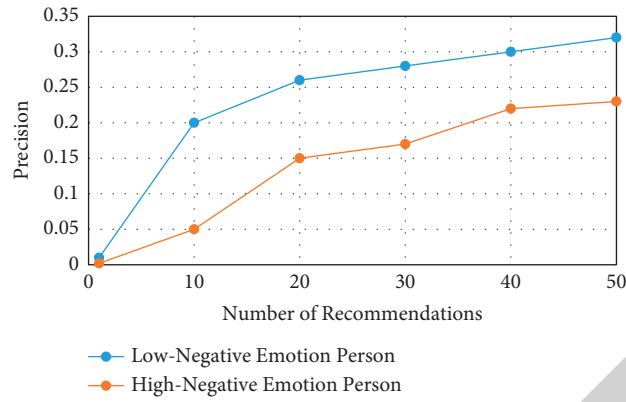


FIGURE 14: Precision measure of different negative emotion person of recommendations.

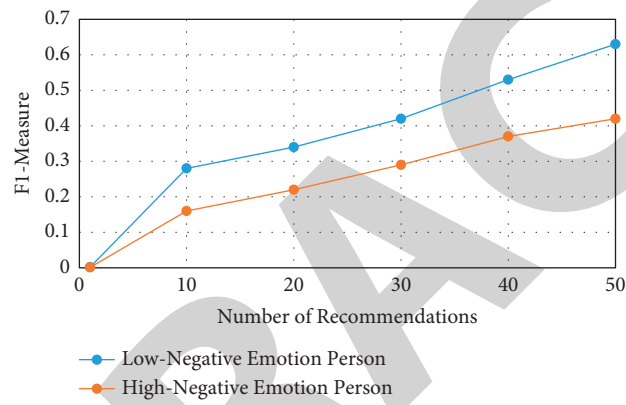


FIGURE 15: F1 measure of different negative emotion person of recommendations.

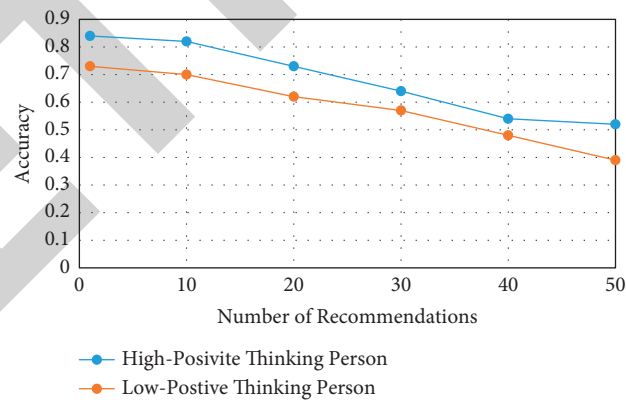


FIGURE 16: Accuracy measure of different positive thinking person of recommendations.

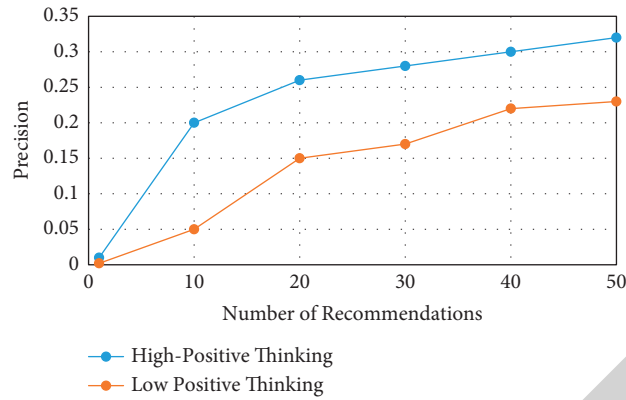


FIGURE 17: Precision measure of different positive thinking person of recommendations.

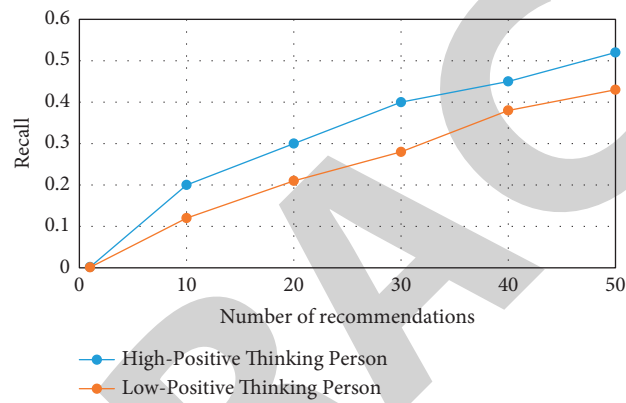


FIGURE 18: Recall measure of different positive thinking person of recommendations.

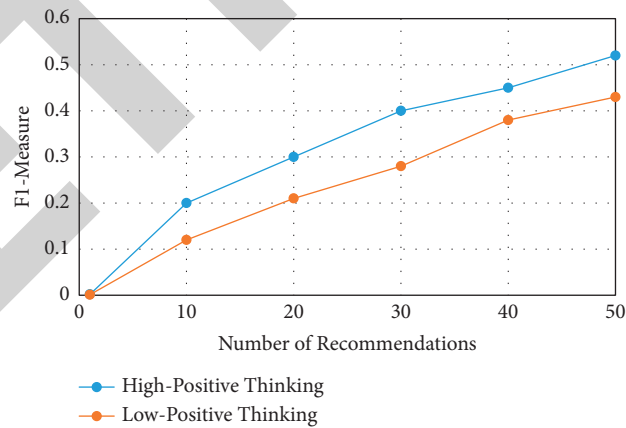


FIGURE 19: F1-measure of different positive thinking person of recommendations.

TABLE 4: Correlation between accuracy and other attributes.

Attribute 1	Attribute 2	Partial correlation
Accuracy (AC)	Hardness (HD)	0.882156
Accuracy (AC)	Stress (ST)	-0.51087
Accuracy (AC)	Positive emotion (PE)	0.42498
Accuracy (AC)	Negative emotion (NE)	-0.377452
Accuracy (AC)	Positive thinking (PT)	0.462025

It is possible to determine the plane of regress of $X1$ in terms of $X2, X3, X4, X5,$ and $X6$, and it is given by the following:

$$\left(\frac{X1}{\sigma_1}\right)\omega_{11} + \left(\frac{X2}{\sigma_2}\right)\omega_{12} + \left(\frac{X3}{\sigma_3}\right)\omega_{13} + \left(\frac{X4}{\sigma_4}\right)\omega_{14} + \left(\frac{X5}{\sigma_5}\right)\omega_{15} + \left(\frac{X6}{\sigma_6}\right)\omega_{16} = 0,$$

$$\omega = \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ r_{21} & 1 & r_{23} & r_{24} & r_{25} & r_{26} \\ r_{31} & r_{32} & 1 & r_{34} & r_{35} & r_{36} \\ r_{41} & r_{42} & r_{43} & 1 & r_{45} & r_{46} \\ r_{51} & r_{52} & r_{53} & r_{54} & 1 & r_{56} \\ r_{61} & r_{62} & r_{63} & r_{64} & r_{65} & 1 \end{bmatrix} \begin{bmatrix} 1 & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ r_{21} & 1 & r_{23} & r_{24} & r_{25} & r_{26} \\ r_{31} & r_{32} & 1 & r_{34} & r_{35} & r_{36} \\ r_{41} & r_{42} & r_{43} & 1 & r_{45} & r_{46} \\ r_{51} & r_{52} & r_{53} & r_{54} & 1 & r_{56} \\ r_{61} & r_{62} & r_{63} & r_{64} & r_{65} & 1 \end{bmatrix}. \quad (6)$$

We note here that the corresponding cofactor of the element in the i^{th} row and j^{th} column is ω_{ij} for all i, j ranging from 1 to 5.

The coefficient of multiple correlations in six variate distributions in which each of the variables $X1, X2, X3, X4, X5,$ and $X6$ in terms of the mutual correlations between i, j is given by the following:

$$R_{1.23456}^2 = 1 - \frac{\omega}{\omega_{11}}. \quad (7)$$

The coefficient of multiple correlations estimates the closeness of association between the observed values and the expected value of a variable obtained from the multiple linear regression of that variable on other variables.

$R_{1.23456}^2$ is a simple correlation between $X1$, and the error estimate e_1 is 0.23456. It must lie between -1 and $+1$. However, in our case, it will be $0 \leq R_{1.23456}^2 \leq 1$. It establishes that the correlation between the hardness and accuracy is highly correlative. Therefore, it is considered to be one of the important attributes taken for analysis.

5. Experimental Learning

The experiment showed several fascinating features of an emotion's influence in human-computer interaction. Computer systems or software systems have been unable to personalize or comprehend the user's emotions. We observed several challenges that exchange-system designers must address to personalize and optimize the interaction system and users.

- (i) Understand the relation between user interaction and the hardness of the person.
- (ii) Understand the relation between user interaction and the stress rate of the person.
- (iii) Understand the relation between user interaction and the positive emotion of the person on a particular day.

(iv) Understand the relation between user interaction and the negative emotion of the person on a particular day.

(v) Understand the relation between user interaction and the positive thinking attitude of the person.

(vi) Discovering the method for providing implicit feedback to capture these attributes.

(vii) Discovering the new method of mobile interaction mechanisms to each set of users.

6. Implications in Human-Emotional System Development: How to Incorporate Emotional Knowledge

To provide better interaction to the computer user, the system needs some emotional knowledge about them. Now, this knowledge is collected through explicit feedback. However, we need to develop new devices, mechanisms, and methods to manage these knowledge sets automatically. The most crucial fact about human emotion is that it is both instantaneous and intrinsic. Our approach analyzes the people's relativity hardness, stress, positive emotion, negative emotion, and positive thinking attitude. It is hard to understand the positive and negative emotions because it is instantaneous. The system uses four ways to collect emotional details. The first method is to observe the user's emotional status through the new hardware devices, utilizing relative measures like blood pressure. The second method is to identify facial expressions, and the third method is to collect explicit feedback from the user. The fourth method is information personalization. A commonly used metaphor for accepting the user's emotional status is consuming a personal assistant interface-agent collaborating with the user in some work environment. Over some time, the system will acquire a complete emotional attitude of the user.

In general, users' emotional information can be fed to the system, and it is possible to adapt the user for both

system and application software development cases. All these emotional statuses are kept under the category of the application-independent information domain.

7. Conclusions

This paper has studied the relativity between user interaction and emotional attributes, such as hardness, stress, positive emotion, negative emotion, and positive thinking. The study presented the results of an experiment to learn the impact of incorporating emotional knowledge into the system.

This work gives an idea about the understanding of user-emotion status while interacting with the system. This study has implications on designing an interface with emotional knowledge to improve the system's overall efficiency. There are seven significant issues for the incorporation of emotional understanding, which are as follows:

- (i) Understand the relation between user interaction and the hardness of the person.
- (ii) Understand the relation between user interaction and the stress rate of the person.
- (iii) Understand the relation between user interaction and the positive emotion of the person on a particular day.
- (iv) Understand the relation between user interaction and the negative emotion of the person on a particular day.
- (v) Understand the relation between user interaction and the positive thinking attitude of the person.
- (vi) Discovering the method for providing implicit feedback to capture these attributes.
- (vii) Discovering the new method of machine interaction mechanisms to each set of users.

We have developed techniques for learning and incorporating emotional information to address the issues outlined above. We are also focusing on dynamic user profile techniques that analyze user interactions to determine emotional user status. One can use the results of these algorithms in future systems.

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] C. Peter and R. Beale, "Affect and emotion in human-computer interaction," LNCS, Springer, Berlin, Germany, 2008.
- [2] M. Jeon, "Emotions and affect in human factors and human-computer interaction: Taxonomy, Theories, approaches, and methods," *Emotions and Affect in Human Factors and Human-Computer Interaction*, Elsevier, Amsterdam, Netherlands, pp. 3–26, 2017.
- [3] J. Aday, W. Rizer, and J. M. Carlson, "Neural mechanisms of emotions and affect," in *Emotions and Affect in Human Factors and Human-Computer Interaction*, pp. 27–87, Elsevier, Amsterdam, Netherlands, 2017.
- [4] H. A. Elfenbein, B. Sigal, and E. C. Noah, "Supplemental material for the social perception of emotional abilities: expanding what we know about observer ratings of emotional intelligence," *Emotion*, vol. 15, no. 1, 2015.
- [5] Z. Kövecses, "Happiness in context," *Where Metaphors Come From*, pp. 155–175, 2015.
- [6] D. Ellis and I. Tucker, "Artificial intelligence and emotion," *Emotion in the Digital Age*, Routledge, Oxfordshire, England, UK, pp. 32–55, 2020.
- [7] J. Vallverdú, "Why robots must have synthetic emotions? The role of emotions in the artificial cognitive systems," *Proceedings*, vol. 1, no. 3, p. 272, 2016.
- [8] J. Shen, O. Rudovic, S. Cheng, and M. Pantic, "Sentiment apprehension in human-robot interaction with NAO," in *Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, Xi'an, China, September 2015.
- [9] S. Freed, "Prevailing Prejudices pertaining to artificial intelligence," *AI and Human Thought and Emotion*, Auerbach publications, Boca Raton, FL, USA, pp. 65–82, 2019.
- [10] A. Angeli and G. Johnson, "Emotional intelligence in interactive systems," *Design and Emotion*, Taylor and Francis Group, Oxfordshire, England, UK, pp. 262–266, 2003.
- [11] J. S. De Freitas, R. Gudwin, and J. Queiroz, "Emotion in artificial intelligence and artificial life research: facing problems," in *Proceedings of the Intelligent Virtual Agents, 5th International Working Conference, IVA 2005*, p. 501, Kos, Greece, September 2005.
- [12] S. Mahlke, M. Minge, and M. Thüring, "Measuring multiple components of emotions in interactive contexts," in *Proceedings of the CHI '06 extended abstracts on Human factors in computing systems - CHI EA '06*, Montréal Québec Canada, April 2006.
- [13] P. Desmet, "Measuring emotion: development and application of an instrument to measure emotional responses to products," *Human-Computer Interaction Series*, vol. 3, pp. 391–404, 2018.
- [14] R. Gupta, P. K. Shukla, and P. K. Shukla, "Performance analysis of anti-phishing tools and study of classification data mining algorithms for a novel anti-phishing system," *International Journal of Computer Network and Information Security*, vol. 7, no. 12, pp. 70–77, 2015.
- [15] S. Ran, S. Hsu, and T. Jung, "Examining the relationship between EEG dynamics and emotion ratings during video watching using adaptive mixture independent component analysis," in *Proceedings of the 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Toronto, Ontario, Canada, April 2020.
- [16] S. Hsu, Y. Zi, Y. C. Wu, P. M. Jackson, and T. Jung, "Exploring mental state changes during hypnotherapy using adaptive mixture independent component analysis of EEG," in *Proceedings of the 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Cleveland, Ohio, USA, October 2018.
- [17] P. Tiwari and P. Shukla, "Artificial neural network-based crop yield prediction using NDVI, SPI, VCI feature vectors," in *Information and Communication Technology for Sustainable Development. Advances in Intelligent Systems and Computing*,

Retraction

Retracted: A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] N. Arunkumar, V. Pandimurugan, M. S. Hema et al., "A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9441357, 11 pages, 2022.

Research Article

A Versatile and Ubiquitous IoT-Based Smart Metabolic and Immune Monitoring System

N. Arunkumar ¹, V. Pandimurugan ², M. S. Hema ³, H. Azath ²,
S. Hariharasitaraman ², M. Thilagaraj ⁴, and Petchinathan Govindan ⁵

¹Department of Biomedical Engineering, Rathinam Technical Campus, Coimbatore-641021, India

²School of Computing Science and Engineering, VIT Bhopal University, Kotri Kalan, Ashta, Near Indore Road, Bhopal, Madhya Pradesh 466114, India

³Anurag University, School of Engineering, Department of Information Technology, Venkatapur, Ghatkesar Road, Hyderabad, Telangana-500088, India

⁴Department of Electronics and Instrumentation Engineering, Karpagam College of Engineering, Coimbatore, India

⁵Department of Electrical and Electronics Technology, Ethiopian Technical University, Addis Ababa, Ethiopia

Correspondence should be addressed to Petchinathan Govindan; petchinathan.govindan@etu.edu.et

Received 26 December 2021; Revised 31 January 2022; Accepted 5 February 2022; Published 2 March 2022

Academic Editor: Deepika Koundal

Copyright © 2022 N. Arunkumar 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 the present medical age, the focus on prevention and prediction is achieved using the medical internet of things. With a broad and complete framework, effective behavioral, environmental, and physiological criteria are necessary to govern the major healthcare sectors. Wearables play an essential role in personal health monitoring data measurement and processing. We wish to design a variable and flexible frame for broad parameter monitoring in accordance with the convenient mode of wearability. In this study, an innovative prototype with a handle and a modular IoT portal is designed for environmental surveillance. The prototype examines the most significant parameters of the surroundings. This strategy allows a bidirectional link between end users and medicine via the IoT gateway as an intermediate portal for users with IoT servers in real time. In addition, the doctor may configure the necessary parameters of measurements via the IoT portal and switch the sensors on the wearables as a real-time observer for the patient. Thus, based on goal analysis, patient situation, specifications, and requests, medications may define setup criteria for calculation. With regard to privacy, power use, and computation delays, we established this system's performance link for three common IoT healthcare circumstances. The simulation results show that this technique may minimize processing time by 25.34%, save energy level up to 72.25%, and boost the privacy level of the IoT medical device to 17.25% compared to the benchmark system.

1. Introduction

One of the main paradigms of networking is the internet of things (IoT), which spreads through a variety of claims, providing central access to and convergence of information [1]. Users and approved workers, such as medical practitioners, can have access to information according to the mission description for each person. Privacy and confidentiality, sensitive data security, and limited functionality are essential for healthcare [2]. IoT can connect the internet to a variety of sensors, cars, homes, and computers, enabling people to exchange statistics, evidence, and services.

This allows for information synthesis that can make the study, usability, and comfort of usage of data in submissions very important. IoT versatility has brought a number of new developments toward better data access and increased resource utilization and information sharing, among multiple causes, to boost complete data quality performance [3]. This is becoming possible thanks to advanced protocol networking technology innovations, high internet concentration, and consumer accessibility of large infrastructure. As a result, people are more concerned with consolidated acquisition and evaluation of data to save time and energy [4].

IoT's clever cities, intelligent houses, infrastructure, and economic surveillance are all critical issues. Healthcare has been one of the most critical issues, with accelerated industrialization, urbanization, and an ever-growing rate of senior citizens in Europe in recent years. The emphasis in healthcare is progressively moving from conventional methodologies, such as postdiagnostic therapy, to preventive and predictive health security. This pattern involves constant and systematic criteria for the monitoring of the individual's historical data from various fields of healthcare. This clinical internet of things is the foremost emphasis of the modern century of healthcare, after positive experiments with electronic (e-health) and mobile (m-health) [5].

An IoT-based medical platform is able to integrate and merge variables (on the server side) from various fields that can lead to protecting healthcare [6]. In 2016, the WHO stated that the second field in connection with healthcare that causes one in seven deaths was environmental criteria control involving physicochemical components. The larger and more advanced facilities most frequently follow these criteria. The facilities are only limited to being scattered by high operating costs, complicated calibration/recalibration, and advanced facilities in certain countries and towns. These stations, on the one hand, are not available anywhere for the environment, but on the other hand, they only provide an overview of the area [7]. The overview of medical healthcare is shown in Figure 1.

In this assessment, the environmental parameter category includes toxic/hazardous gases, sound, ultraviolet conditions, temperatures, moisture, and air pressure. Environmental sensitivity and physiological parameters are, therefore, two of the most critical areas for monitoring in the field of healthcare. In p2Health, the physiological and biochemical variables must be constantly tracked and customized. Several studies have identified adverse effects on the health status of environmental contaminants, particularly on patients with physiological parameters and vital indications [8].

Depending on the time of exposure, accumulation and quantity of toxins, and the clinical state of patients, the sensitivity of people with cardiac disease and cardiovascular disease to chemical air pollution is a major factor in their breathing rate and heart failure. The harmful effects of environmental contaminants are not only restricted to environmental criteria but also physical criteria. Physical environmental metrics classify the highest noise frequency, UV, weather, moisture, and intensity. Various studies have shown that noise has adverse effects on sleep efficiency [9]. It is given a higher weight in the case of people suffering from chronic illnesses. In monitoring patients with chronic obstructive pulmonary disease, it is critical, for example, to track the sound level as an environmental parameter.

The high level of sound will contribute to sleep disturbance and thereby affect the physiological state of the patient. Exposure to UV indexes above the threshold can also pose a health risk, particularly in patients with skin cancer and COPD. Similarly, air and moisture can create painful problems. Environmental toxins can have a direct or indirect impact on physiological parameters [10]. However,



FIGURE 1: Overview of medical healthcare.

continuous control, data synchronization, and the analysis of the relationship between environmental and physiological parameters are a leap forward to studying the effects and weights of each parameter on the other. The effect may differ from one parameter to the next.

The internal and external actors should be carefully investigated and determined. For p2Health to incorporate mIoT, customizable control of the parameters is taken into account by means of effective wearables. Wireless network nodes are an inseparable IoT tier that contains multiple sensor nodes and is used for inter/intradata communication between nodes and levels. In addition, data from different topics of focus are measured, gathered, and transmitted, even inside the WSN. The composed information is communicated from the wireless sensor node to the receiver through an IoT entry to construct a database according to the physiological and environmental indicators of a person 24/7. The WSN can be extended to stretch the observation border, as appropriate, by a variety of nodes to various extents of medical concern.

All the key components of the WSN are wearable sensor nodes that are used for custom surveillance. The functional structure, with temperament proportion, breathing frequency, blood pressure, and physique illness, is most often used for the physique in the Tuner Frame Zone System. While in customized healthcare, the acquisition, observation and loading, and transfer of every information collection have already been demonstrated, and both physiological and environmental monitoring by means of a durable and effective approach remain unintegrated.

Furthermore, an effective wearable is lacking in the measuring of ambient parameters. Integration of data must be illustrated in an appropriate approach and framework using powerful strategies/models aimed at full and then customized observations of active healthcare constraints. This explanation is specified and then applied to the IoT, which must also be able to host a massive amount of data in real time. In addition, data acquisition should be modular to allow for bidirectional coordination between licensed staff and end users for required medical commands [11]. Smart interpretation of data should also be available through the algorithms applied.

1.1. The Main Objective of This Study

- (i) We designed an IoT-based platform for a suitable wearability mode to comprehensively measure environmental, physiological, and behavioral factors.

- (ii) We offer a privacy-controlled download method to determine the discharge rate and local calculation rate for the data processing. This approach takes into account the present condition of the radio channel, the amount and importance of fresh healthcare sensing data, the estimated level, the bacterial level, and the computational work history to reduce computer delays, save on IoT device energy consumption, and enhance data protection.
- (iii) For medical analysis of parameter interactions, we integrate, synchronize, and process physiological and environmental parameters. This also involves seeing the data on the server.

1.2. Organization of the Study. In the introduction part are given the essential information about patient monitoring and how the IoT devices are most helpful for monitoring health in different aspects. The second part of this study is made of a detailed literature review for the existing works given for support of defining the problem. The third part of this study proposed a system and flow diagram of the framework of the immune monitoring system given with various parameters. The final part has the result and discussion with various analyses like temperature, utility, and respiration analysis like lot parameters compared and discussed.

2. Existing Work and Literature Review

IoT devices provide edge computing with lower energy use and computation delays, which enables computation-intensive and latency-sensitive applications. In order to lower the computer overhead for resource-limited mobile devices, the binary download as suggested picks a data transmission rate under a stochastic Wi-Fi channel with a single edge. In order to decrease energy usage under latency constraints on a multiuser grid, the partial reloading schema as presented employs time division and orthodontic frequency division. In order to lower the implementation latency and work error rate for the scenario with just a known server, the mobile offload strategy as presented leverages the Lyapunov optimization, provided that both the model transmission delay and the local implementation model are recognized.

Geman et al. presented that the attackers could be inquisitive about data security, such as where the user is located and how their IoT device is being used. Depending on how far the user is from the edge node, an edge device can learn about location information and IoT device usage patterns [1]. The privacy level is tied to the amount of sensing data that has to be sent and the offload rate. Koo and Hur mentioned and received the calculation reports from the device, the IoT device will first do the local processing, and after this, it will compare the current channel states with the previously saved channel states and data size to assess the privacy level that was obtained [12]. The IoT device incorporates techniques to purposely limit the offloading rate while the channel is healthy and to raise the offloading rate

when the channel is unhealthy in order to safeguard user privacy. Privacy is shown in red, since the consumption pattern indicates the difference between the amount of real sensing data and the amount of offloading data that takes place when increased wireless channel strength is present. Whether the IoT device stays in certain places with severely degraded radio channels is shown as location privacy.

The effective approach within IoT needs the careful handling of many problems by the application of WBAN in healthcare for environmental and physiological control. Due to critical continuous surveillance, it is of serious concern to consume power for short-term WBAN connectivity and long-range data transfer from the mobile to the server. The most critical requirements are wearability mode, flexible approach to sensor acceptance, the possibility of expansion, data collection range, accessibility, and fusion. Another part of healthcare is the relationship between physiological and environmental indicators, which involves data examination and can only be seen by data analysis by appropriate algorithm decisions based on an individual's continuous supervision [1].

Dawood et al. elaborated, reflects on the most current and important work in the monitoring of environmental, behavioral, and environmental biological restrictions. All other works offered in this segment fulfill the needs of the IoT in this field. An IoT-based tropospheric monitoring network has been established in [13]. Kim et al. suggested using a platform designed for the evaluation of environmental parameters like fine dust and ozone in various modes of contact in the short/long term. Each unit conveys a container containing information, and the position and operating position through the LTE system in this ecological measurement stage. This parameter is calculated by an applied board, and the data are analyzed on the server. The number of planning boards was used to gather data to track various air pollutants on this site.

A description of the IoT-based multisensor platform for atmospheric parameters to monitor nitrous gas (NO_x), carbon dioxide (CO), and ozone (O₃), and temperature, moisture, and air compression is presumed. The stage can track environmental contaminants at a low concentration. This infrastructure is built around two shields, one of which contains the ATmega328. One shield is used for gas sensors; the other shield is used for temperature, moisture, and wind speed sensors. These are planned as a fixed framework for server transmission based on the Wi-Fi networking protocol. The machine is Linux based for data observation [2].

The IoT-based solution for indoor air quality surveillance (IAQM) based on ATmega1281 was stated in [6]. In this job, a complete IAQM device enables carbon dioxide calculation and sulfur dioxide (SO₂). It presents nitrous gas (NO₂), O₃, chlorine (Cl₂), and environmental relative humidity. The boards in a star configuration are linked to a gateway via Zigbee. The transitional port passes the data via Wi-Fi to the cloud. As a front-end amplifier aimed at the devices, the sensors use the LMP91000 chip. All devices remain situated and arranged for protection. That makes this platform a fixed boxing station, useful for domestic surveillance [3].

A low-cost wireless network of sensors is presented with a protocol and Wi-Fi for long-range data transmission. A framework for calculating particulate matter has been established in this study. While, due to its limited portability, this device is best suited for regular environmental monitoring, its extensive monitoring supports many of the environmental parameters. A wearable single-gas detector was conceived and implemented for volatile organic compounds based on a capacitive micromachined ultrasonic transducer [4].

This watch has long-term surveillance, is low powered, and has a detection limit of 120 parts per trip. The data are transmitted via a low-power network to a mobile phone or Raspberry Pi 3 and through Wi-Fi from the gates to the cloud. Even though the wearability of the member nodes is efficient, the identification of the single aspect limits the scenarios for this instrument. P was approached with common sense as a handheld prototype of Olivier et al. This common-sense version tests many environmental toxins, namely, CO, NO, O₃ gas sensors, and sensors monitoring light, humidity, and body position, and O₃ gas sensors.

Data are transmitted from the prototype via a component from SparkFun to the smartphone. The information received is viewed on a mobile phone and transmitted via a Cinterion GPRS radio to a host server. These data are also presented and evaluated on the web server. There will be a customized CO₂ and O₃ ambient air detector. Ahamed et al. presented a strong calibration method that was considered in the production of W-air according to the ambient physical and textile parameters and breathing emissions. In order to estimate CO₂, W-air uses the V OCs CCS811 sensor. The relationship between VOCs and CO₂ in this prototype is expected to be solid [14].

The data obtained from each gas sensor were transferred to a smartphone for viewing at a time interval of five seconds and one minute, respectively. The IoT monitoring devices for environmental parameters are not confined to the above, but the methods proposed have been the most relevant and most recent research in this field. Various other works were discussed on indoor/outdoor environment control portable, mobile, and/or stationary equipment. This type of work addresses the issues like relative moisture and relative humidity raised in the preceding paragraph. The boards in a star configuration are linked to a gateway via Zigbee. The intermediate port passes the data via Wi-Fi to the cloud.

Sensors use flaws via front-end equipment to drive themselves. Because both sensors are arranged for protection, the device is useful as an immovable container position for domestic surveillance. In [15], a low-cost network of environmental sensors is demonstrated using a long-range data transfer protocol and Wi-Fi. An established platform for metal oxide (CO), (NO₂), hydrogen (H₂), ammonia (NH₃), and methane is presented in this work for the calculation of particulates, light, colors, steam, UV light, and methane. While, due to its limited portability, this device is best suited for regular environmental monitoring, its extensive monitoring supports many of the environmental parameters.

In 2012, a wearable single-gas detector was developed and applied to organic volatile compounds on the basis of a

capacitive micromachine ultrasonic transducer. This watch is built for sustained surveillance, is low powered, and has a LOD of 120 ppb. The data are sent in low-power mode to a smartphone or a Raspberry Pi 3. An advanced and portable medical surveillance system with multiple sensors is available. The chest is a control panel, ECG, temperature sensor, accelerometer, acceleration engine, light-emitting color changes, and pushbutton. It includes a chest-placed unit and a color-changing light-emitted diode. In reality, facial recognition and clinical specifications are tracked in this system along with location monitoring.

The user's health status can be monitored with the embedded vibration engine by identifying capacitive touch patterns. The incorporated color-shifting lead can be used to provide the holder with more insight into the present state of health. Finally, for emergencies, a pushbutton is issued. The e-health detector platform V2.0, Arduino, and Raspberry Pi are the first biometric shield. This system monitors the heartbeat, oxygen in the blood, air supply, body temperature, ECG, galvanic surface concentration, blood pressure, patient location, and muscle/electrochemicals. This device can measure pulse (EMG) [5].

Data are gathered through different alternatives: real-time patient status reporting and patient critical data transfer to a health department for assessment. Depending on the request, data transfer is enabled on the e-health sensor. A multisensor fusion approach is the foundation of the structural system. A service provider model is used in particular [16].

There are various challenges that must be correctly handled in order to propose an optimal solution for IoT with wearable implementation in healthcare for environmental and physiological monitoring. Due to important, ongoing monitoring, power consumption is a major issue for short-range communication in wearables and long-range data transfer from smartphones to the cloud. The most essential needs are wearability mode, flexibility in sensor adoption solutions, extension options and data collecting ranges, usability, and fusion. In medical care, the connection between physiological and environmental indicators requires data research and can only be proved by data analysis through adequate algorithm choices based on the continuous monitoring of a person.

3. Proposed System

A modular cellular data framework is presented to ensure preventive and workplace control of everyday practices in the healthcare IoT strategy. In order to track the ubiquitous criteria, portable handsets with comfortable wearability are used in comparison to the modular data collection strategy. To do this, a system has already been established for robust environmental parameter tracking. The solution is a mix of operating systems that specifically aims to offer functionality and design abilities for all the main facets of the platforms.

The following criteria and demands are discussed, given the high diversity of this area of study and the monitoring criteria focusing on physiological and psychological well-being and also on the effects of environmental parameters on

the ground [17]. We spoke earlier about the need for continuous and systematic data tracking and processing in preventive medicine. Therefore, wearables must be wearable and compact and lightweight with a comfortable style of wearability. Research into the field of measuring accurate information in the real world, in particular, shows that measuring settings require less user control. For these components, special attention must be given to product selection and prototype design. The proposed system architecture is shown in Figure 2.

The wearable devices remain prudently selected after the accessible device rendering follows the stated guidelines for physiological and psychological parameter tracking. The proposed system is designed to fulfill these criteria in the environmental surveillance sector as a handheld prototype. These three research areas will effectively contribute to cardiovascular and workplace health by integrating them. A thorough monitoring of environmental and physiological criteria is required in order to obtain comprehensive monitoring of healthcare. Pervasive surveillance involves a range of sensors that lead to scale, weight, and thus to wearability growth.

In order to ensure that wearability is conveniently preserved, the architecture techniques in combination with efficient software creation have to be consistent with the careful hardware concept, in this respect, 3D space usage for the system and efficient goods [18]. A thorough and personal collection of parameter settings, based on the specified parameter sets and the assessment by approved staff, is needed for the broadness of investigational scenarios within the area of preventive and occupational monitoring. It should not be restricted to specific case studies or topic classes.

There is also a strong need for a broad range of operative capabilities to combine multiple devices and sensor nodes from different vendors. The modular information gathering system is called an IoT portal. In this model, the access point is modular to adapt to external designs and devices in various fields to supplement the data solution where required [19]. The usage techniques for all data sources and, therefore, both data routes are taken into consideration in one framework. This hybrid data usability has a considerable effect on the structure of the device as data are distributed.

The incorporation of different systems and their individual configuration is essential to allow the flexibility of such information bases in response to research issues. Furthermore, the single mixture and setup also impact proper process control readiness, which requires a stronger device structure. The practical definition follows a dispersed device method that disassociates basic purposes from the domain cloud toward the separate IoT entry, such as information contribution, information gathering, and processing. Therefore, its functionality centers on data collection and information control, including customer-related tracking process setup.

The highest level includes cloud systems with the internal p2Health-Cloud system and cloud solutions from external providers as gateways into indirectly open results. The p2Health-Cloud allows user-specific research to be planned,

through the management of so-called measures, including data range, data delivery frequency, the appropriate formats, and so on. This configuration seeks unique sensor node solutions but sets the necessary parameters. The p2Health-Cloud is carried out in the field of data science, primarily by algorithms for the detection of potential major associations with the goal of elevating information synthesis and information analysis, and it takes into account data comparisons for both local and national reference databases [20].

The results will be given to the user and supervisor through a web interface where required. The second level is focused on personal mobile devices that act as IoT gateways and carry out measuring tasks. This involves connecting directly accessible data sources to the needed sensor node and indirectly accessible external cloud solutions, collecting and preparing data, and providing data for P2Health-Cloud as provided by the measurement procedures [21]. The collection of the appropriate data sources depends on the user's comprehensive monitoring of the registered sensor nodes and can be taken into account, if possible, as alternate nodes.

The gateway then becomes the data ability to concentrate on each inquiry and must adapt the data collection process to the collected data, transmitted data size, and power usage via a remote sensor node setup. At this step, the incorporation of external cloud solutions encompasses the information collected and is hence a major portion of the gateway message initiative [22]. The entry also handles wellbeing server decentralization and subordinate synthesis of data, including information synchronization, information configuration, and processing. The third stage contains the devices used by the handler in connection with the WBAN. At this stage, all entered data source classes are viewed while the implicitly available data sources react as flight recorders that do not provide any additional features.

The directly available sources, by comparison, also have network configurations that allow changing between various modes or allow comprehensive calculation and procedure structure configurations by way of allowing or average settings for warnings and data compression processing choices. This allows the measuring method to be optimized in relation to the requisite control and, if available, initial handling of information [23]. These sensor nodes most often include a level of the sensor that enables more sensors or modular substitution to be linked. This makes it easy to substitute or connect single sensor modules to the sensor node. Therefore, IoT devices may be customized for the relevant specific application with hardware.

IoT healthcare devices employ several sensors to monitor health data such as heart rate and ECG and to give healthcare recommendations, including telling users whether they should seek immediate medical attention and providing telehealth assistance [24]. The IoT gadget is powered by both the battery and the energy-harvesting module, which lets it locally complete compute jobs, remotely offload jobs, and store the remainder for later.

A newly generated sensed information from the IoT device over a particular time variable l for the data size $D_1^{(l)}$ would be handling the data received earlier of the buffer level

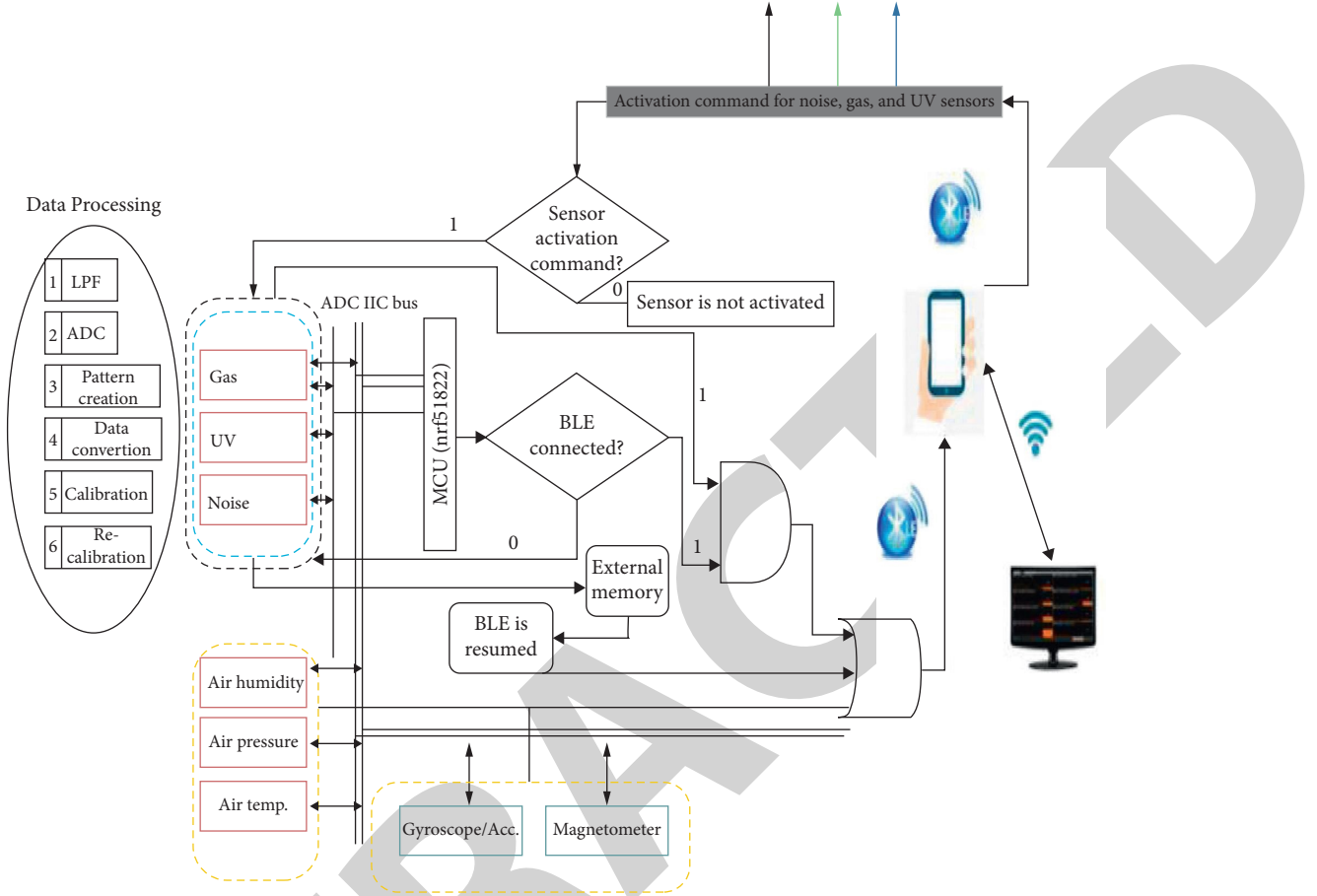


FIGURE 2: Proposed system architecture.

value $D_0^{(l)}$. For simplicity, we are just omitting the time variable l . The computational partition method and the overall computation tasks have been divided into the captured data with the capacity $D_1^{(l)} + D_0^{(l)}$ by M number of similar computational tasks [25]. The sensed information can be prioritized based upon the analytical model that can be represented as Y^l .

Let the scenario be considered. The edge computer is receiving $Y_1^{(l)}$ from the IoT device-generated computational task through a communication pipeline with a power gain value of $g^{(l)}$, that has been executed by the computer over the local module of $\gamma_0^{(l)}$ in a computational speed over g bps. Also, during the process of the forthcoming computational tasks, the buffering of the processed tasks is like $\{y_0^{(l)}, y_1^{(l)}\} \in \{m_0/M, m_1/M\}_{0 \leq m_0, m_1 \leq M}$. The communication channel's power value can be estimated through a model, that is,

$$P(g^{l+1} = n | g^l = m) = i_{nm} \quad \forall m, n \in F, \quad (1)$$

where F represents the collection of states present in the communication channel [26]. The total amount of energy that was consumed by the device has been represented as \mathfrak{N} for handling every sensed information received from the sensor devices. The flow model of the proposed architecture is shown in Figure 3.

There is the possibility that an attacker who is inquisitive about the user's privacy, including the user's location and use behaviour, might get the results of the calculation sent to the IoT device via the edge device. For the sake of location privacy and usage pattern analysis, an edge device can track the activity of the IoT device by offloading history depending on the channel state of the user at different distances [27]. The privacy level is connected to the amount of sensor data that needs to be handled and the rate at which it may be transferred.

Q is the present state's discounted long-term utility or Q function, which is utilized to select the offloading policy. The present state $t^{(l)}$ of the healthcare sensing data, radio channel status, anticipated renewable energy generated in the time slot, battery level of the IoT device, and computation history are taken into consideration while choosing the offloading policy [28]. This plan employs the channel model already established and simulates real-world events to help a person find the best course of action.

An IoT device analyzes the importance of healthcare data represented by $Y^{(l)}$ and calculates the network maximum power to the edge device $g^{(l)}$ while monitoring the data of size $D_1^{(l)}$ at the time slot l . Using this information, the IoT device calculates the harvested energy density σ^{-l} and watches the standard battery capacity of the IoT device $c^{(l)}$. Also, the current state has been represented by

$t^{(l)} = \{D_1^{(l)}, Y^{(l)}, g^{(l)}, \sigma^{-l}, D_0^{(l)}, c^l\}$. The complete set of states obtained can be collectively known as B .

To create an analogous number of calculation tasks, a fresh and buffered set of healthcare data, with size $D_1^{(l)} + D_0^{(l)}$ partitioned into M sets, is applied the computation division technique. We choose the transfer policy $y^{(l)} = [y_0^{(l)}, y_1^{(l)}] \in B$ to establish a compromise between exploration and exploitation by employing the greedy approach [29]. A further particular aspect of the offloading policy that optimizes $R(t^{(l)}, y)$ is that it is picked with a low probability compared to the other plausible offloading policies. It retains the remainder in the buffer, locally analyzes the data, and transfers one more of data $y_1^{(l)} (D_1^{(l)} + D_0^{(l)})$ to the IoT device.

The IoT device evaluates the relationship between the cost of sensed data and the amount of transferring information and also the current channel statuses to determine the privacy level that was reached once local processing is completed.

4. Results and Discussion

In customized healthcare, we aspire to track extended and ubiquitous criteria in comfortable wearability modes. We are working on identification and mitigation for the identification of earlier disease functions. Users, doctors, health criteria, and requirements were taken into consideration in the implementation of the solution. Thus, for a joint contribution to users and medicines, a scalable IoT portal has been introduced. In the one hand, medicinal products will identify functions, quantify parameters, talk to the consumer, track the data in real time, and adjust wearable to the appropriate research and therapeutic issues. The customers are not exclusively limited to particular vendors on the other hand.

The customer should choose solutions that are easy to adapt to programs. The data for transaction time are shown in Table 1. In reality, the handler ensures not have to be suitable in the workaround; nonetheless, instead, the IoT entry is compatible with the user. In specific, this approach can be used in professional conditions, particularly in communities where children are vulnerable to unsafe situations that could jeopardize their safety [30]. This involves miners, technicians/chemical workers, and heavy-duty manufacturing and construction workers/technicians. Physicians would have full access to a broad variety of parameters, which will lead to a wide variety of situations and important clinical correlations between parameters and diagnoses.

The transmission analysis is shown in Figure 4. Two sets of data are provided in support of the approach. The solution has been checked in the Life Science Automation Center in a chemical/analytical laboratory to ensure the reliability of measured data, real-time setup, sensor activation, server task specification, sufficient data transmission, and selection.

The subject was interested in a chemical reaction in a relatively noisy atmosphere while wearing Equivital, Fitbit, and Ubiquisense. The privacy level comparison and the energy consumption comparison are shown in Figures 5 and 6,

respectively. This test was conducted under all safety and durability legislations. Only a small amount of data obtained is displayed here because of the limited space.

Here, air humidity, heat, NO₂, and Ubiquisense noise, RR interval, breathing rate, pulse rates, and Equivital skin temperature were calculated. The data that are sent to the server demonstrate the sensitivity and correct functionality of the prototype for measuring, collecting, and transmitting data. The temperature analysis is shown in Figure 7.

The NO₂ limit aimed at the model is from 3 to 49.2 ppm, and the LODs are 0.4 ppm. Nitrous acerbic was physically applied to Cu in this experiment, resulting in a chemical process and a release of NO₂. This was achieved in several rounds of several dozens by the chemical technician. In high levels and a maximum reaction in low concentrates of 5 seconds, the reply and regeneration period of the device show the rapid response at once.

In this procedure, up to 21.8 ppm was detected in the spectrum of gas concentration. The sound level was simultaneously calculated. The utility comparison is shown in Figure 8. With 1 dB resolution, the system is limited from 32 to 85 dB.

The respiration analysis is shown in Figure 9. Noise during the test exceeded not more than 44 dB then, and in this respect, the usual noise was around 50 dB. These atmospheric conditions of the laboratory indicate air temperature and humidity. Vital signs were simultaneously tracked for the physiological parameters. The heart rate ranges between 74 bpm and 93 bpm. One of the essential indicators of a coordinated respiration rate followed a similar trend with a variable emotion amount and swing amount of 8 to 19 per minuscule. The measurement result of the R-wave peak, as the most important peak, is also calculated [31]. The time across every two peaks in RR is the time measured. From this parameter, the heart rate variability can be derived.

The temperature of the skin was eventually revealed. The data communication efficiency is shown in Figure 10. At the start of the procedure, the skin temperature of the technician is 26.3°C. At the conclusion of the test, the temperature has risen to 29.3°C. Both Fitbit parameters have been sent and synchronized to the server. They will be located in different server directories. The amount of information obtained below the constraint is shown by the various selection charges of the devices. During the 18-minute testing period, 2,250 noise samples (1.9 Hz), 0.985 moisture, and mid-air heat samples (1 Hz), and then 595 NO₂ (0.5 Hz) samples were obtained for environmental monitoring.

The performance of the proposed system is shown in Figure 11. Similarly, 72 breathing rates, cardio activity, skin temperature, and 1.450 RR interval samples (1.34 Hz) were obtained (every 15 seconds). We also submitted the solution to many practical experiments where all the functionalities of the solution are checked. Consequences of around 6 hours and 55 minutes stand summarized here for results and interpretation.

The computational latency comparison is shown in Figure 12. The doctors specified the tasks and accordingly adjusted the unit in this experiment. In order to best assess

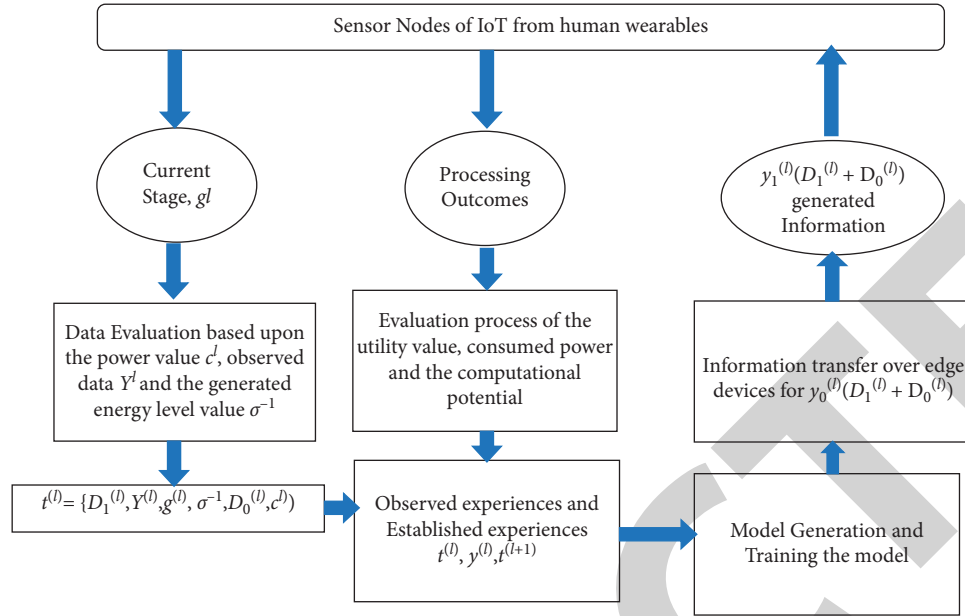
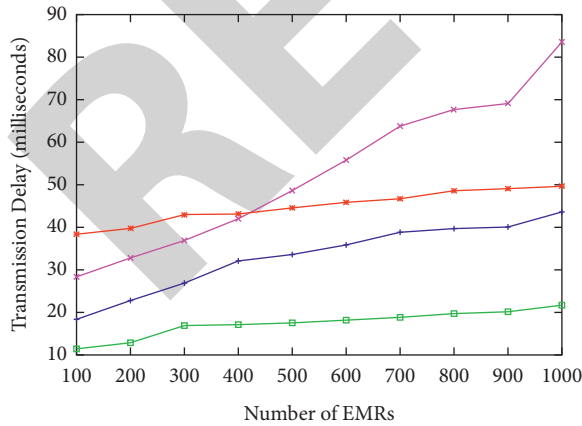


FIGURE 3: Flow model of the proposed architecture.

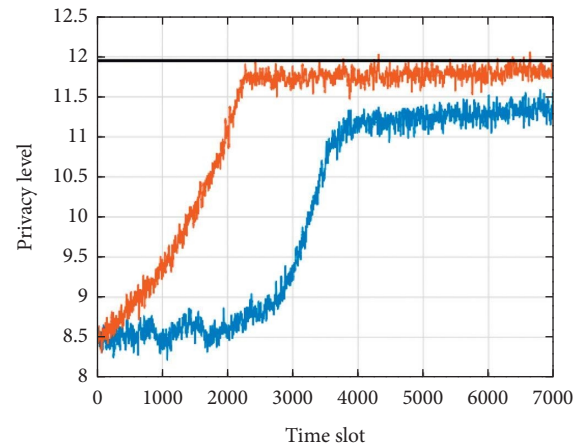
TABLE 1: Transaction time (ms).

No. of EMRs	Masood et al.	Menaka et al.	Omtosho et al.	Proposed approach
100	20.43	20.13	23.50	18.54
200	20.63	22.50	25.67	19.02
300	23.18	24.02	26.33	21.78
400	24.50	27.43	26.92	21.50
500	24.34	27.24	27.00	22.33
600	25.47	31.72	23.50	23.00
700	26.33	32.33	25.67	25.93
800	28.78	35.77	30.38	27.33
900	30.01	37.00	31.81	27.89
1000	32.50	39.98	32.62	29.72



— Menaka et. al. [26]
 — Omtosho et. al. [27]
 — Masood et. al. [20]
 — Proposed Approach

FIGURE 4: Transmission analysis.



— CMDP [17]
 — RL
 — Theoretical value

FIGURE 5: Privacy level comparison.

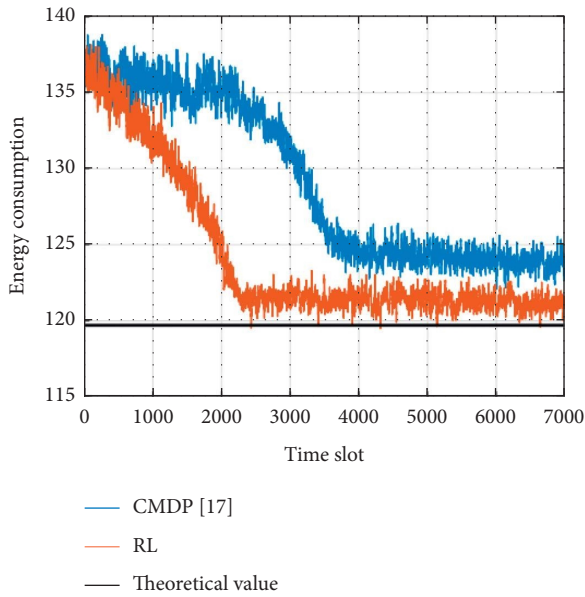


FIGURE 6: Energy consumption comparison.

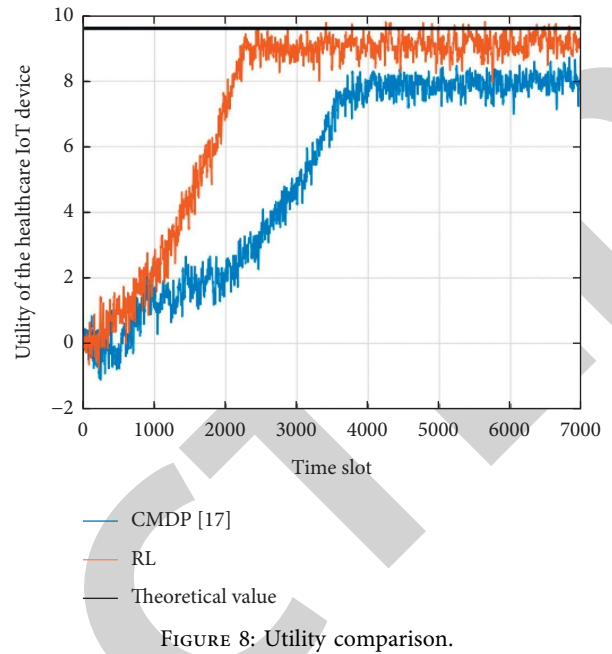


FIGURE 8: Utility comparison.

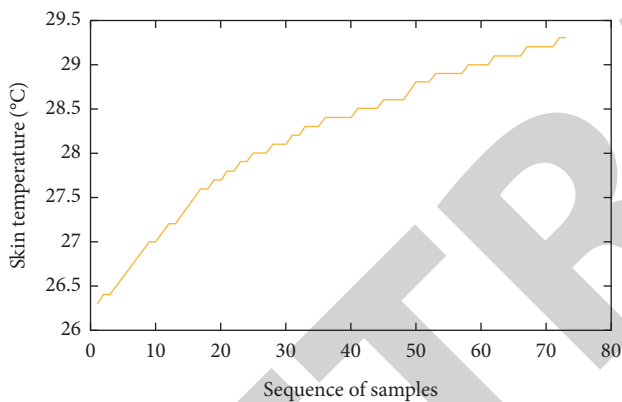


FIGURE 7: Temperature analysis.

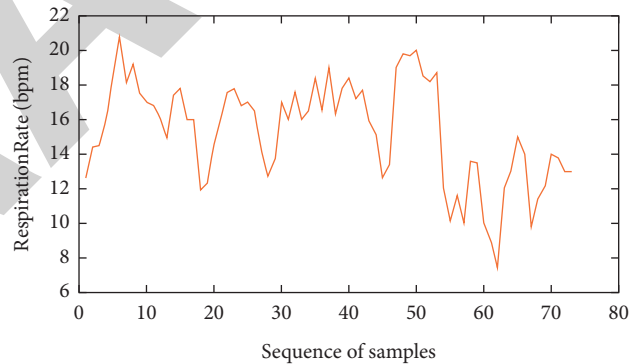


FIGURE 9: Respiration analysis.

the wearables, all measures of environmental-physiological activity have been setup to measure the largest possible spectrum. In summary, the customer before sleep was told to bring and act as his normal routing both goods and prototypes.

The findings show the user’s tracking while sleeping. The final QVR comparison is shown in Figure 13. This indicates that the customer has a natural state during his night. There are some noise bursts; however, the effects immediately vanish, nonetheless and this happens on an insufficient solitary period and, therefore, can originate significant meddling. Heart frequency, respiratory rate, and slumber levels are seen and then analyzed. While the parameter analysis indicates the normal set, the individual has failed to deeply sleep. This study, however, aims to encourage the assessment, definition, proof of principle, and application of practical usefulness, and approaches and techniques instead of medical inquiries.

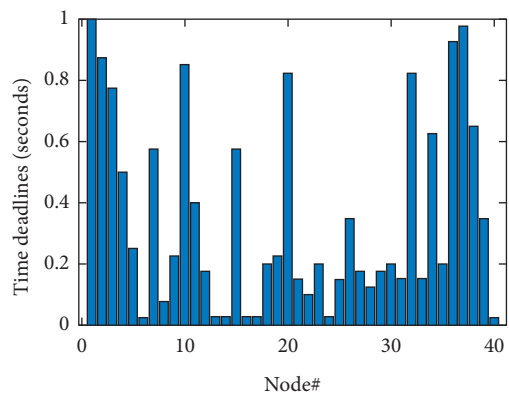


FIGURE 10: IoT data communication efficiency.

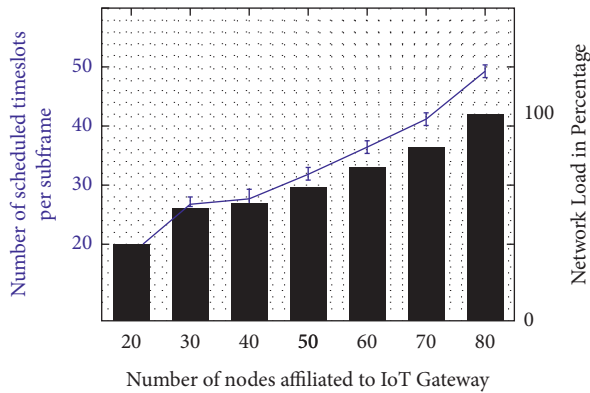


FIGURE 11: Performance of the proposed system.

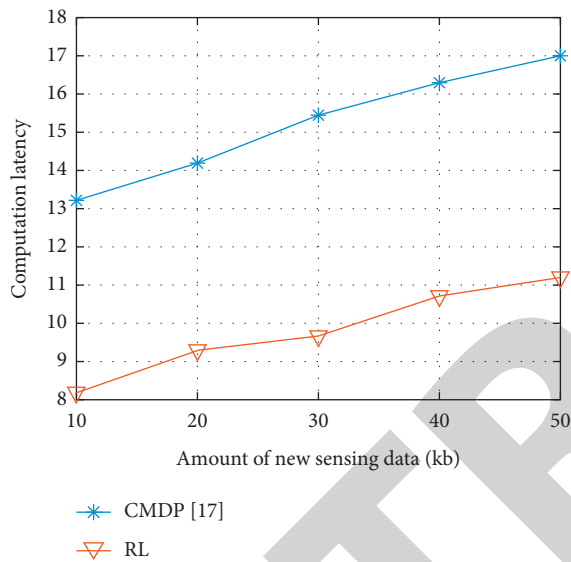


FIGURE 12: Computational latency comparison.

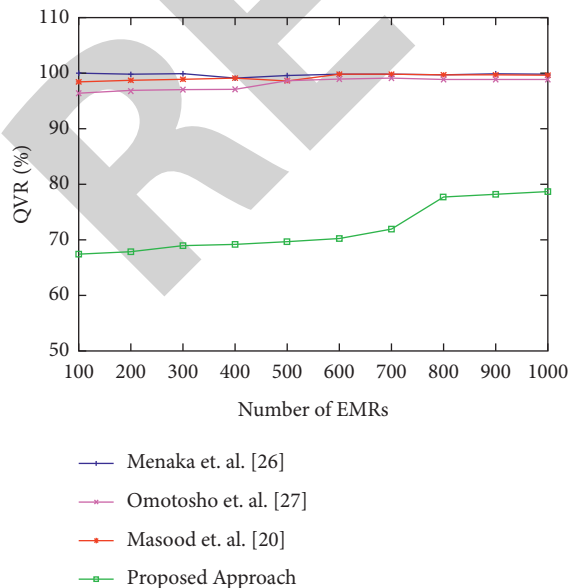


FIGURE 13: QVR comparison.

5. Conclusions

A health-monitoring IoT-powered healthcare device has been presented as a privacy-aware technique to estimate energy consumption, the privacy breach, and the IoT computation and energy expenditure in an offloading and local computing paradigm. This technique decides the offloading policy to use at the edge device from every time slot by considering the privacy level, energy usage, and processing delay. The suggested methodology incorporates a transfer learning technique, which is well known in the radio channel modeling community, and a learning architecture, both of which serve to enhance the system's learning capabilities for dynamic healthcare IoT systems. We implemented a robust, widespread, readily accessible, and easy-to-use IoT-based infrastructure for monitoring several types of environmental and physical parameters with the accuracy improvement in the transaction response of records by 26.75%, and the QVR value of the proposed structure reduced the inference and noise up to 70% in the framework. Additionally, we brought in a wrist-worn prototype with sensors that were only active in low-light environments. To support an effective end-to-end connection between the user and the physician, an IoT gateway, an intermediary hub between sensor nodes and servers, has been built. Instead of doing clinical studies, the focus was on the method validation, technological definition, and usability. In the future, the scope can include more parameters to the wearability model with the stripped antenna that can incorporate the framework for a better immune monitoring system in antiquated peoples.

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 there are no conflicts of interest.

References

- [1] O. Geman, I. Chiuchisan, I. Ungurean, M. Hagan, and M. Arif, "Ubiquitous healthcare system based on the sensors network and android internet of things gateway," in *Proceedings of the IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)*, pp. 1390–1395, Guangzhou, China, October 2018.
- [2] J. Botero-Valencia, L. Castano-Londono, D. Marquez-Viloria, and M. Rico-Garcia, "Data reduction in a low-cost environmental monitoring system based on lora for wsn," *IEEE Internet of Things Journal*, vol. 6, 2018.
- [3] M. Ekström, G. C. Riise, and H. A. Tanash, "Risk of cancer after lung transplantation for copd," *International Journal of Chronic Obstructive Pulmonary Disease*, vol. 12, 2017.
- [4] N. D. Elhayatmy and A. S. Ashour, "Internet of things based wireless body area network in healthcare," in *Internet of*

Retraction

Retracted: An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. A. AlZubi, S. Tiwari, K. Walia, J. M. Alanazi, F. I. AlZobi, and R. Verma, "An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2933015, 9 pages, 2022.

Research Article

An Efficient Stacked Deep Transfer Learning Model for Automated Diagnosis of Lyme Disease

Ahmad Ali AlZubi ¹, Shailendra Tiwari ², Kuldeep Walia ³, Jazem Mutared Alanazi ¹,
Firas Ibrahim AlZobi ⁴ and Rohit Verma⁵

¹Computer Science Department, Community College, King Saud University, Riyadh, Saudi Arabia

²Department of Computer Science and Engineering, Thapar Institute of Engineering & Technology, Patiala, India

³Sri Guru Granth Sahib World University, Fatehgarh Sahib, Punjab, India

⁴Department of Information Systems and Networks, Faculty of Information Technology,
The World Islamic Sciences & Education University, Amman, Jordan

⁵School of Computing, National College of Ireland, Dublin, Ireland

Correspondence should be addressed to Ahmad Ali AlZubi; aalzubi@ksu.edu.sa

Received 2 January 2022; Revised 26 January 2022; Accepted 7 February 2022; Published 28 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Ahmad Ali AlZubi 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.

Lyme disease is one of the most common vector-borne infections. It typically causes cardiac illnesses, neurologic illnesses, musculoskeletal disorders, and dermatologic conditions. However, most of the time, it is poorly diagnosed due to many similarities with other diseases such as drug rash. Given the potentially serious consequences of unnecessary antimicrobial treatments, it is essential to understand frequent and uncommon diagnoses that explain symptoms in this population. Recently, deep learning models have been used for the diagnosis of various rash-related diseases. However, these models suffer from overfitting and color variation problems. To overcome these problems, an efficient stacked deep transfer learning model is proposed that can efficiently distinguish between patients infected with Lyme (+) or infected with other infections. 2nd order edge-based color constancy is used as a preprocessing approach to reduce the impact of multisource light from images acquired under different setups. The AlexNet pretrained learning model is used for building the Lyme disease diagnosis model. To prevent overfitting, data augmentation techniques are also used to augment the dataset. In addition, 5-fold cross-validation is also used. Comparative analysis indicates that the proposed model outperforms the existing models in terms of accuracy, f-measure, sensitivity, specificity, and area under the curve.

1. Introduction

Lyme disease is one of the most common vector-borne infections, generally due to one of the three pathogenic genospecies of the spirochete *Borrelia* [1, 2]. It typically causes cardiac illnesses, neurologic illnesses, musculoskeletal disorders, dermatologic conditions, etc. [3, 4]. However, most of the time, it is poorly diagnosed due to many similarities with other diseases such as drug rash [5], pityriasis rosea rash [6], and ringworm [7]. Figure 1 shows the example of Lyme disease along with other similar diseases. It is clearly found that the drug rash, pityriasis rosea rash, and ringworm visually seem to be similar and so

many times, Lyme disease is either underdiagnosed or overdiagnosed.

Overdiagnosis or underdiagnosis of Lyme disease leads to unnecessary antibiotic treatments. Numerous problems and adverse consequences because of medicines being given to patients longer than recommended, needless antibiotics, or unusual therapies for Lyme disease were reported, like cholecystitis, catheter-associated bloodstream infection, clots from venous catheters, *Clostridioides difficile* infections, and death. Given the potentially serious consequences of unnecessary antimicrobial treatments, it is essential to understand frequent and uncommon diagnoses that explain symptoms in this population [2–4].



FIGURE 1: Images of Lyme disease along with other similar diseases.

Therefore, it is required to build such a framework or model which can clearly distinguish among patients infected with Lyme (+) or infected with other infections. Recently, the imaging dataset of Lyme patients has been published on the Kaggle website [8]. Many researchers have utilized it for distinguishing between Lyme infections and infections with other rashes using deep learning models. However, these models suffer from the overfitting problem. Also, suitable preprocessing techniques are required to improve the quality of images under consideration to achieve efficient results. To overcome these problems, in this paper, an efficient stacked deep transfer learning model is proposed to classify Lyme patients.

The main contributions of this paper are as follows:

- (a) An efficient stacked deep transfer learning model is proposed to classify Lyme patients
- (b) 2nd order edge-based color constancy is used as a preprocessing approach to reduce the impact of multisource light from images acquired under different setups
- (c) The AlexNet pretrained learning model is used
- (d) Data augmentation techniques are also used to augment the dataset

The remaining paper is organized as follows: Section 2 presents the literature work, Section 3 presents the proposed model, Section 4 discusses various experimental results, and Section 5 concludes the paper.

2. Related Work

In [9], an ensemble deep learning pipeline (EDLP) was designed by using the 34-layer ResNet model. ResNet was used to extract the features from the limited skin disease dataset. Eleven skin conditions were classified. It has achieved 91.7% precision and 92.55% recall, respectively. In [10], skin cancer was evaluated from the rashes. A convolutional neural network (CNN) was utilized to predict images of rashes or skin cancer. It has achieved an accuracy of 80.2% for 20 epochs. In [11], a mobile-enabled expert system named i-Rash was designed for the diagnosis of inflammatory skin lesions. It can predict the given image as psoriasis, eczema, acne, and healthy. i-Rash was trained using pretrained SqueezeNet.

In [12], a lightweight attention-based deep learning model (LWADL) was designed to predict eleven skin diseases. LWADL achieved better accuracy as compared to VGG19, VGG16, ResNet50, and InceptionV3. In [13], a UNet-based dense CNN (UNet-dCNN) model was designed. MobileNetV2 was also used to achieve better results. It was utilized for histopathological image-based skin cancer diagnosis. It has shown an average accuracy of 87.7%. In [14], a multitask deep learning model was designed. It was utilized for automatic analysis and classification of skin lesions. A focal loss and a Jaccard distance-based loss function were designed. A three-phase joint training approach was used to assure significant feature learning.

In [15], a pretrained ResNet-50 model was utilized for the classification of melanoma. Ensemble learning was also considered to obtain better results. In [16], a fully automated deep ensemble model (FADEM) was designed for lesion boundary segmentation. DeeplabV3+ and Mask R-CNN were ensembled and tested on the ISIC-2017 dataset. FADEM has achieved specificity and sensitivity of 97.94% and 89.93%, respectively. In [17], ensembling of CNN models was achieved by integrating with a test-time regularly spaced shifting. The proposed ensemble model was used for the classification of skin lesions. In [18], an adaptive dual attention module-based CNN model was designed for the segmentation of skin lesions. The proposed model has shown better performance during the computation of potential features in classifying the skin lesions' boundary.

From the existing literature, it is found that the existing models suffer from overfitting and hyperparameter tuning problems. In addition, preprocessing techniques are required to reduce the impact of multiple light sources on the images. Also, none of the researchers have focused on the diagnosis of Lyme disease. Therefore, in this paper, an efficient model is designed to achieve better results.

3. Proposed Methodology

This section discusses the proposed methodology. Initially, the obtained images are improved by using the 2nd order edge color constancy. Thereafter, a ResNet-based model is trained to achieve better results [18]. Figure 2 shows the proposed automated Lyme disease diagnosis model. It clearly shows that the proposed model is decomposed into three phases, i.e., data augmentation, 2nd order edge-based color constancy, and an AlexNet-based pretrained model for extracting the features which are used for the classification using fully connected layers. To achieve regularization and to prevent overfitting, dropouts are also used. Binary cross-entropy is used as a loss function.

3.1. Color Constancy. Color constancy has the ability to restore the impact infected multiple light sources. Thus, the obtained images are independent of colors of the light source. In this paper, a 2nd order-edge based color constancy approach is used. It states that the distribution of color derivatives exhibits the principal dissimilarity in the direction of a light source [18]. Minkowski's norm is then applied to the computed derivatives to predict the direction of a light source [19, 20]. A step-by-step algorithm for the 2nd order edge-based color constancy is presented in Algorithm 1.

3.2. Proposed ResNet Model. Residual network (ResNet) is a well-known pretrained model used as a backbone for classifying many imaging datasets. It allowed us to successfully build an enormously deep model with more than 150 layers. Before ResNet, the existing models suffered from vanishing gradients whenever we tried to train them deeply. It has achieved better results with the help of skip connections. It prevents the vanishing gradient problem by using a substitute shortcut route for the gradient to flow

through. It allows the model to build an identity function that assures the topmost layer will achieve better performance, the same as the lower layer.

Figure 3 shows the ResNet model in the paper. Initially, the images obtained from the color constancy model are utilized for building the trained model. It utilizes various convolution layers, followed by ReLU, normalization, and pooling operations. After using 5 convolution layers, a fully connected layer is utilized along with ReLU and dropouts. Finally, after using three fully connected layers, the softmax function is used to obtain the results.

4. Experimental Analysis

The experiments for the proposed model are performed on the online MATLAB 2021a using a benchmark Kaggle dataset. Comparisons are also performed by considering the competitive models. In addition, we have also validated the proposed model with and without considering the 2nd order edge-based color constancy. Table 1 demonstrates the hyperparameter setting of the proposed model.

4.1. Dataset. In this paper, the Lyme disease (Silent Epidemic) dataset [8] obtained from Kaggle is used for experimental purposes. It composes images of the erythema migrans, referred to as bull's eye rash. It is one of the utmost protuberant signs of Lyme disease. The dataset also includes other kinds of rashes that may be often confused with Lyme disease by medical staff. For training, there are 206 Lyme (-ve) and 151 Lyme (+ve) images available. For testing, there are 51 Lyme (-ve) and 36 Lyme (+ve) images available.

Therefore, data augmentation is used to augment the dataset. Since the obtained images were captured using different machines under different light sources, using these images directly for diagnosis may result in poor performance of the model. Therefore, in this paper, to prevent the effect of multiple light sources, color constancy is used. It can restore the impact of color light sources from the images to achieve better performance of the models.

4.2. Training and Validation Analysis. Figure 4 shows the training and validation analysis of the proposed model without the use of 2nd order edge-based color constancy. It clearly shows that the proposed model has achieved 95.71% validation accuracy. But also, it is found that the proposed model suffers from the overfitting issue since the training accuracy is 100%. Therefore, still, there is room for improvement in it.

Figure 5 shows the training and validation accuracy of the proposed model with 2nd order edge-based color constancy. It is found that the proposed model with color constancy achieves 98.69% validation accuracy. Therefore, the proposed model is least affected by the overfitting problem. Besides, the validation accuracy of the proposed model has shown better convergence speed than in the results shown in Figure 4.

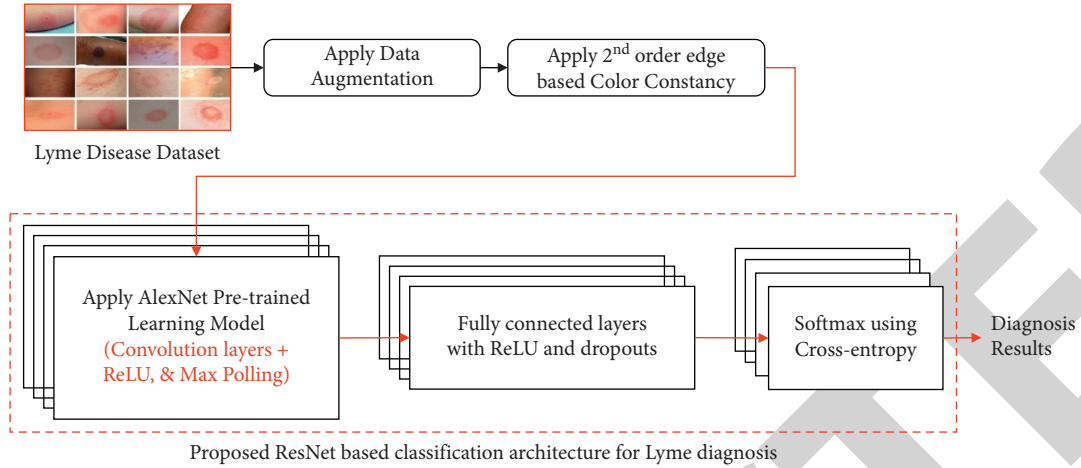


FIGURE 2: The proposed automated Lyme disease diagnosis model.

Begin

Input: Color image (C_I) having size M and N Minkowski's norm (m_n) = 5 Sigma (S_g) = 2 differential order(d_o) = 2.

Step 1: Divide C_I into R_I , G_I , and B_I .

Step 2: Remove saturated color points. It represents those pixels that are greatly influenced by the light direction [21, 22].

Step 3: Computed aggregated color values as [23]:

$$R_T = \sum \sum (R_I) \quad T_{R=} \sum \sum (I_R) \quad T_{R=} \sum \sum (I_R)$$

$$G_T = \sum \sum (G_I)$$

$$B_T = \sum \sum (B_I)$$

Here, R_T , G_T , and B_T define the aggregated color values of R_I , G_I , and B_I , respectively.

Step 4: Compute the average of all color channels as [24]:

$$A_C = R_T, G_T, B_T / 3 * M * N$$

Step 5: Compute the impact of color saturation as [25, 26]:

$$R_A = A_C / \text{mean}(R_T)$$

$$G_A = A_C / \text{mean}(G_T)$$

$$B_A = A_C / \text{mean}(B_T)$$

Step 6: Remove the color saturation as [18, 27]:

$$S_R = R_A * R_I$$

$$S_G = G_A * G_I$$

$$S_B = B_A * B_I$$

Step 7: Evaluate the effect illuminance (C_p) using 2nd order edge-based approach as:

$$C_p = (C_I)^{m_n}$$

Step 8: Evaluate the impact of light on each color channel as [18, 28]:

$$N_R = \sum \sum (C_p(:, :, 1) * m_R)^{1/m_n}$$

$$N_G = \sum \sum (C_p(:, :, 2) * m_G)^{1/m_n}$$

$$N_B = \sum \sum (C_p(:, :, 3) * m_B)^{1/m_n}$$

Here, m_R , m_G , and m_B show the mask containing the saturated pixels.

Step 9: Evaluate the aggregated impact of normalized whiteness in the color channels as [18, 29]:

$$I_L = \sqrt{N_R^2 + N_G^2 + N_B^2}$$

Step 10: Compute the impact of light source as:

$$S_R = S_R / I_L$$

$$S_G = S_G / I_L$$

$$S_B = S_B / I_L$$

Step 11: Compute the restored color channels as

$$R_O = (R_I / S_R * \sqrt{3})$$

$$G_O = (G_I / S_G * \sqrt{3})$$

$$B_O = (B_I / S_B * \sqrt{3})$$

Step 12: Return concatenated (R_O , G_O , and B_O)

ALGORITHM 1: Second-order edge-based color constancy algorithm.

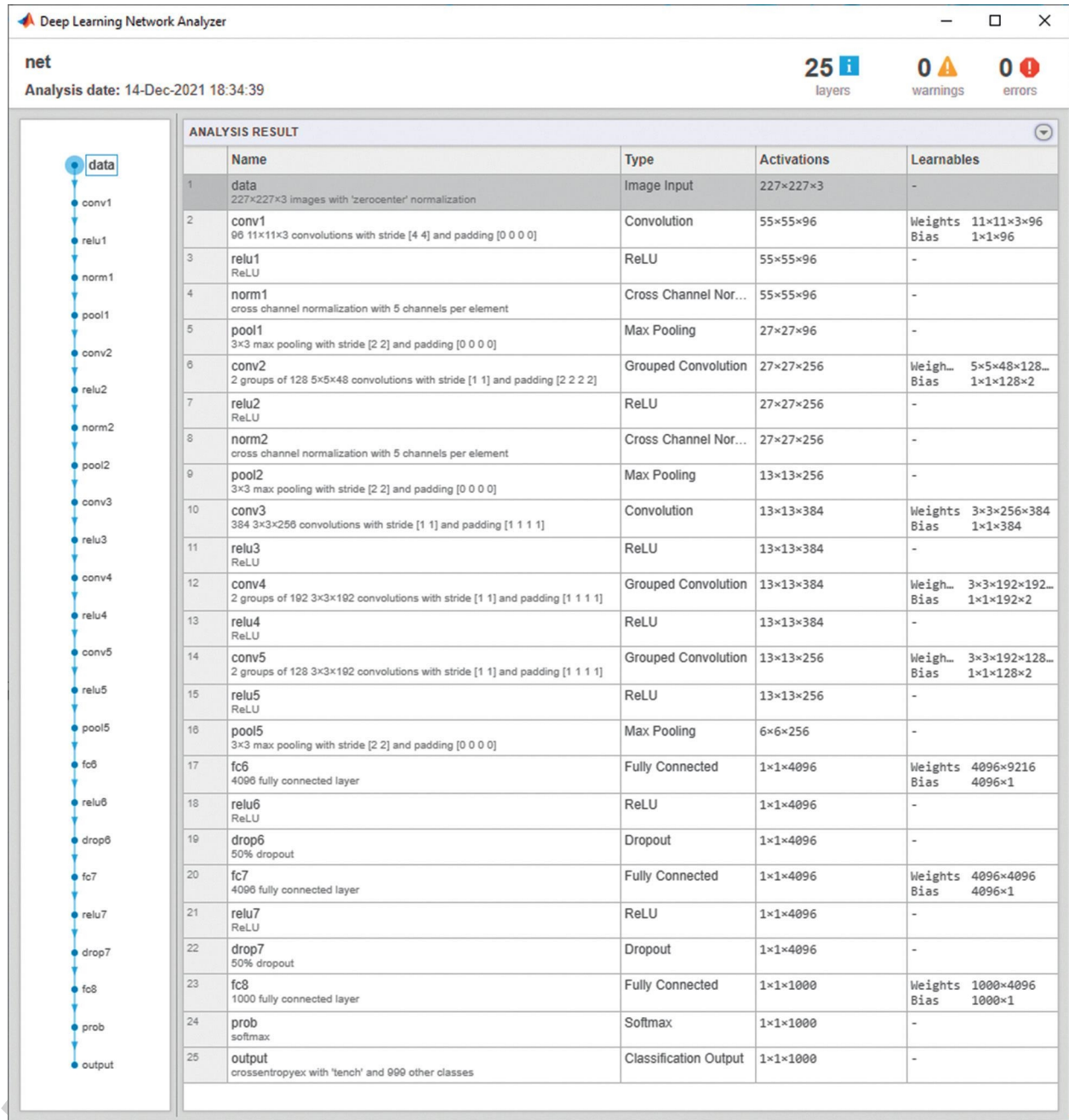


FIGURE 3: The proposed ResNet-based classification architecture for Lyme diagnosis.

TABLE 1: Hyperparameter setting of the proposed model.

Parameter	Value/type
Gradient decay factor	0.9
Squared gradient decay factor	0.99
Epsilon	1.00E-08
Initial learning rate	3.00E-04
Learning rate drop factor	0.1
Learning rate drop period	10
L2 regularization	1.00E-04
Gradient threshold	L2-norm
Maximum epochs	200
Minimum batch size	64
Validation frequency	50

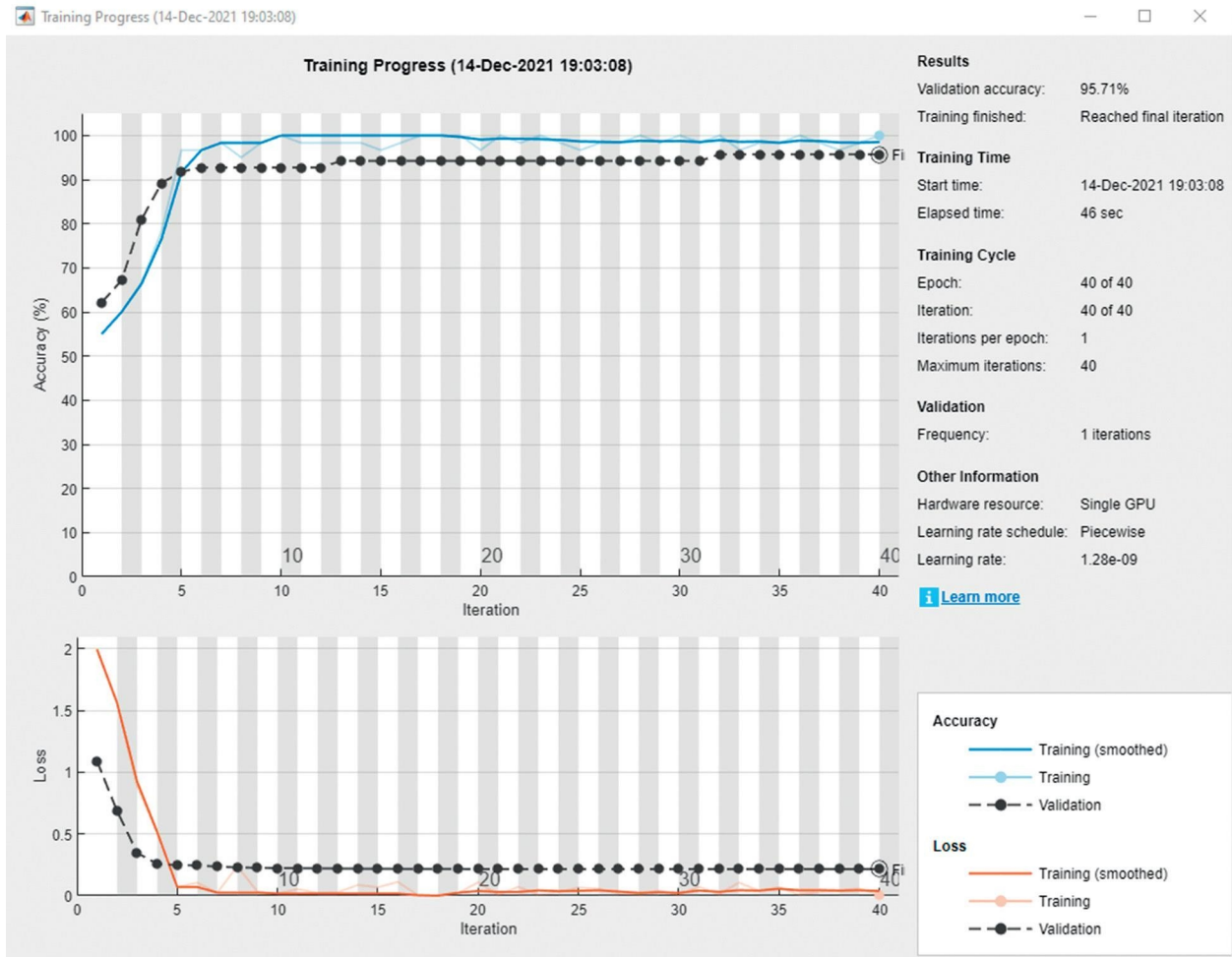


FIGURE 4: Training and validation analysis of the proposed model without using the 2nd order edge-based color constancy.

4.3. Confusion Matrix Analysis. Confusion matrix analyses are widely accepted to compute the performance of the various classification models. It utilizes the concepts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) that can be computed using the actual and predicted classes.

Figure 6 shows the confusion matrix of the proposed model without the use of 2nd order edge-based color constancy. It is found that without the use of color constancy, the proposed model achieves an average accuracy of 95.7%. It is found that this model has achieved better results, but with a 4.3% error rate.

Figure 7 demonstrates the confusion matrix analysis of the proposed model with 2nd order edge-based color constancy. It is found that the proposed model has achieved 98.7% accuracy, which is 2.0% better than the proposed model without the use of 2nd order edge-based color constancy (EBCC). There are only 7 cases that are poorly classified by the proposed model, compared to 23 in the model with the use of 2nd order edge-based color constancy.

4.4. Discussion. This section presents the discussion of the proposed and competitive models when they are applied to the Lyme disease dataset. The hyperparameters of the competitive models are obtained from their published papers. These competitive models are CNN [10], EDLP [9], SqueezeNet [11], LWADL [12], Unet-dCNN [13], ResNet-50 [15], FADEM [16], and Ensemble CNN [17]. Table 2 demonstrates the testing analysis of the proposed and competitive models. It is found that the proposed model achieves remarkably better performance than the competitive models. Bold values represent the better-performing model. The proposed model has achieved an average improvement in terms of accuracy, f-measure, sensitivity, specificity, and area under the curve (AUC) as 2.9787%, 2.7891%, 3.0875%, 2.1578%, and 2.1579%, respectively.

It is found that the proposed model achieves remarkable performance for Lyme disease. Therefore, the proposed model can be used for real-time diagnosis of Lyme disease and can help doctors treat patients with the correct medicines.

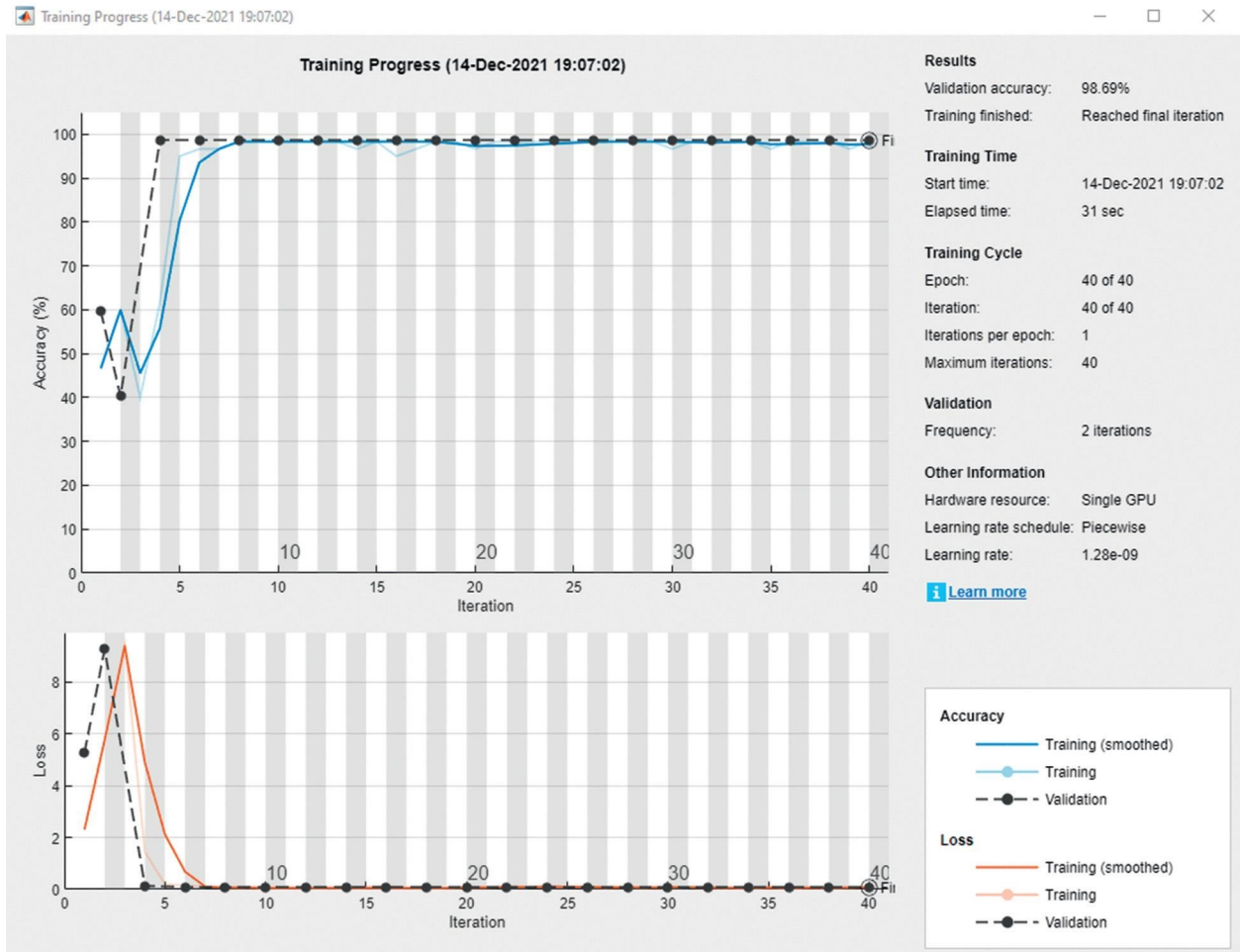


FIGURE 5: Training and validation analysis of the proposed model with 2nd order edge-based color constancy.

Output Class	Lyme (+ve)	197 36.8%	19 3.5%	91.2% 8.8%
	Lyme (-ve)	4 0.7%	316 59.0%	98.8% 1.2%
		98.0% 2.0%	94.3% 5.7%	95.7% 4.3%
		Lyme (+ve)	Lyme (-ve)	
		Target Class		

FIGURE 6: Confusion matrix analysis of the proposed model without 2nd order edge-based color constancy.

Output Class	Lyme (+ve)	213 39.7%	3 0.6%	98.6% 1.4%
	Lyme (-ve)	4 0.7%	316 59.0%	98.8% 1.2%
		98.0% 1.8%	99.1% 0.9%	98.7% 1.3%
		Lyme (+ve)	Lyme (-ve)	Target Class

FIGURE 7: Confusion matrix analysis of the proposed model with 2nd order edge-based color constancy.

TABLE 2: Training analyses among the proposed and competitive Lyme rash classification models.

Model	TP	FP	TN	FN	Accuracy	f-measure	Sensitivity	Specificity	AUC
CNN [10]	196	20	294	26	90.4306	91.5032	87.9069	93.3333	91.0679
EDLP [9]	192	24	303	17	89.6551	94.3708	92.4444	92.2330	92.3221
SqueezeNet [11]	194	22	294	26	89.3719	92.1921	87.6777	93.3130	91.1111
LWADL [12]	210	6	295	25	97.0731	92.4471	88.8392	98.0769	94.2164
Unet-dCNN [13]	184	32	297	23	86.0262	92.6282	89.5454	90.0311	89.8336
ResNet-50 [15]	202	14	288	32	93.1372	89.8089	85.5855	95.2702	91.1196
FADEM [16]	195	21	292	28	90.1408	91.3580	87.2727	93.3753	90.8752
Ensemble CNN [17]	195	21	310	10	90.7894	96.633	95.3917	93.1818	94.0952
Proposed without EBCC	197	19	316	4	90.5940	98.6486	97.8609	93.8906	95.3815
Proposed with EBCC	213	3	316	4	98.6111	98.6711	98.1566	99.0000	98.6460

5. Conclusion

This paper proposes an efficient stacked deep transfer learning model that can efficiently distinguish between patients infected with Lyme (+) or infected with other infections. 2nd order edge-based color constancy was used as a preprocessing approach to reduce the impact of multisource light from images acquired under different setups. The AlexNet pretrained learning model was utilized for building the Lyme disease diagnosis model. Data augmentation techniques were also used to augment the dataset. Extensive comparative analyses have shown that the proposed model outperforms the competitive models in terms of accuracy, f-measure, sensitivity, specificity, and AUC of 2.9787%, 2.7891%, 3.0875%, 2.1578%, and 2.1579%, respectively.

Data Availability

The used dataset is freely available at <https://www.kaggle.com/sshikamaru/lyme-disease-rashes>.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the study.

Acknowledgments

This work was supported by the Researchers Supporting Project (No. RSP-2021/395), King Saud University, Riyadh, Saudi Arabia.

References

- [1] T. Kobayashi, Y. Higgins, M. T. Melia, and P. G. Auwaerter, "Mistaken identity: many diagnoses are frequently misattributed to Lyme disease," *The American Journal of Medicine*, 2021.
- [2] R. Patel, K. L. Grogg, W. D. Edwards, A. J. Wright, and N. M. Schwenk, "Death from inappropriate therapy for Lyme disease," *Clinical Infectious Diseases*, vol. 31, no. 4, pp. 1107–1109, 2000.
- [3] N. S. Marzec, C. Nelson, P. R. Waldron et al., "Serious bacterial infections acquired during treatment of patients given a diagnosis of chronic Lyme disease - United States," *MMWR. Morbidity and mortality weekly report*, vol. 66, no. 23, pp. 607–609, 2017.
- [4] M. C. Reid, R. T. Schoen, J. Evans, J. C. Rosenberg, and R. I. Horwitz, "The consequences of overdiagnosis and overtreatment of Lyme disease: an observational study," *Annals of Internal Medicine*, vol. 128, no. 5, pp. 354–362, 1998.

Retraction

Retracted: Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] F. W. Alsaade and M. S. Alzahrani, "Classification and Detection of autism Spectrum Disorder Based on Deep Learning Algorithms," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8709145, 10 pages, 2022.

Research Article

Classification and Detection of Autism Spectrum Disorder Based on Deep Learning Algorithms

Fawaz Waselallah Alsaade  and Mohammed Saeed Alzahrani 

College of Computer Science and Information Technology, King Faisal University, P.O. Box 4000, Al-Ahsa, Saudi Arabia

Correspondence should be addressed to Fawaz Waselallah Alsaade; falsaade@kfu.edu.sa

Received 1 January 2022; Revised 14 January 2022; Accepted 7 February 2022; Published 28 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Fawaz Waselallah Alsaade and Mohammed Saeed Alzahrani. 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.

Autism spectrum disorder (ASD) is a type of mental illness that can be detected by using social media data and biomedical images. Autism spectrum disorder (ASD) is a neurological disease correlated with brain growth that later impacts the physical impression of the face. Children with ASD have dissimilar facial landmarks, which set them noticeably apart from typically developed (TD) children. Novelty of the proposed research is to design a system that is based on autism spectrum disorder detection on social media and face recognition. To identify such landmarks, deep learning techniques may be used, but they require a precise technology for extracting and producing the proper patterns of the face features. This study assists communities and psychiatrists in experimentally detecting autism based on facial features, by using an uncomplicated web application based on a deep learning system, that is, a convolutional neural network with transfer learning and the flask framework. Xception, Visual Geometry Group Network (VGG19), and NASNETMobile are the pretrained models that were used for the classification task. The dataset that was used to test these models was collected from the Kaggle platform and consisted of 2,940 face images. Standard evaluation metrics such as accuracy, specificity, and sensitivity were used to evaluate the results of the three deep learning models. The Xception model achieved the highest accuracy result of 91%, followed by VGG19 (80%) and NASNETMobile (78%).

1. Introduction

Autism spectrum disorders (ASD) refer to a group of complex neurodevelopmental disorders of the brain such as autism, childhood disintegrative disorders, and Asperger's syndrome, which, as the term "spectrum" implies, have a wide range of symptoms and levels of severity [1]. These disorders are currently included in the International Statistical Classification of Diseases and Related Health Problems under Mental and Behavioral Disorders, in the category of Pervasive Developmental Disorders [2]. The earliest symptoms of ASD often appear within the first year of life [3–6] and may include lack of eye contact, lack of response to name calling, and indifference to caregivers. A small number of children appear to develop normally in the first year, and then show signs of autism between 18 and 24 months of age [5], including limited and repetitive patterns of behavior, a narrow range of interests and activities, and weak language

skills. As these disorders also affect how a person perceives and socializes with others, children may suddenly become introverted or aggressive in the first five years of life as they experience difficulties in interacting and communicating with society. While ASD appears in childhood, it tends to persist into adolescence and adulthood [7].

Advanced information technology that uses artificial intelligence (AI) models has helped to diagnose ASD early through facial pattern recognition. Yolcu et al. [8] used the convolutional neural network (CNN) algorithm to train data for extracting components of human facial expressions and proposed the use of such algorithm to detect facial expressions in many neurological disorders [9]. In 2018, Haque and Valles [10], using deep learning approaches, updated the Facial Expression Recognition 2013 dataset to recognize facial expressions of autistic children. Rudovic et al. [11] presented the CultureNet deep learning model that was used to identify 30 videos.

Several studies have discovered important features of autism in numerous ways to detect autism, such as feature extraction [12], eye tracking [13], facial recognition, medical image analysis [9], and voice recognition [14]. However, facial recognition plays a more significant role in detecting autism than the person's emotional state. Facial recognition is a common way to identify persons and to prove that they are normal or abnormal. It involves mining pertinent information to disclose behavior patterns [15, 16]. Duda et al. [17] introduced a new method of producing samples of differentials between autism and attention deficit hyperactivity disorder (ADHD) and using such differentials to recognize autism. Sixty five samples of differentials on facial expressions of social responsiveness were collected for the datasets. Deshpande et al. [18] designed metrics for studying brain activities to identify autism. AI and soft computing approaches have also been applied to detect autism. Different studies have been conducted on detecting autism, but few of them have focused on brain neuroimaging. Parikh et al. [19] used machine learning approaches to develop a system for extracting the characteristics of autism. They gathered their dataset from 851 persons whom they classified as having and not having ASD. Thabtah and Peebles [20] used rule-based machine learning (RBML) to detect ASD traits. Al Banna et al. [21] developed a smart system for monitoring ASD patients during the coronavirus disease 2019 (COVID-19) pandemic.

Many real-life applications of AI and machine algorithms have been designed to help solve social problems. AI has been used in all health applications to help doctors control diseases such as autism. Artificial neural networks have been particularly focused on as a way to extract features of ASD patients that can be used to discriminate between persons with and without ASD. Among the techniques that have been used or proposed to detect autism in children are deep learning techniques, particularly, CNNs and recurrent neural networks (RNNs) [22, 23], and the bidirectional long short-term memory (BLSTM) model [24]. Lately, more studies have been conducted to diagnose ASD using machine learning approaches such as [24, 25], brain imaging [26–28], analysis of data on physical biomarkers [29–33], assessment of the behavior of persons with autism, and assessment of clinical data using the machine learning approach [33].

Our study demonstrated the use of a well-trained classification model (based on transfer learning) to detect autism from an image of a child. With the advent of high-specification mobile devices, this model can readily provide a diagnostic test of putative autistic traits by taking an image with cameras. The main contributions of our research are as follows:

- (i) Three pretrained deep learning algorithms were applied for ASD detection: NASNETMobile, Xception, and VGG19
- (ii) The Xception model showed the best performance of the three pretrained deep learning algorithms
- (iii) A system was designed to help health officials to detect ASD through eye and face identification

- (iv) The developing system has been validated and examined using various methods.

The rest of this paper is organized as follows. Section 2 describes materials and methods of pretrained deep learning models. Section 3 provides experiments. Section 4 and 5 provide results and discussion. The conclusion of papers is presented in Section 6.

2. Materials and Methods

This study proposes a deep learning model based on transfer learning, namely, Xception, NASNETMobile, and VGG1 9 to detect autism using facial features of autistic and normal children. Facial features can be used to determine if a child has autism or is normal. The models extracted significant facial features from the images. One of the advantages offered by deep learning algorithms is the ability to extract very small details from an image, which a person cannot notice with the naked eye. Figure 1 shows the framework of our study, from the data acquisition to the data preprocessing and loading, to the model preparation and training, and to the model performance test.

2.1. Dataset. This study analyzed facial images of autistic children and normal children obtained from Kaggle platform, which is publicly accessible online [34]. The dataset consisted of 2,940 face images, half of which were of autistic children and the other half were of nonautistic children. This dataset was collected through Internet sources such as websites and Facebook pages that are interested in autism. Table 1 shows the distribution of the split dataset samples. The splitting of the input data is presented in Figure 2.

2.2. Preprocessing. The purpose of the data preprocessing was to clean and crop the images. Because the data were collected from Internet resources by Piosenka [34], they had to be preprocessed before they could be used to train the deep learning model. The dataset creator automatically cropped the face from the original image. Then, the dataset was split into 2,540 images for training, 100 for validation, and 300 for testing. To scaling, the normalization method was applied; the dataset was rescaling the parameters of all the images from the pixel values [0, 255] to [0, 1].

2.3. Convolutional Neural Network Models. AI has been remarkably developed to assist humans in their daily life, for example, through medical applications, which are based on a branch of AI called “computer vision.” Hence, the CNN algorithm has contributed to the detection of diseases and to behavioral and psychological analysis.

2.3.1. Basic Components of the CNN Model. The convolutional neural network (CNN) is one of the most famous deep learning algorithms. It takes the input image and assigns importance to learnable weights and biases in order to recognize the class of the image. The neuron can be said to be a simulation of the communication pattern of the neurons of

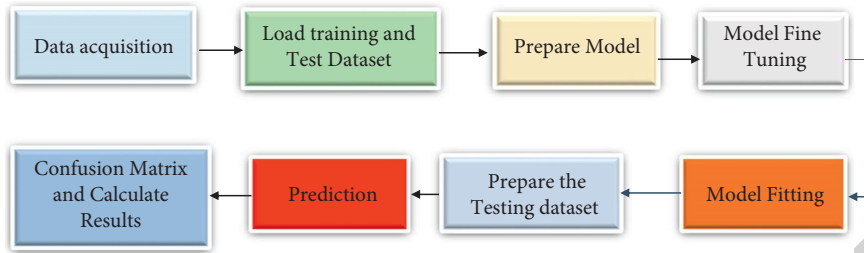


FIGURE 1: The framework of our study.

TABLE 1: Splitting of the dataset for training, testing, and validation.

Total face images	Training set (%)	Validation set (%)	Testing set (%)
2,940	2,540	100	300

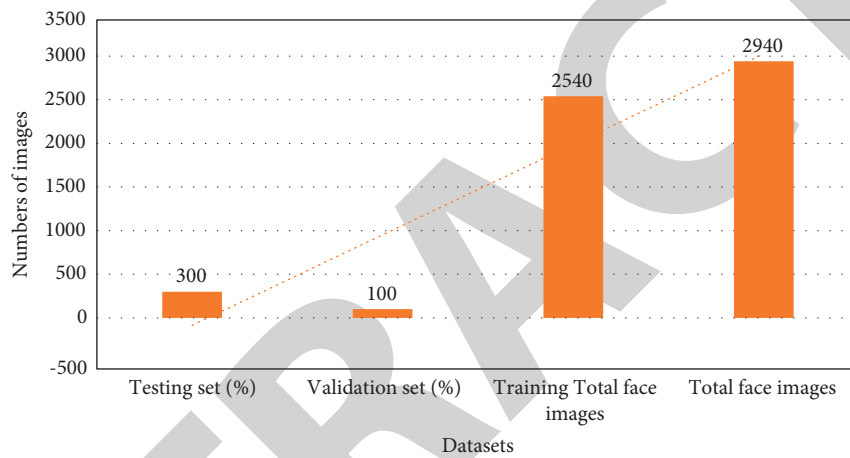


FIGURE 2: Splitting of the dataset.

the human brain through the interconnection and communication between cells. In this section, we will explain the basic components of the CNN model: the input layer, convolutional layer, activating function, pooling layer, fully connected layer, and output prediction.

2.3.2. *Convolutional Layer with a Pooling Layer.* The input of the convolutional layer is an image as a matrix of pixel values. The objective of the convolutional layer is to reduce the images into a form that can be easily processed without losing their important features that will help to detect autism. The first layer of the CNN model is responsible for extracting low-level features such as edges and color. The build of the CNN model allowed us to add more layers to it to enable it to extract the high-level features that will help it understand images. Due to the large number of parameters outputted from the convolution layer, which may significantly prolong the arithmetic operations of the matrices, the number of weights was reduced by using one of the following two techniques: max pooling or average pooling. Max pooling is based on the maximum values in each window in the stride, while average pooling is based on the mean value of each window in the stride. In this study, the model was based on max pooling. Figure 3 shows the convolutional layer and the max pooling and average pooling

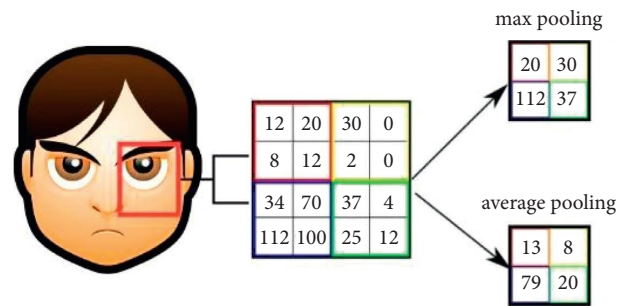


FIGURE 3: The convolution layer and max pooling and average pooling.

processes. The slide window of the kernel extracts the features from the input image and converts the image into a matrix, after which the number of the parameters is mathematically reduced through max pooling and average pooling.

2.3.3. *Fully Connected Layer and Activation Function.* The fully connected (FC) layer is a nonlinear combination of the high-level features that received the input from the hidden layers and are represented as outputs. In the FC layer, the input image is represented as a column vector. The

training of the model has two paths: the forward neural network and backpropagation. The forward neural network feeds from flattened output layer. In the backpropagation, the neural network minimizes the loss errors and learns more features by applying the number of the training iterations. Most deep learning models show high performance while increasing the number of the hidden layers and the training iterations, which allows the neural network to extract the low-level features deeply. The softmax classifier receives the parameters from the FC layer and calculates the properties to predict the output, as shown in Figure 4. A Softmax output of 0 means the image belongs to class 0 and a Softmax output of 1 means the image belongs to class 1. In this study, class 0 is the autism class and class 1 is the normal class.

2.4. Deep Learning Models. This paper is based on three pretrained models for autism detection using facial feature images: NASNetMobile, VGG19, and Xception.

2.4.1. Xception Model. The Xception model was trained on the ImageNet dataset [2, 3] for the image recognition and classification task. Xception is a deep CNN that provides new inception layers. The inception layers are built from depthwise convolution layers and are followed by a pointwise convolution layer. Transfer learning has two concepts: feature extraction and fine-tuning. In this study, the feature extraction method used the pretraining model, which was trained on the standard dataset to extract the feature from the new dataset and to remove the top layers of the model. The new top layers were added to the model for custom classification based on the number of classes. Fine-tuning has been used to adapt the generic features to a given class to avoid overwriting. Figure 5 shows the network used in the Xception model architecture for extracting the image features for the dataset. The Xception architecture used the features maps, followed by a global max pooling layer and two dense layers, 128 and 64, respectively, with a rule activation function. Then, the output of the dense layers was passed to the flatten layer, which took the input as a feature map and the output as a vector. Batch normalization was used to enhance the output by avoiding overfitting. Keras supported the early-stopping method, which stopped the training when the validation loss of the model did not improve. In this model, the RMSprop optimizer was used to reduce the error learning rate or loss during the training of the parameters of the CNN model. The last layer for the output prediction used the Softmax function.

2.4.2. Visual Geometry Group Network (VGG) Model. VGG19 stands for the visual geometry group network (VGG19), a deep artificial neural network model with a 19 artificial multilayers process. Primarily, VGG19 is based on the CNN technique, is commonly implemented on the ImageNet dataset, and is valuable because of its straightforwardness, as 3×3 convolutional layers are attached to its upper side to upsurge with the gravity level. To reduce the

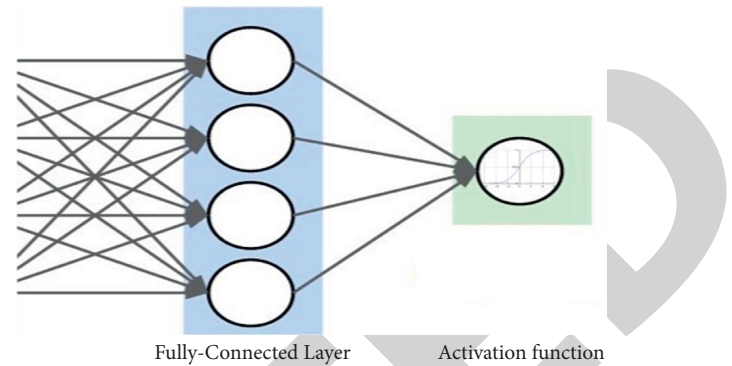


FIGURE 4: The fully connected layer and the activation function.

input volume size, max pooling layers were employed as assigners in VGG19. In the VGG19 structure, two FC layers were adopted with 4,096 neurons to associate each layer with others in the model. The structure of VGG19 is presented in Figure 6.

- (i) Input layer: the function of this layer is to take and accept an image input with a size of $224 \times 224 \times 3$. For the ImageNet race, the inventors of the model picked out the center $224 \times 224 \times 3$ patch in each image to retain the input size of the image dependable.
- (ii) Convolutional layers: VGG's convolutional layers leverage a minimal receptive field, i.e., 3×3 , the minimum probable size that still releases up/down and left/right. The convolution stride is fixed at 1 pixel to conserve the spatial resolution after the convolution (the stride is the number of pixel shifts over the input matrix).
- (iii) Hidden layers: all the hidden layers in the VGG network use ReLU. VGG does not frequently influence Local Response Normalization (LRN), as it increases the memory utilization and the training time. Furthermore, it does not enhance the overall accuracy of the model.
- (iv) Fully connected layers: of the three FC layers, the first two had 4,096 channels each, and the third had 1,000 channels.

3. Experiments

The results of the deep learning models are presented in this section, and the significant results of developing system are declared.

3.1. Experimental Setup. The experiment was executed on different libraries of python and hardware devices for developing an intelligent autism detection system (ADS). Table 2 shows the main requirements for the design of the ADS.

3.2. Evaluation Metrics. This study uses different types of performance evaluation metrics for the three pretrained models such as accuracy, sensitivity, and specificity and a

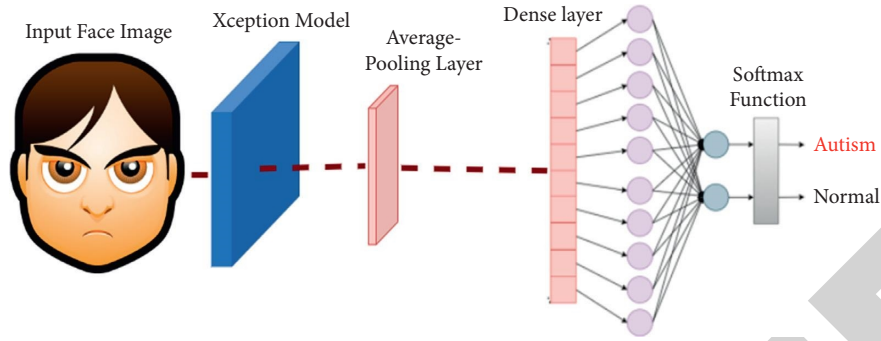


FIGURE 5: Xception model architecture in our study.

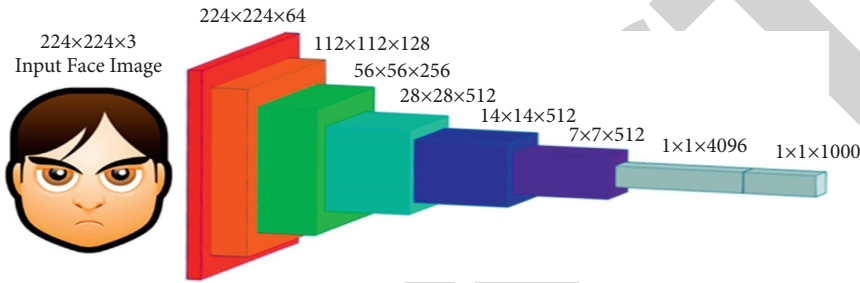


FIGURE 6: VGG19 model.

TABLE 2: Experimental setup.

Hardware	Software/libraries
Processor core I7	TensorFlow library
	Keras library
	Panda seaborn
8 GB RAM	Matplotlib
	Numpy

Note. The model parameters were set with 100 epochs and with a batch size of 32. To avoid overfitting and to optimize the training time, the early stop strategy was used, whereby the training was stopped after 28 epochs. The parameters of the deep learning model are shown in Table 3.

TABLE 3: Parameters used in the pretrained deep learning models.

Parameter name	Value
Global max pooling layer size	3 * 3
Dense layer	128, 64
Batch size	32
Number of epochs	100
Output classification layer	Softmax
Optimizer	ADAM
Activation function	Rule

confusion matrix. A confusion matrix is a type of measure of classification performance that represents a table of the true and false values of the testing results. In the confusion matrix of the Xception model, the True Positives were 132 autistic children out of 150 autistic children, the False Negatives were 18 children classified as autistic children, the True Negatives were 141 correctly classified children out of the 150 normal children, and the False Positives were 9 children.

The equations for these metrics are as follows:

$$\text{Accuracy} = \frac{TP + TN}{FP + FN + TP + TN} \times 100\%,$$

$$\text{Specificity} = \frac{TN}{TN + FN} \times 100\%, \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FP} \times 100\%,$$

where TP is the True Positive, FP is the False Positive, TN is the True Negative, and FN is the False Negative. Specificity is the capacity of the model to correctly identify the normal children, and sensitivity is the capacity of the model to correctly identify autistic children.

4. Results

This section presents the testing results of the experiments conducted to detect ASD. Table 4 summarizes the testing results of the used deep learning models.

In these experiments, three different pretrained deep learning models, namely, Xception, VGG19, and NASNetMobile were implemented to detect ASD. Each model was trained and tested to extract the traits that categorize children as autistic and as normal based on their facial features. Figure 7 shows the confusion metrics of the three deep learning models. They show that the Xception model had the highest testing accuracy, 91%, of the three models, and the NASNetMobile model had the lowest performance level of 78%. Although the dataset was collected from Internet sources by the data generator, which clearly showed the difference in ages and the quality of the photography, the Xception model showed

TABLE 4: Testing results of the pretrained deep learning models.

Model name	Specificity	Sensitivity	Accuracy
Xception	0.94	0.88	0.91
VGG19	0.83	0.78	0.80
NASNETMobile	0.75	0.82	0.78

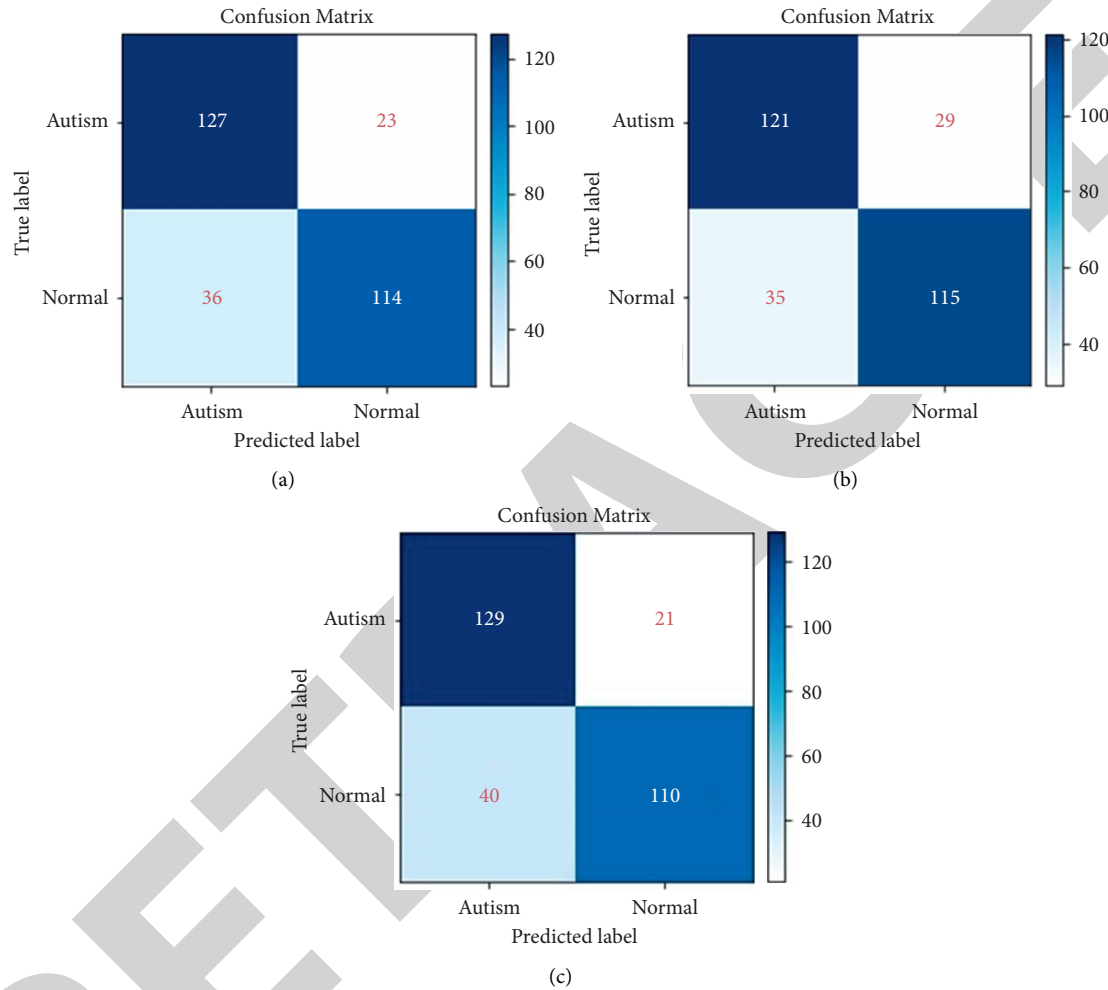


FIGURE 7: Confusion matrices for the. (a) VGG19 model, (b) NASNetmobile model, and (c) Xception model.

the highest accuracy, with only a small percentage of errors.

The performance of the VGG19 model for the training and validation of the data for ASD detection is presented in Figure 8(a), with the y -axis representing the score percentage and the x -axis indicating the number of epochs. In the training process, the accuracy of the VGG19 model rose from 56% to 85% after 25 epochs, and in the validation process, its accuracy was 82%. Its training and validation losses are presented in Figure 8(b). Its training loss was 3.5, and its validation loss was 2.5.

Figure 9 shows the performance of the NASNetMobile model in detecting ASD. The graphic plot shows that the model did not have good results. Its training and validation accuracy percentages are shown in Figure 9(a), with the y -

axis representing the score percentage and the x -axis indicating the number of epochs. Its accuracy in the training stage was 70–90% and in the validation stage was 65–78%. It had good results in the training phase but poorer results in the validation phase than another deep learning model. Its entropy training and validation losses are presented in Figure 9(b). Its training loss was 4.0, and its validation loss was 3.2. Its performance plot shows overfitting in the training process, which explains why its accuracy percentage was lower.

The Xception model achieved 91% accuracy for the testing data and 100% accuracy for the training data. Figure 10(a) shows its accuracy performance. Its accuracy in the training process was 100% and in the validation process was from 70% to 91%. It was observed that the Xception

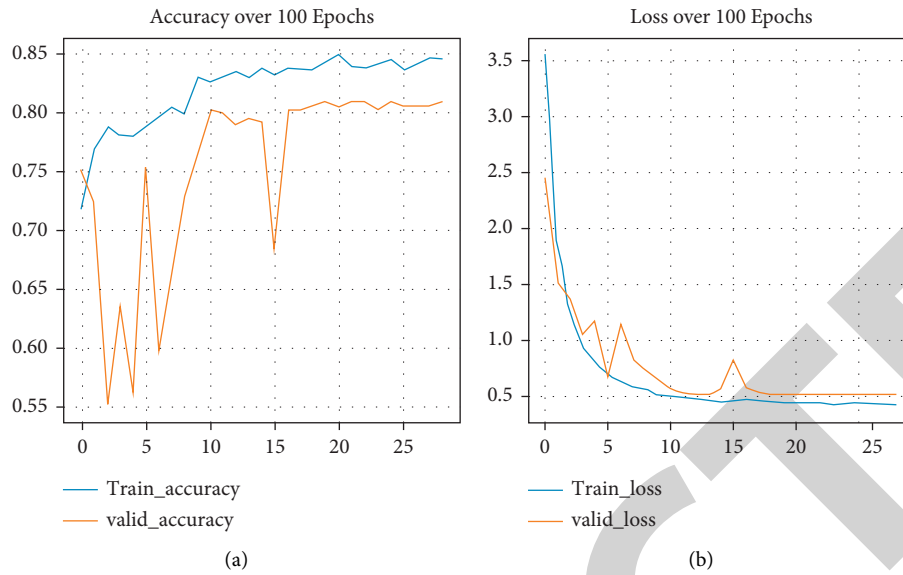


FIGURE 8: Overall performance plot for the VGG19 model. (a) Model performance. (b) Model loss.



FIGURE 9: Overall performance plot for the NASNetMobile model. (a) Model performance. (b) Model loss.

model is the appropriate deep learning model for detecting ASD. Its validation loss is shown in Figure 10(b) as 2.0.

5. Results and Discussion

People with autism face difficulties and challenges in understanding the world around them and in understanding their thoughts, feelings, and needs. The world surrounding a person with autism seems to him like a horror movie, and he finds some sounds and lights and even smells and tastes of foods frightening and sometimes painful. Thus, when a sudden change occurs in their world, they are terrified that no one else can understand.

Diagnosing autism is very important to save the lives of many children. Developing an intelligence system based on AI can help identify autism early. In this study, three advanced deep learning models, namely, Xception, VGG19,

and NASNetMobile were considered for use in diagnosing autism. The empirical results of these models were presented, and it was noted that the Xception model attained the highest accuracy of 91%.

The results of the comparative prediction analysis of the Xception model and the existing system are presented in Table 5. Musser [35] used the VGGFace model that attained 85% accuracy. Tamilarasi et al. [36, 37] used deep learning ResNet50 to detect autism but with small data, that is, only 49 images. Its accuracy was 89%. Jahanara et al. [38] introduced the VGG19 model to identify autism through facial images. The VGG19 model achieved 84% accuracy in the validation accuracy process.

In this study, three deep learning algorithms were used to detect autism. We observed that the Xception model showed significant accuracy, 91%, compared with existing systems. Figure 11 shows the results of the comparison of

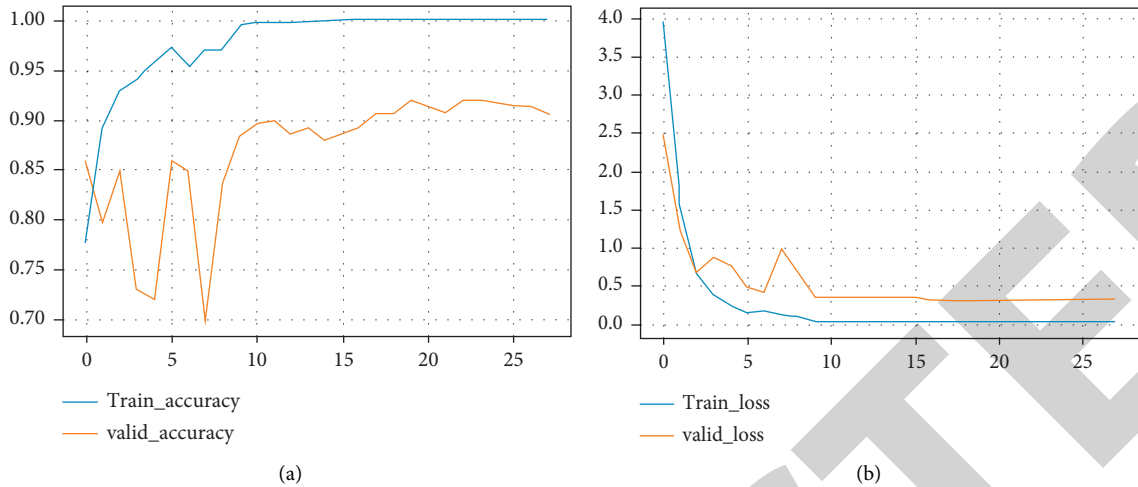


FIGURE 10: Plot of the training and validation accuracy of the Xception model. (a) Accuracy performance. (b) Model loss.

TABLE 5: Comparison of the results of our system and existing models.

Author	Datasets	Model	Accuracy (%)
Han et al. [35]	Same as our dataset	VGGFace	85.0
Haque et al. [10]	The dataset included 19 TD and 20ASD children	ResNet50	89.2
Beary et al. [38]	Same as our dataset	VGG19	84.0
Developing model		MobileNet	91.0

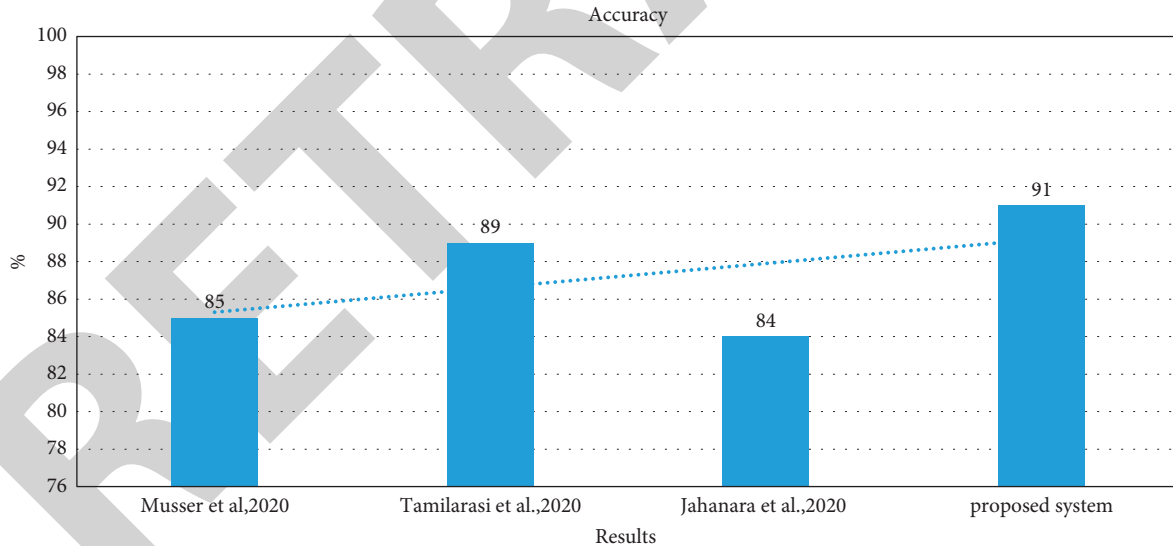


FIGURE 11: Comparison results of the developing system with existing systems.

our developing system with various existing systems. Overall, our developing system outperformed all the existing systems.

6. Conclusions

Interest in child autism has risen due to the advances in global health know-how and capacities. Moreover, the number of autistic children has increased in recent years,

due to which researchers and academics have intensified their efforts to uncover the causes of autism and to detect it early in order to give autistic people behavioral development treatment programs that should help them integrate into society and leave the isolation of the autistic world.

This paper evaluated the performance of three deep learning models in detecting ASD through facial features: NASNETMobile, Xception, and VGG19. Each model was trained on a publicly available dataset on the Internet, and

the best result for classification accuracy was achieved by the Xception model (91%). The results of the model classification showed us the possibility of using such models based on deep learning and computer vision as automatic tools for specialists and families to accurately and more quickly diagnose autism. Computer techniques contribute to the successful conduct of complex behavioral and psychological analyses for autism diagnosis that require a longer time and great effort.

Data Availability

The dataset used to support the findings of the study are available at <https://www.kaggle.com/cihan063/autism-image-data> (accessed on 10 December 2021).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Faisal University for funding this research work through project no. NA00075.

References

- [1] C. S. Paula, S. H. Ribeiro, E. Fombonne, and M. T. Mercadante, "Brief report: prevalence of pervasive developmental disorder in Brazil: a pilot study," *Journal of Autism and Developmental Disorders*, vol. 41, no. 12, pp. 1738–1742, 2011.
- [2] L. C. Nunes, P. R. Pinheiro, M. C. D. Pinheiro et al., *A Hybrid Model to Guide the Consultation of Children with Autism Spectrum Disorder*, A. Visvizi and M. D. Lytras, Eds., Springer International Publishing, View at: Google Scholar, pp. 419–431.
- [3] Apa–American Psychiatric Association, "Diagnostic and statistical manual of mental disorders (DSM -5)," 2020, <https://www.psychiatry.org/psychiatrists/practice/dsm>.
- [4] R. Carette, F. Cilia, G. Dequen, J. Bosche, J.-L. Guerin, and L. Vandromme, "Automatic autism spectrum disorder detection thanks to eye-tracking and neural network-based approach," in *Proceedings of the International Conference on IoT Technologies for Healthcare*, pp. 75–81. 4, Springer, Angers, France, 24–25 October 2017.
- [5] L. Kanner, "Autistic disturbances of affective contact," *Nerv. Child*, vol. 2, pp. 217–250, 1943.
- [6] E. Fombonne, "Epidemiology of pervasive developmental disorders," *Pediatric Research*, vol. 65, no. 6, pp. 591–598, 2009.
- [7] "International statistical classification of diseases and related health problems (ICD)," <https://www.who.int/standards/classifications/classification-of-diseases>.
- [8] G. Yolcu, I. Oztel, S. Kazan et al., "Deep learning-based facial expression recognition for monitoring neurological disorders," in *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1652–1657, Kansas City, MO, USA, 13–16 November 2017.
- [9] G. Yolcu, I. Oztel, S. Kazan et al., "Facial expression recognition for monitoring neurological disorders based on convolutional neural network," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 31581–31603, 2019.
- [10] M. I. U. Haque and D. Valles, "A facial expression recognition approach using DCNN for autistic children to identify emotions," in *Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 546–551, Vancouver, Canada, 1–3 November 2018.
- [11] O. Rudovic, Y. Utsumi, J. Lee et al., "CultureNet: a deep learning approach for engagement intensity estimation from face images of children with autism," in *Proceedings of the 2018 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 339–346, ISSN 2153-0866, Madrid, Spain, 1–5 October 2018.
- [12] M. S. Satu, F. Farida Sathi, M. S. Arifen, M. Hanif Ali, and M. A. Moni, "Early detection of autism by extracting features: a case study in Bangladesh," in *Proceedings of the 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, pp. 400–405, Dhaka, Bangladesh, 10–12 January 2019.
- [13] Q. Guillon, N. Hadjikhani, S. Baduel, and B. Rogé, "Visual social attention in autism spectrum disorder: insights from eye tracking studies," *Neuroscience & Biobehavioral Reviews*, vol. 42, pp. 279–297, 2014.
- [14] T. Akter, M. S. Satu, L. Barua, F. F. Sathi, and M. H. Ali, "Statistical analysis of the activation area of fusiform gyrus of human brain to explore autism," *International Journal of Computer Science and Information Security*, vol. 15, pp. 331–337, 2017.
- [15] S. Schelinski, K. Borowiak, and K. von Kriegstein, "Temporal voice areas exist in autism spectrum disorder but are dysfunctional for voice identity recognition," *Social Cognitive and Affective Neuroscience*, vol. 11, no. 11, pp. 1812–1822, 2016.
- [16] X. Jiang and Y.-F. Chen, "Facial image processing," in *Applied Pattern Recognition*, H. Bunke, A. Kandel, and M. Last, Eds., Springer, Berlin/Heidelberg, Germany, pp. 29–48, 2008.
- [17] M. Duda, R. Ma, N. Haber, and D. P. Wall, "Use of machine learning for behavioral distinction of autism and ADHD," *Translational Psychiatry*, vol. 6, no. 2, Article ID e732, 2016.
- [18] G. Deshpande, L. E. Libero, K. R. Sreenivasan, H. D. Deshpande, and R. K. Kana, "Identification of neural connectivity signatures of autism using machine learning," *Frontiers in Human Neuroscience*, vol. 7, p. 670, 2013.
- [19] M. N. Parikh, H. Li, and L. He, "Enhancing diagnosis of autism with optimized machine learning models and personal characteristic data," *Frontiers in Computational Neuroscience*, 2019.
- [20] F. Thabtah and D. Peebles, "A new machine learning model based on induction of rules for autism detection," *Health Informatics Journal*, vol. 26, no. 1, pp. 264–286, 2020.
- [21] M. H. Al Banna, T. Ghosh, K. A. Taher, M. S. Kaiser, and M. Mahmud, "A monitoring system for patients of autism spectrum disorder using artificial intelligence," in *Proceedings of the International Conference on Brain Informatics*, pp. 251–262, Springer, Padua, Italy, July 2020.
- [22] M. Li, D. Tang, J. Zeng et al., "An automated assessment framework for atypical prosody and stereotyped idiosyncratic phrases related to autism spectrum disorder," *Computer Speech & Language*, vol. 56, pp. 80–94, 2019.
- [23] F. B. Pokorny, B. W. Schuller, P. B. Marschik et al., "Earlier identification of children with autism spectrum disorder: an automatic vocalisation-based approach," in *Proceedings of the Annual Conference of the International Sp*, August 2017.

Retraction

Retracted: 5G Massive MIMO Signal Detection Algorithm Based on Deep Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] L. Yan, Y. Wang, and N. Zheng, "5G Massive MIMO Signal Detection Algorithm Based on Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 9999951, 9 pages, 2022.

Research Article

5G Massive MIMO Signal Detection Algorithm Based on Deep Learning

Lichao Yan , Yi Wang , and Ning Zheng 

School of Intelligent Engineering, Zhengzhou University of Aeronautics, Zhengzhou, Henan 450046, China

Correspondence should be addressed to Lichao Yan; lcyan002480@zua.edu.cn

Received 28 December 2021; Revised 26 January 2022; Accepted 3 February 2022; Published 27 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Lichao Yan 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.

Aiming at the problems of poor signal detection effect caused by many interference factors in large-scale MIMO technology scene, this paper proposes a 5G massive MIMO signal detection algorithm based on deep learning. Firstly, the MIMO system model based on neural network is constructed, and Deep Neural Network (DNN) detection is introduced into the receiver of the traditional MIMO system to obtain the information bits or codewords and channel state information transmitted by transmitters. Then, the end-to-end training method is adopted to make neural network learn the mapping relationship of information bits or codewords transmitted by system transceivers. Furthermore, DNN detector is improved based on Simplified Message Passing Detection (sMPD) algorithm, and the correction factor is updated continuously to optimize network parameters to realize the accurate detection and decoding of the MIMO system. Finally, the proposed algorithm is experimentally analyzed based on the TensorFlow deep learning framework. Experimental results show that when signal-to-noise ratio is 10 dB, the bit error rate and mean square error are lower than 0.005 and 0.1, respectively.

1. Introduction

At present, as one of mainstream research directions in the field of wireless communication, intelligent communication is actively introducing artificial intelligence technology into all levels of the wireless communication system, which has become an effective way to explore the field of intelligent communication [1]. As one of key technologies of the 5G mobile communication system, large-scale Multiple Input Multiple Output (MIMO) technology has high spectrum utilization and link reliability [2]. However, due to the increase in number of antennas configured at transceiver end of the MIMO system, the signal processing process at receiving end of the communication system becomes more complicated, which will make signal detection face a huge challenge of high computational complexity [3, 4]. Therefore, designing signal detection algorithms with low computational complexity and high detection performance is an urgent problem for massive MIMO systems [5]. The popular research direction of the MIMO system has turned to the suboptimal detection algorithm.

The purpose of massive MIMO signal detection is to recover the original signal at the transmitter under noise and interference environments. It is an important link to improve the overall performance of the system and reduce the complexity of the system [6]. Low-complexity signal detection algorithms have become a research hotspot in recent years [7]. For classic linear detection algorithms such as ZF detection, the computational complexity is too high due to its large number of matrix inversion operations [8], the detection performance of linear detection algorithm is not ideal, so it is difficult to apply to the actual massive MIMO system [9]. Reference [10] proposed a low-complexity massive MIMO detection method based on approximate expected propagation. This method uses the approximation method of the channel hardening phenomenon to eliminate matrix inversion during iteration and reduces the computational complexity of accurate detection while maintaining good performance. In reference [11], a linear detector based on conjugate gradient of multiple search directions was proposed in the massive MIMO uplink system in order to reduce the complexity of data detection. It determined the

search direction in each subdomain by the projection method, which improves the performance of the algorithm. However, the loop effect caused problems such as slow or nonconvergence, which leads to a decrease in algorithm detection performance. Reference [12] proposed a system model that considers the reverberation effect of radar systems on massive MIMO receivers using the receiver for uplink to perform channel estimation and data detection. The experimental results proved the effectiveness of the proposed detection scheme. Reference [13] proposed an iterative precoding scheme based on Chebyshev acceleration. The simulation results showed that the proposed method has lower complexity while ensuring higher detection performance. However, it was difficult to be widely used in actual promotion and application due to the limitations of the coding scheme.

As deep learning has made great achievements in other fields, there are more and more research studies on applying deep learning to communication systems [14]. Orthogonal approximation message passing detection neural network uses the idea of the orthogonal approximation message passing detection algorithm, while increasing adjustable parameters using deep learning optimization methods to improve its detection performance. Reference [15] proposed a heuristic method for pilot/data power optimization. This method could improve algorithm performance through space-time block coding when the access point does not have any channel state information. However, when applied to a high-order modulation system, the computational complexity will increase as the number of users and modulation order increase. Reference [16] algorithm could speed up the convergence speed by reducing the spectral radius of the rederived iterative matrix. Reference [17] proposed a MIMO detection scheme based on deep learning using a novel training algorithm. The algorithm accelerated the training by time and spectrum correlation in the real channel, so that it has lower computational complexity than the existing methods on real channel. However, each neural network layer needs to perform matrix inversion operation, which brings high computational complexity. The computational advantages of deep learning improve the detection performance of massive MIMO signals. However, the method of extending the iterative algorithm to neural network is more dependent on the channel environment [18].

This paper proposes a massive MIMO signal detection algorithm based on deep learning to address the problems of large number of MIMO antennas in the 5G communication system, which makes signal detection difficult. Compared with traditional detection algorithms, its innovations are summarized as follows:

- (1) Due to the complexity of DNN parameters, the proposed algorithm uses the Simplified Message Passing Detector (sMPD) algorithm to optimize network parameters iteratively to reduce the computational difficulty and improve network signal analysis capabilities.

- (2) In order to realize the accurate detection of massive MIMO signal, the proposed algorithm constructs a detection system based on neural network. Among them, DNN detector improved by the sMPD algorithm is used to analyze MIMO signal to achieve low error signal transmission.

2. System Model and Problem Analysis

Consider a massive MIMO system in a single-cell multiuser Time Division Duplex (TDD) mode. The base station is equipped with M antennas, and the uplink channel estimation process is shown in Figure 1.

In the uplink, the base station receives signals simultaneously transmitted by N single antenna user in the same time-frequency resource, which can be expressed as follows:

$$\mathbf{y}(t) = \mathbf{G}\mathbf{x}(t) + \boldsymbol{\delta}(t), \quad (1)$$

where \mathbf{G} represents the flat fading channel matrix of $M \times N$ between the base station and N users, $\mathbf{x}(t)$ represents the signal vector of $N \times 1$ sent by N users at time, and $\boldsymbol{\delta}(t)$ represents additive white Gaussian noise with a mean value of 0 and a variance of $\sigma_n^2/2$.

Assuming that each user transmits an orthogonal pilot sequence of length L ($L \geq N$) in the channel coherence time, the pilot signal matrix received by base station is as follows:

$$\mathbf{Y} = \mathbf{G}\boldsymbol{\Psi} + \hat{\boldsymbol{\delta}}, \quad (2)$$

where $\boldsymbol{\Psi} = [\varphi(1), \varphi(2), \dots, \varphi(L)]$ is the pilot matrix of dimension $N \times L$ that contains all user training sequences and $\hat{\boldsymbol{\delta}}$ is an additive white Gaussian noise matrix whose elements are all Gaussian random variables with independent identically distributed mean value 0 and variance $\sigma_n^2/2$.

According to the system model, Rice flat fading channel matrix \mathbf{G} can be further expressed as follows:

$$\mathbf{G} = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{G}} + \sqrt{\frac{1}{\kappa+1}}\tilde{\mathbf{G}}, \quad (3)$$

where $\bar{\mathbf{G}}$ represents the deterministic component matrix containing line-of-sight signal and $\tilde{\mathbf{G}}$ represents Rayleigh random component matrix containing the scattered signal. $\kappa \geq 0$ is the Rice fading factor; when $\kappa = 0$, the channel is Rayleigh channel, that is, $\mathbf{G} = \tilde{\mathbf{G}}$.

For the deterministic component $\bar{\mathbf{G}}$, it can be expressed as

$$[\bar{\mathbf{G}}]_{mn} = e^{-j(m-1)\frac{2\pi d}{\lambda}\sin(\theta_n)}, \quad (4)$$

d and λ represent the base station antenna spacing and radio wave wavelength, respectively, and θ_n represents the angle of arrival of n user.

For Rayleigh random component $\tilde{\mathbf{G}}$ containing the scattered signal, the finite scattering channel model is considered. Assuming that each user has q paths to base

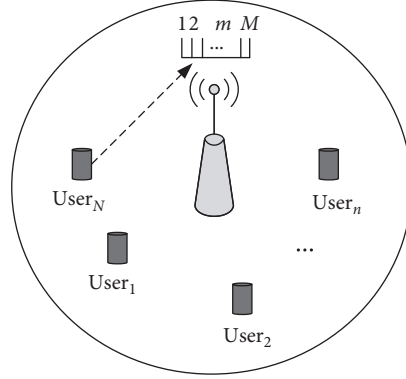


FIGURE 1: Estimation scenario of large-scale MIMO uplink channel.

station, the channel vector \mathbf{g}_n from the n user to the base station can be expressed as follows:

$$\mathbf{g}_n = \frac{1}{\sqrt{Q}} \sum_{q=1}^Q h_{nq} \phi(\theta_q) \quad (5)$$

$$\phi(\theta_q) = \left[1, e^{-j2\pi d/\lambda \cos(\theta_q)}, \dots, e^{-j2\pi \frac{(M-1)d}{\lambda} \cos(\theta_q)} \right]^T,$$

where h_{nq} is the random transmission gain of q path. In order to facilitate the analysis without loss of generality, it is assumed that the path gain obeys a complex Gaussian distribution with a mean value of 0 and a variance of 1. θ_q is the random arrival angle of q path; $\phi(\theta_q)$ is the direction vector under ULA model; and d and λ represent the base station antenna spacing and radio wave wavelength, respectively.

In the massive MIMO system, it is usually relevant on the base station side but not on the user side. This is due to the limited number of scatterers on base station side, while the users have relatively abundant scatterers [19, 20]. Therefore, it can be assumed that the antenna direction vectors of users are the same, and the random component channel matrix can be expressed as follows:

$$\mathbf{G} = \Omega \tilde{\mathbf{H}}, \quad (6)$$

where $\Omega = [\phi(\theta_1), \phi(\theta_2), \dots, \phi(\theta_q)]$ is the full-rank matrix of $M \times Q$ containing Q direction vectors; $\tilde{\mathbf{H}}$ is the path gain matrix of $Q \times N$; $[\tilde{\mathbf{H}}]_{q,n} = h_{nq}$.

According to the nature of rank, in Rice flat fading channel model, the rank of \mathbf{G} is $\text{rank}(\mathbf{G}) \leq \text{rank}(\tilde{\mathbf{G}}) + \text{rank}(\tilde{\mathbf{G}})$. The number of antennas M and the number of users N are usually large ($M \gg N$), while the number of effective paths Q is relatively small. Therefore, in the physical finite scattering model, the rank $\text{rank}(\tilde{\mathbf{G}}) = \min\{M, N, Q\} = Q$ of Rayleigh random component $\tilde{\mathbf{G}}$. The rank $\text{rank}(\mathbf{G}) \leq N + Q$ of channel matrix \mathbf{G} is significantly smaller than the matrix dimension.

3. MIMO Signal Detection Model Based on Neural Network

3.1. MIMO System Model. Taking the MIMO system as the main research object, M_t and M_r antennas are installed in the transceiver end of the system. The model structure of the MIMO communication system is shown in Figure 2. In order to accurately detect and decode the transmitted information at the receiving end of MIMO system, aiming at this goal, it mainly combines DNN and autoencoder neural network model. The two different neural network structures and physical layer framework of the traditional MIMO wireless communication system are organically integrated. Its purpose is to process the information bits or code words received by the MIMO system.

The mathematical model of the entire MIMO system can be expressed by the following linear equation:

$$\mathbf{y} = \mathbf{G}\mathbf{x} + \delta, \quad (7)$$

where $\mathbf{x} = [x_1, \dots, x_{M_t}]^T \in \mathbb{C}^{M_t \times 1}$ represents the information vector sent by the transmitting end of the system and $\mathbf{y} = [y_1, \dots, y_{M_r}]^T \in \mathbb{C}^{M_r \times 1}$ is the information vector received by the receiving end of the MIMO system.

3.2. Input and Output of Neural Network. The number of neurons in input layer of a neural network generally depends on the attributes of sample data. In the proposed neural network model, matrix \mathbf{G} and vector \mathbf{y} are used as the input of neural network. Before input, the data are preprocessed, matrix \mathbf{G} is converted into a $(M \times N)$ dimensional column vector, and the column vector is combined with column vector \mathbf{y} . Since each element of column vector is a complex number, the real part and the imaginary part are separated and connected in series as the network input \mathbf{I} , $\mathbf{I} = [g_{1,1}, g_{2,1}, \dots, g_{M,N}, y_1, y_2, \dots, y_M]^T$. The number of input neurons of neural network is $2(N \times M + M)$, and the activation function of output layer is selected according to digital modulation mode of the MIMO system [21]. Taking binary phase shift keying as an example, since the transmitted signal symbol is 0 or 1, the activation function is

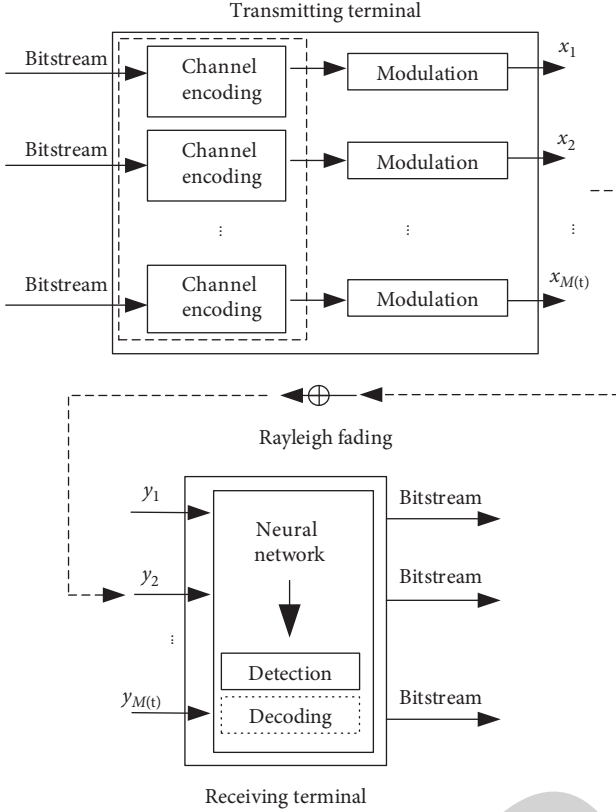


FIGURE 2: MIMO system model based on neural network.

selected as Sigmoid function. The reason is that the output of Sigmoid is mapped in the range of (0, 1), which has a strong correlation with the transmitted symbol. For the modulation method, a linear activation function can be selected [22]. After neural network is output, the output is recombined into the transmission signal pattern, and symbol decision is performed, and the result is detection signal x^* of MIMO system receiver based on neural network. The flow of entire data in the receiver is shown in Figure 3.

3.3. Designed DNN Detector Based on sMPD Algorithm. A detection network is designed based on the sMPD algorithm, which relies on the sMPD algorithm. That is, using the existing algorithm as a starting point, each iteration of the sMPD algorithm is expanded to each layer in the network, and the optimization method of deep learning is used to obtain the best training parameters to obtain a neural network that meets the signal detection performance requirements [23, 24]. From this perspective, deep learning provides a powerful tool. It can make the improved MPD get the best correction factor to achieve higher performance. In addition, DNN detector designed based on the sMPD algorithm constructs a DNN by expanding each iteration of iterative sMPD to each layer of detection network. Among them, correction factor set is the parameters to be optimized, and the best correction factor is obtained by “learning” by deep learning training [25].

DNN is one of typical deep learning models, which can map input $x_0 \in \mathbb{C}^{M_i \times 1}$ to output $y_0 \in \mathbb{C}^{M_o \times 1}$:

$$y = f(x_0, \omega), \quad (8)$$

where ω represents the parameter that approximates optimal function, which maps the input to desired output.

DNN usually has multiple hidden layers between the input and output layers, and a multilayer neural network combines many functional units. For a DNN with K layers, the output of $(k-1)$ is used as the input of k layer, and the mapping function of k layer can be defined as follows:

$$x_k = f^{(1)}(x_{k-1}; \omega_k), \quad (9)$$

where ω_k represents the training parameters of the k layer and $f^{(1)}(x_{k-1}; \omega_k)$ is the mapping function of the k layer.

Due to the similarity of the factor graph model and DNN structure, DNN can be designed by unfolding iterative algorithm. Each iteration in the algorithm corresponds to each layer in the network, as shown in Table 1. The comparison of MPD’s factor graph model and DNN structure confirms that they have similar structures. Based on a reasonable design of detection network, the MPD algorithm can be further improved through the optimization method of deep learning.

In the damped MPD algorithm, the damping factor for each iteration can be different. When selecting the normalization/offset factors of sMPD, these factors can be expanded to have a different symbol probability $p_{ij}^{(k)}$ for each message. In fact, a different correction factor can be set for each message in each iteration, and the calculation of prior probability can be expressed in a unified manner. Specifically, by increasing the damping factor, $p_{ij}^{(k)}$ is defined as follows:

$$p_{ij}^{(k)} \leftarrow (1 - \beta_{ij}^{(k)}) p_{ij}^{(k)} + \beta_{ij}^{(k)} p_{ij}^{(k-1)}, \quad (10)$$

where $\beta_{ij}^{(k)}$ is the weighted average damping factor of message calculated in the current iteration and the message obtained in the previous iteration.

To further increase the correction factor $\tau_{ij}^{(k)}$ and bias factor $\mu_{ij}^{(k)}$, (10) is converted to the following:

$$p_{ij}^{(k)} \leftarrow (1 - \beta_{ij}^{(k)}) (\tau_{ij}^{(k)} p_{ij}^{(k)} - \mu_{ij}^{(k)}) + \beta_{ij}^{(k)} p_{ij}^{(k-1)}. \quad (11)$$

The damped sMPD algorithm requires multidimensional parameters to approximate MPD, and these multidimensional parameters are designed to further improve performance. However, it will lead to an increase in the number of parameters to be optimized, especially when the number of transmitting and receiving antennas is large. For traditional methods, this is a complex optimization problem. However, it can be solved by deep learning, a powerful optimization tool [26].

In summary, the first iteration of the improved MPD algorithm can be summarized by two kinds of messages, including Log-Likelihood Ratio (LLR) $\zeta^{(k)}$ and probability $p^{(k)}$ and the set of correction factor $\tau^{(k)}$. A complete iteration in the sMPD algorithm can be mapped to a layer in DNN to construct DNN-MIMO detector and use $\tau^{(k)}$ as the training

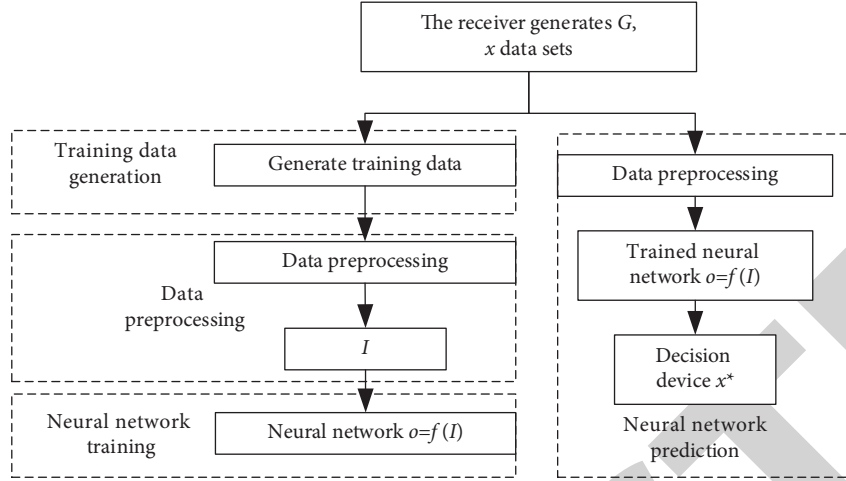


FIGURE 3: Flow process of data flow.

TABLE 1: MPD diagram model and DNN structure.

MPD	DNN
Node	Neuron
Send signal x	Input data x
Receive signal y	Output data y
The k -th iteration	The k -th hidden layer
Confidence message	Hidden layer information
Data update rules	Hidden relationship of each layer
Correction factor	Parameter ω

parameter set to be learned. Specifically, the DNN detector based on sMPD is expressed as follows:

$$\begin{cases} \zeta^{(k)}, p^{(k)} = f^{(k)}(\zeta^{(k-1)}, p^{(k-1)}, \tau^{(k)}) \\ o = \sigma(\zeta^{(K)}) \end{cases}, \quad (12)$$

where $f^{(k)}(\zeta^{(k-1)}, p^{(k-1)}, \tau^{(k)})$ is the mapping relationship of l -th iteration in the improved MPD algorithm, o is the final output of DNN, and $\sigma(\cdot)$ is the normalization function. The output $\zeta^{(K)}$ of the K final iteration is rescaled within the range of $[0, 1]$.

For the sMPD algorithm, $\zeta^{(k)} = \{\zeta_j^{(k)}\}$ and $p^{(k)} = \{p_j^{(k)}\}$, and training factor $\Delta = \{\alpha^{(1)}, \dots, \alpha^{(K)}, \omega^{(1)}, \dots, \omega^{(K)}, b^{(1)}, \dots, b^{(K)}\}$ include attenuation, scaling, and bias factors. This sMPD-based DNN detector is called Improved MPD (IMPD). The overall structure of IMPD detector is shown in Figure 4. The hidden layer in DNN is expanded by the sMPD algorithm. These hidden layers are also different according to the iterative process of the selected MPD algorithm, thus forming a variety of different DNN detectors.

In order to measure the detection performance of signal detection network designed for the massive MIMO system, cross entropy is used to express the error between output of neural network and real transmitted signal. And the minimum SGD method is used to minimize loss function to determine the best correction factor set (Δ).

The IMPD detector includes two stages of learning and detection. First, neural network is trained to obtain the

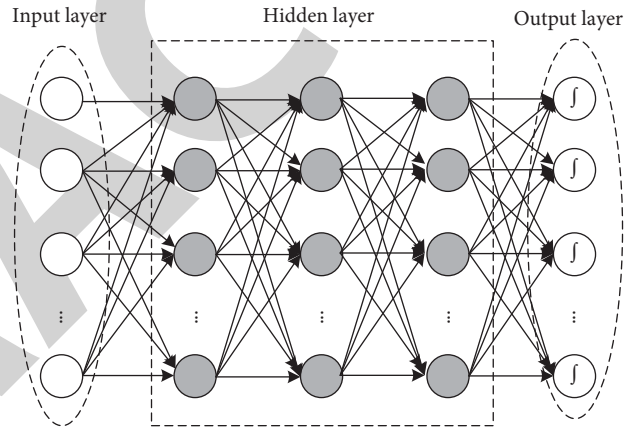


FIGURE 4: Structure of IMPD detector.

optimized correction factor. After that, trained DNN detector is used for signal detection. This DNN based on MPD training is similar to the MPD algorithm, except that the optimized correction factor is obtained through the training process. Therefore, IMPD detector can be regarded as a special case of the MPD algorithm and requires additional training preprocessing. However, this is all done offline on the computer, and it will not increase the complexity of performing real-time detection tasks.

3.4. Batch Normalization and Dropout. The proposed neural network model is relatively complicated, and there are too many parameters that need to be updated, so it is prone to overfitting during the training process. Overfitting means that the function fitted by the neural network fits the training set perfectly, but the results of validation set and test set containing new data are poor. That is, the training data are overfitted without considering the generalization ability of network.

The overfitting phenomenon can be solved by batch normalization and dropout. Batch training refers to training a batch of data at a time during the neural network training process. The purpose of training is to reduce the loss

function value of entire batch of data. This can reduce the randomness caused by a single data training as much as possible. A batch of data is called a batch, and all training sets the batch value to 1000. Dropout was first proposed by Hinton in 2012. It refers to randomly ignoring a certain percentage of neurons during neural network training. The neural network model using Dropout randomly selects a certain number of neurons for training during the training process and does not need to rely too much on certain local features. This can reduce the interaction between feature detectors (neurons) and make the model more generalized. Using Dropout in the training process is equivalent to different neural networks for each training, and each neural network will also give different results. However, as the training process progresses, most of neural network output results will be correct. Then, a small number of wrong results will not affect the total output. The Dropout operation was used during training.

4. Experiment and Analysis

With the help of advanced deep learning libraries such as TensorFlow, the proposed algorithm is experimentally analyzed, and the simulation parameter settings are shown in Table 2.

4.1. Performance Comparison of Deep Learning and Linear Detection Algorithms. The performance of the used deep learning method is compared with that of the traditional linear detection algorithm (maximum likelihood algorithm and zero-forcing algorithm). The result of bit error rate (BER) is shown in Figure 5.

It can be seen from Figure 5 that the detection performance based on the deep learning method applied to MIMO system signal detection is lower than the performance of the optimal maximum likelihood detection algorithm. However, it is significantly better than zero-forcing detection algorithm performance. Since the principle of the traditional linear detection algorithm is relatively simple, the proposed algorithm shows better performance advantages after testing the neural network with better training effect. When BER is 0.01, signal-to-noise ratio (SNR) difference between the deep learning method and maximum likelihood detection algorithm is about 2.5 dB. BER is in the range of 0.1 to 0.01. This deep learning method has a performance gain of nearly 2 dB.

4.2. Performance Comparison of Deep Learning and Belief Propagation Decoding Algorithms. In order to further verify the performance of the deep learning method in channel decoding in the MIMO communication system, the decoding performance of the Back Propagation (BP) algorithm of polarization code is selected and compared. In the experiment, the number of iterative cycles of the BP decoding algorithm is set to 200. After the iteration loop, the BER result obtained is shown in Figure 6.

It can be seen from Figure 6 that the performance of channel decoding based on deep learning is better than that of the traditional BP decoding algorithm. When BER is 0.01,

TABLE 2: Simulation parameters.

Parameter	Value
Pilot sequence length	55
Number of base station antennas M	100
Signal-to-noise ratio (dB)	10
Number of users N	50

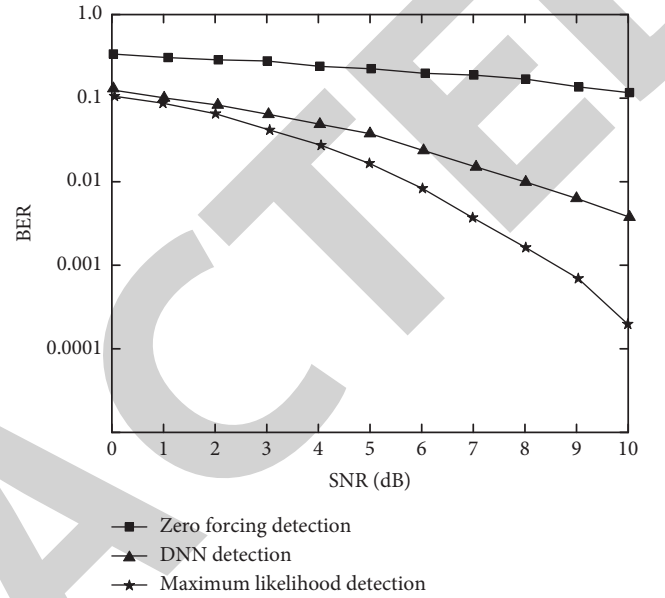


FIGURE 5: Performance comparison between deep learning and linear detection algorithms.

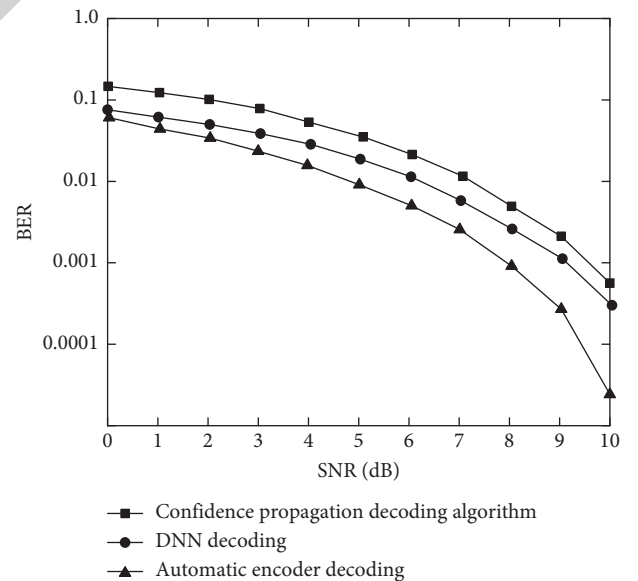


FIGURE 6: Performance comparison between deep learning and belief propagation decoding algorithms.

the decoding method using autoencoder neural network has a performance gain of about 1.5 dB. The decoding performance of DNN also has a performance gain of nearly 1 dB. This demonstrates that the deep learning method can improve channel decoding performance of the MIMO system.

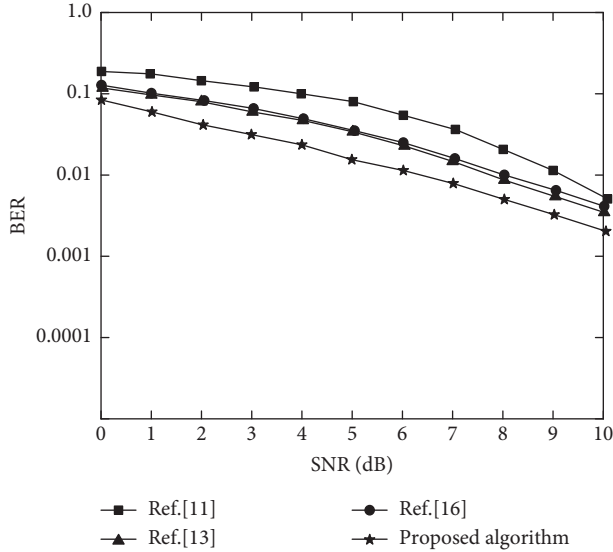


FIGURE 7: Detection performance comparison of different algorithms.

4.3. Detection Performance Comparison of Different Algorithms. In order to demonstrate the performance of the proposed algorithm, the algorithms in reference [11], reference [13], and reference [16] are compared. In SNR range of 0~10 dB, the trained model will be tested, and BER results under different SNRs are shown in Figure 7.

It can be seen from Figure 7 that in the same SNR range, the detection performance of the proposed algorithm is better than that of other comparison algorithms. When SNR is 10 dB, its BER is lower than 0.005, because the proposed algorithm constructs a detection system based on neural network and uses the DNN detector improved by the sMPD algorithm to analyze MIMO signals, so low error signal transmission is realized. The improved linear detection algorithm used in reference [11] is affected by loop effect, and its detection performance is not good. When SNR is 9 dB, its BER is still higher than 0.01. Reference [16] carried out parallel iteration based on the linear detection algorithm, and the detection speed and performance have been improved to a certain extent. However, the data processing capability of the linear detection algorithm itself is insufficient, so there is a significant gap compared with IMPD of the proposed algorithm. Reference [13] used the iterative precoding algorithm of Chebyshev acceleration for signal detection. Only from the iterative aspect of optimization, its detection performance improvement effect is not obvious. Thus, when BER is 0.01, the detection performance of the proposed algorithm and reference [13] had a performance gain of nearly 1.5 dB. And within this SNR range, the detection performance of the MIMO system shows the most obvious advantages.

Besides, mean square error (MSE) results of four comparison algorithms under different SNRs are shown in Figure 8.

It can be seen from Figure 8 that the MSE of the proposed algorithm is significantly lower than that of other comparison algorithms, especially better than that of the

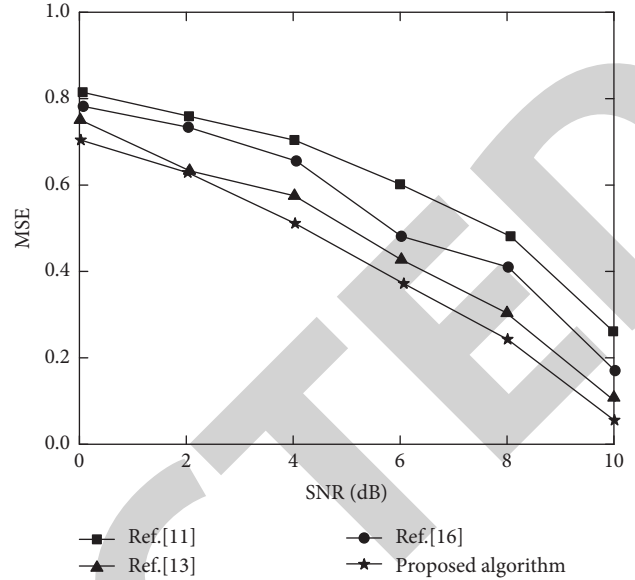


FIGURE 8: MSE curves of different algorithms.

linear detection algorithm of reference [11]. Based on the linear detection algorithm, the neural network is used to compensate for the lack of channel estimation, which is closer to the real channel distribution characteristics. Reference [16] incorporated parallel iteration on the basis of linear detection, and MSE of detection result is reduced, but its data learning ability is somewhat weaker than that of deep learning algorithms. Reference [13] used the iterative precoding algorithm to achieve MIMO signal detection, and its detection accuracy is high. When SNR is 8 dB, its MSE is lower than 0.4. However, the influence of parameters in the deep learning algorithm and the interference of iteration errors are not fully considered, so the detection effect of the proposed algorithm is better. The proposed algorithm uses designed IMPD detector for MIMO signal detection and increases the correction factor and parameter training to reduce detection error. When SNR is 10 dB, its MSE is lower than 0.1, which can meet the requirements of actual environment.

5. Conclusion

Massive MIMO technology has become one of key technologies of 5G communication systems due to its high spectrum utilization and link reliability. However, as the number of antennas in the massive MIMO system increases, signal detection will face the challenge of high computational complexity. To this end, this paper proposes a 5G massive MIMO signal detection algorithm based on deep learning. Besides, the sMPD algorithm is used to optimize the network parameters of DNN, and the MIMO system model based on neural network is constructed. The trained DNN detector learns to process the transmission signal to achieve accurate decoding and detection of 5G massive MIMO signal. Experimental results based on TensorFlow deep learning framework show that

- (1) The DNN detector designed based on the sMPD algorithm has better detection performance and reduces BER through the iterative optimization of correction factor and network parameters. Moreover, the detection performance advantage is obvious.
- (2) The proposed algorithm builds the MIMO system with the help of neural network learning advantages and uses batch training and dropout to improve the generalization ability of the model, which can accurately and efficiently realize signal detection. When SNR is 10 dB, its BER and MSE are lower than 0.005 and 0.1, respectively.

Since the algorithm in this paper is proposed under the assumption that the channel environment is in an ideal state, the interference of some objective factors is ignored. Further, in-depth research is needed for more nonideal channel environments and realistic system scenarios. In addition, all the proposed improved algorithms only consider the signal detection problem itself, without considering the communication error rate performance that combines it with channel estimation, channel coding, and other issues.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61801435, in part by the Scientific and Technological Key Project of Henan Province under Grant 212102210559, in part by Henan Province Science Foundation for Youths under Grant 212300410296, and Scientific and technological project in Henan Province under Grant 202102210334.

References

- [1] Q. Liang, J. Mu, M. Jia, W. Wang, X. Feng, and B. Zhang, "Communications, Signal Processing, and Systems," in *Proceedings of the 2017 International Conference on Communications, Signal Processing, and Systems, ICCSP 2017, Harbin, 2017*.
- [2] R. Hayakawa and K. Hayashi, "Error recovery for massive MIMO signal detection via reconstruction of discrete-valued sparse vector," *IEICE-Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100.A, no. 12, pp. 2671–2679, 2017.
- [3] T. X. Doan, C. D. Ho, and H. Q. Ngo, "Massive MIMO under multi-keyhole channels: does the use-and-then-forget bounding technique work?" *Physical Communication*, vol. 47, no. 11, pp. 101384–101398, 2021.
- [4] Y. Ji, F. Meng, J. Jin et al., "Quantum version of MMSE-based massive MIMO uplink detection," *Quantum Information Processing*, vol. 19, no. 2, pp. 1–28, 2020.
- [5] K. Imran, "A robust signal detection scheme for 5G massive multiuser MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 9597–9604, 2018.
- [6] S. Liang, X. Wang, and L. Ping, "Semi-blind detection in hybrid massive MIMO systems via low-rank matrix completion," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5242–5254, 2019.
- [7] Y. Lee, "Decision-aided Jacobi iteration for signal detection in massive MIMO systems," *Electronics Letters*, vol. 53, no. 23, pp. 1552–1554, 2017.
- [8] S. Said, W. Saad, M. Shokair, and S. El-araby, "BER performance of signal detection for massive multi user spatial modulation systems," *International Journal of Wireless Information Networks*, vol. 27, no. 3, pp. 484–493, 2020.
- [9] X. Liu and J. Zhang, X. Tong, G. Cheng, Y. Dian, and Z. J. Shu, Barzilai-Borwein based signal detection algorithm for massive MIMO," *Systems Engineering and Electronics*, vol. 40, no. 8, pp. 1861–1865, 2018.
- [10] X. Tan, Y.-L. Ueng, and Z. Zhang, "A low-complexity massive MIMO detection based on approximate expectation propagation," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 8, pp. 7260–7272, 2019.
- [11] Z. Albataineh, "Low-complexity near-optimal iterative signal detection based on MSD-CG method for uplink massive MIMO systems," *Wireless Personal Communications*, vol. 7, no. 4, pp. 1–15, 2020.
- [12] C. D. Andrea, S. Buzzi, and M. Lops, "Communications and radar coexistence in the massive MIMO regime: uplink analysis," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 19–33, 2019.
- [13] K. K.-C. Lee, Y.-H. Yang, and J.-W. Li, "A low-complexity AEPDF-assisted precoding scheme for massive MIMO systems with transmit antenna correlation," *Journal of Signal Processing Systems*, vol. 92, no. 5, pp. 529–539, 2020.
- [14] T.-B. Nguyen, M.-T. Le, and V.-D. Ngo, "Low complexity lattice reduction aided detectors for high load massive MIMO systems," *Wireless Personal Communications*, vol. 109, no. 3, pp. 1805–1825, 2019.
- [15] M. Karlsson, E. Bjornson, and E. G. Larsson, "Techniques for system information broadcast in cell-free massive MIMO," *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 244–257, 2019.
- [16] Z. Mai and Y. Chen, "A novel fast linear iteration detection algorithm in MU-massive MIMO systems," *International Journal of Wireless Information Networks*, vol. 26, no. 3, pp. 1–6, 2019.
- [17] M. Khani, M. Alizadeh, J. Hoydis, and P. Fleming, "Adaptive neural signal detection for massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5635–5648, 2020.
- [18] Y. Xue, Z. Wu, J. Yang et al., "Adaptive preconditioned iterative linear detection and architecture for massive MU-MIMO uplink," *Journal of Signal Processing Systems*, vol. 90, no. 10, pp. 1453–1467, 2018.
- [19] K. Long, V. Leung, H. Zhang, et al., Y. Li, Z. Feng, Complexity analysis of massive MIMO signal detection algorithms based on factor graph," *5G for future wireless networks*, vol. 10, no. 24, pp. 246–256, 2018, [Lecture notes of the institute for computer sciences, social informatics and telecommunications engineering.

Retraction

Retracted: Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] B. Sumathy, A. Kumar, D. Sungeetha et al., "Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 5906797, 10 pages, 2022.

Research Article

Machine Learning Technique to Detect and Classify Mental Illness on Social Media Using Lexicon-Based Recommender System

B. Sumathy,¹ Anand Kumar,² D. Sungeetha,³ Arshad Hashmi,⁴ Ankur Saxena ,⁵ Piyush Kumar Shukla,⁶ and Stephen Jeswinde Nuagah ⁷

¹Department of Instrumentation and Control Engineering, Sri Sairam Engineering College, Chennai, India

²School of Computer Science and Applications, REVA University, Bangalore, India

³Department Electronics and Communication Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India

⁴Department of Information Systems, Faculty of Computing and Information Technology in Rabigh (FCITR), King Abdulaziz University, Jeddah, Saudi Arabia

⁵Indus Institute of Information & Communication Technology, Indus University, Ahmedabad, Gujarat, India

⁶Computer Science & Engineering Department, University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, (Technological University of Madhya Pradesh), Bhopal 462033, India

⁷Department of Electrical Engineering, Tamale Technical University, Tamale, Ghana

Correspondence should be addressed to Stephen Jeswinde Nuagah; jeswinde@tatu.edu.gh

Received 10 November 2021; Accepted 17 December 2021; Published 26 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 B. Sumathy 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 emergence of social media has allowed people to express their feelings on products, services, films, and so on. The feeling is the user's view or attitude towards any topic, object, event, or service. Overall, feelings have always influenced people's decision-making. In recent years, emotions have been analyzed intensively in natural language, but many problems still have to be watched. One of the most important problems is the lack of precise classification resources. Most of the research into feeling gradation is concerned with the issue of polarity grading, although, in many practical applications, this relatively grounded feeling measure is insufficient. Design methods are therefore essential, which can accurately classify feelings into a natural language. The principal goal of the research is to develop an overflow of grammatical rules-based classification of Indian language tweets. In this work, three main challenges are identified to classify feelings in Indian language tweets and possible methods for tackling such issues. Firstly, it has been found that the informal nature of tweets is crucial for the classification of feelings. Based on the tweets, the mental illness of the person has been classified. Therefore, to categorize Indian language tweets, a combination of grammar rules based on adjectives and negations is proposed. Secondly, people often express their feelings with slang words, abbreviations, and mixed words. A technique called field tags is used to include nongrammatical arguments such as slang words and diverse words. Thirdly, if a tweet is more complex, the morphological richness of the Indian language results in a loss of performance. The grammar rules are embedded in N-gram techniques and machine learning methods. These methods are grouped into three approaches, which functionally predict Indian language tweets with syntactic words.

1. Introduction

The increase in social media and users' numbers has made it possible to express one's opinion in natural language. Social media feeling analysis in recent years has been an active field

of research. A model needs to identify different social media [1] users' dimensions of feelings to analyze this natural language management system. The detailed review of sentimental [2] models of analysis shows that the study can assist the user to classify the operator's feelings on a theme.

Analysis of the emotions is used to find user feelings or opinions. An individual has his own space in social media, such as Twitter, to post an idea or topic or comment on a service. The user review shows that various models of sentiment analysis in natural languages have been developed, film reviews, product reviews, political reviews, and so forth, for feelings analysis. In Twitter, research is being conducted extensively to predict the public mood used in different fields and applications. The classification of sentiment, in general, is divided into 3 types: (i) approach to master learning, (ii) a hybrid teaching approach, and (iii) analysis of sentiment requiring a detailed analysis of techniques of natural language processing, so that training datasets for machine learning and feeling lexical data can be provided for statistical or semantic methods. This study is aimed at developing the user's sensation framework for Tamil tweets. Nonnative English speakers have been highly influenced by social media such as Twitter. There are different discourse challenges for nonnative English speakers when expressing an opinion on social media. The first challenge is to develop grammar rules for classifying feelings in Tamil tweets. The second problem is that there are insufficient resources, such as dataset and feeling lexicons. The last question is to improve slang words' performance, words transliterated in various languages and fields.

It is not very easy to precisely identify user feelings from domain to domain by this domain-dependent word. Based on these hypotheses, the research will be validated.

Hypoproposed work 1: concerns about the inclusion of syntactic methods for the necessary results for further improvement.

Hypoproposed work 2: the rule-based grammar approach can better represent tweet feelings.

Hypoproposed work 3: grammar rules combined with the supervised master method of learning improve results.

The following general and specific objectives have been identified to address the challenges linked to the above-mentioned research problems. Techniques are developed for the classification of Tamil tweets based on grammar rules. Besides, this paper proposes the principal component of the sentiment analysis scheme. The proposed regulations on language grammar for Tamil tweets' classification are a characteristic feature by which user feelings are identified, and tweets are grouped into a set of categories. The work proposed contributes to new grammar rule-based algorithms for the Tamil tweets. These grammatical rules are relevant to user tweet categorization. The main tasks are to reach the elemental powers in classifying tweets and are also generic enough for various fields and systems. The Tamil tweets classification is further developed by incorporating syntactic measures such as domain-specific and tweet tags. The main idea is to add domain words to the user phrase to improve the performance of classification. The work focuses on the variety of gender instead of polarity-based systems. There are few types of research on the sort of user texts by genre. The adjective grammar shows the way for sentiment analysis in a language like Tamil. Although the Tamil language

has complicated instructions, this proposed work invokes only negative guidelines and procedural regulations to categorize tweets into dissimilar categories. The planned grammar rules focus on adjectives, negatives, and connection words only to deal with ungrammatical tweet structure. This paper has proposed a new model combining syntactic, semantic, and supervised methods of learning. In general, the work is more accurate than the existing systems—also, the model is more exposed to different areas and comparison of results.

The purpose of this effort is to propose a new method for Twitter sentiment analysis, which is divided into two stages. First and foremost, there is the tweet jargon, which includes emoticons and other symbols. The emoticons are converted to plain text by using processes that are independent of the language being used. Alternatively, it is readily adaptable to multiple languages. Second, the generated tweets are categorized based on their subject matter. BERT is a language model with the advantage of being pretrained on plain text rather than tweets. The models are based on plain text and are readily accessible in various languages, reducing the need for time and resources to create them. The following advantages are accessible: (1) models may be trained directly on tweets from scratch and (2) available plain text corpora are bigger than tweets only corpora, allowing for higher performance. A case study detailing how the technology was put to use the approach for Italian is provided, along with a comparison to various Italian options that are currently available. Findings demonstrate the efficiency of the technique and suggest that, as a result of its basic foundation from a theoretical perspective, from a methodological standpoint, it has the potential to be useful for other languages as well.

2. Literature Survey

The terms exchange of opinion mining and sentimentality examination are used by most of the current approaches. Opinion mining is defined as identifying the emotional tone underlying a piece of text using natural language processing (NLP) [3].

$$\text{Mining based on Option} = (u, t, i, J). \quad (1)$$

In the above formula, “ u ” is the view objective, “ t ” is the opinion on the goal, “ i ” is the view owner, and “ J ” is the period once the idea was published. It is essential to note from the above definition that feeling mining belongs to the opinion mining sector. Sentiment mining may be of a binary type or theme detection. The term sentiment analysis (SA) is used as the term for classification tasks in this proposed work. A concept of feelings analysis was first introduced in [4]. In sentiment examination, there are three main classification methods: machine learning methods, lexical methods, and hybrid methods [5, 6]. The applied classifier of feelings depends on the data annotated. Usually, these training data are derived from function words to categorize novel information. The results of the classification machine are based on methods of functional selection. Most of the trainings already conducted focus on machine learning. Despite the numerous machine learning methods used in most of the research

studies, the supporting vector machine and the Naïve Bayes classifications are standard. Therefore, the current works related to SVM and NB classifiers are reviewed in the following sections and followed by lexical sentimental analysis procedures.

2.1. Emotion Analysis Using Machine Learning Algorithm. When the choice of feature words or courses is an essential part of using classifiers for sentiment analyses, appropriate plan of contextual features can provide more information and reduce noise opportunities. To do so, various sources of characteristic manufacturing are frequently employed [7]. The SVM is the best method for machine learning [8] to combine several domain model knowledge characteristics, syntactic reliance, previously annotated sentences, and adjectives with standard text characteristics for a performance of 86.0%.

The method of classification of polarity through machine learning algorithms was proposed in most of the reviews [8, 9]. In most of the research projects, SVM is clear to the literature [4, 10, 11] because they are robust and efficient in the analytic sentiment of highly dimensioned information. The authors of [12] considered a new algorithm to find no more than 25 video and audio genre classifications. The videos with these features are classified by SVM. The authors did not take into consideration the high dimension of genre classification.

2.2. Mental Illness Detection Using Lexicon-Based System. The development of lexicon or SentiWordNet is an essential work in the lexical way to describe the “structure that holds information about words and synonyms or related meanings.” The total user sentence or text polarity is then calculated with this lexicon or WordNet with an A-weighted number of all lexical components [13, 14]. Lexicons are built using polar or emotional words. Furthermore, these opposite terms are divided into two or probably three groups, based on their divergence to construct the lexicon (positive, negative, and neutral). For lexical sentimental analysis, lexical resources and knowledge are required in a particular field. The feelings of a given text or review are calculated using the lexicon based word or phrase polarity. Unigrams or N-grams are used for training classifiers in most of the machine learning algorithms [15]. However, unigrams are used in lexicons to assign polarity; therefore, the total value of the complete text polarity is calculated as a unigram. The hybrid approach finally combines both machine learning based processes and lexicon based procedures. Moreover, a method known as a linguistic rule is usually associated with a classification of lexical sentiment [15, 16]. Some research related to hybrid approaches works specifically in a variety. To identify the hybrid method, syntactic features such as word expressions and denials, as well as the structure of the original document, are used [17]. Parts of speech (POS) are methods of identifying grammatical categories of words used in the linguistic based approach. Various POS patterns or targets may be used as functions for the sentence. POS tags are combined of substances, adjectives, or verbs. These tags can then be used to specify a specific polarity

or feeling topic. Natural languages other than English have been widely used recently on social media platforms, such as Twitter and Facebook. Analysis of emotions in foreign languages has grown since a few research projects have been underway to create language resources [18, 19]; for example, in Chinese, Arabic, Hindi, and Tamil, SentiWordNet, which does not exist in English, is the most common resource. However, there is still a lack of resources for sentiment analysis tasks in many natural languages. Still, in sentimental analysis, English is the most widely used language because well-defined resources, such as lexicons, corpus, and dictionaries, are present. In particular, Tamil was used more frequently on Twitter. Researchers must face new challenges to build resources like lexicons or SentiWordNet and natural languages corpus and dictionaries. As a result, there are specific resources available for these languages because research in this area is still lacking. Many linguists and researchers are now developing natural language resources. The use of the NLTK can be taken into consideration to support the SA process. It helps to understand the natural language characteristics such as Tamil and can contribute to a more accurate SA performance. While the use of NLTK is problematic, this gives a new challenge before the SA process to incorporate NLP. A few NLP tools for SA’s natural languages task have recently been developed. The literature reveals the availability of various methods of machine learning for the analysis of feelings. Also, all current investigations on sentiment examination put emphasis on the organization of divergence. As for Tamil tweets, a lack of resources is the main problem in this area. To categorize the sentimentality of Tamil movie tweets, a semantic method can be practical. Three key subjects are finally examined to improve the classification of emotion in Tamil tweets: first, the method for field tags, second, the use of grammar rules by film reviews using syntactic and semantic models, and, third, the machine knowledge methods. A new grammar procedure for the proposed work predicts the Tamil genre class tweet.

3. The Methodology of Proposed Framework

The general framework for the Tamil tweet sentiment analyses using various algorithms is explained. The design of the systems underpinning the work is first described, and then every phase of the plan related to this work is presented. Section 3 provides a note on the structure and justification for the creation of the Tamil language. A short description of the film’s genres follows/product review classifications and accuracy metrics. Figure 1 shows how the Tamil tweets classify their sentiment.

3.1. Proposed Architecture of Proposed Model. Four steps of user tweets’ feelings in Tamil movies are taken for identification. The first step is to collect and prepare tokenizers for all user tweets. The next process is to detect parts of voice tags with tokenized keywords. Finally, the tokenized content procedures will be used to identify the genre category using the natural language toolkit.

Figure 2 shows the general procedure included in this framework for sentimentality analyses.

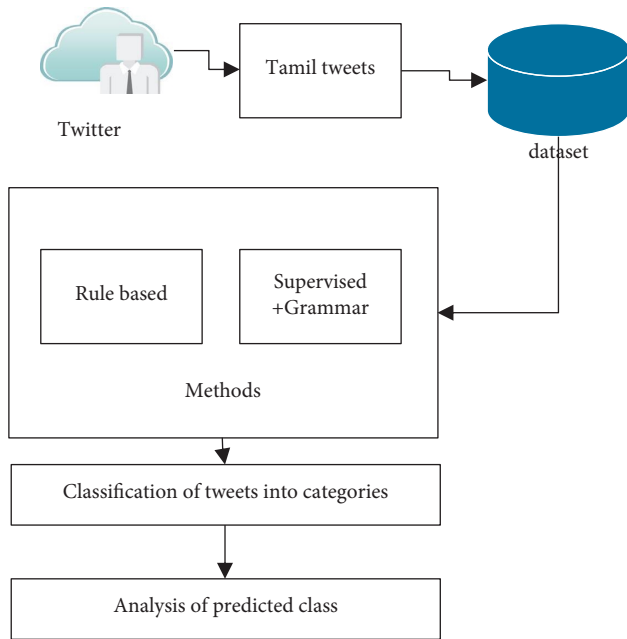


FIGURE 1: Sentiment analysis framework for Tamil movie tweets.

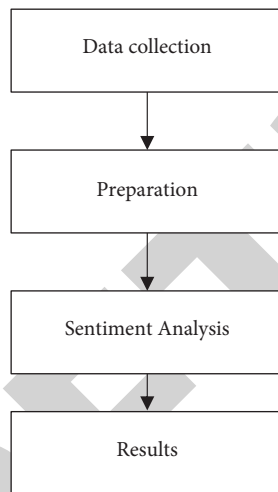


FIGURE 2: Flow chart of proposed work.

3.1.1. Input Data Collection. The primary stage is to gather the data needed for classification—sentiment analysis. There are in the area of sentimental analysis different well-established datasets available in English and the related domain. For natural languages other than English, only limited datasets for feeling analysis are available. In this research, all datasets are extracted from Twitter using the hashtag (#) and then the movie/product’s name using Twitter’s API. However, there is no predefined dataset for Tamil films; an unlabelled dataset for experimental analysis is a significant task.

The last week in July 2016 includes all of the Tamil film tweets used in this proposed work. 100 Tamil films and product tweets were collected (mobile phone). Initially, the

idea was to create a film dataset only but only for the sensation framework to prove that two datasets are made independent of the domain. The body contains 7,346 tweets from Twitter which have been collected and used for all purposes.

3.1.2. Preprocessing Task. The next step is the pretreatment of tweets. To remove conflicting, imperfect, and luminous information, the preprocessing of data is done. To perform all data mining functionality, data needs to be preprocessed. The first job is to delete URLs. Usually, the Uniform Resource Locator does not help in informal words to assess the feeling. For example, take the phrase “I logged on <https://www.amazon.in> as the film is boring.” This phrase is harmful because it is wrong and can become neutral of the amazon text’s occurrence. A technique for removing the Uniform Resource Locator is used to avoid such errors. The following task is to remove retweets. Retweeting is the process of copying a tweet and posting it to a second user. This is usually if a user likes another user’s tweets. Retweets are frequently abbreviated as “RT.” These retweets are redundant data to remove all retweets.

3.1.3. Tokenization Process. Tokenization is a way of dividing words into different words or tokens into user tweets. A phrase, word, or symbol might be a token. The tweet phrases are tokenized into a series of words that can be analyzed with white spaces to remove any specific character or punctuation marks such as # and @. The various Documentary Dictionaries are called token sets produced by combining the full text of a collection.

3.1.4. Sentiment Analysis Models. Supporting vector machines are commonly used to detect sentiment topics on a document level, unchecked approaches like Naïve Bayes [20]. But more advanced models, such as the linguistic rules, are required to categorize the (polarity) opinions and sensations of informal text (gender). The suggested sentiment framework is divided into three functionality-based models. Figure 3 shows two types of feelings investigation.

3.2. Tools. Software access to Twitter is needed to create a tweets corpus. The Twitter REST API is used in this research to access corpus user tweets. This API also provides developers with access to all public tweets and their associated metadata to search for and download streams. However, access to data from the Twitter API is restricted. Authentication methods (OAuth) are used for user prevention misconduct. Various programmers can also use it to understand the use of the API. To access all tweets in real time, Twitter “Firehose” is the only way. Access to the Twitter “Firehose” generally comes from third-party (GNIP and DataSift) managers, although it is not free of charge. The costs of subscription for individual scientists working in sentiment analysis are very high. The streaming API can nevertheless be used to access tweets in real time with a particular number of Twitter data requests. For sentimental

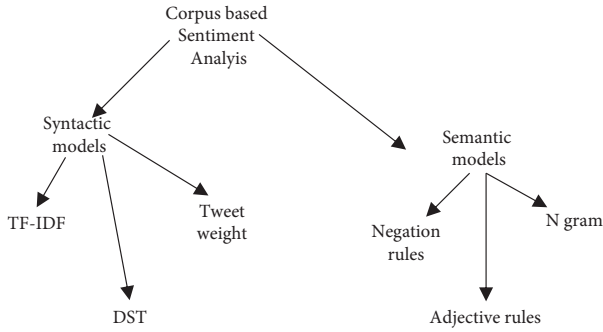


FIGURE 3: Three types of sentiment models.

analysis, the Tweepy Python API is employed to collect Twitter information. If a user wants to use Twitter API directly, the TWIP connection is relatively complex. It enables user-friendly search and download functions. Usually, the relationship with the routine is established.

3.2.1. *Twitter API.* The NLP has developed a portion of the speech tags to classify the words according to their POS. A portion of the speech tagger helps in the analysis of the feelings for the two reasons: (1) it may be used to differentiate words that are not generally felt in POS and (2) words such as pronouns and nouns can be used with POS.

The classification task POS tag in [21] is used in this proposed work. The Python Requests framework was developed to manage HTTP requests using the POS tagger. Figure 4 illustrates an example of a tagged tweet in that POS tagger.

The TF-IDF for document or word classification is a simple unigram model. TF-IDF works well in the classification of documents, such as news articles or reviews [22].

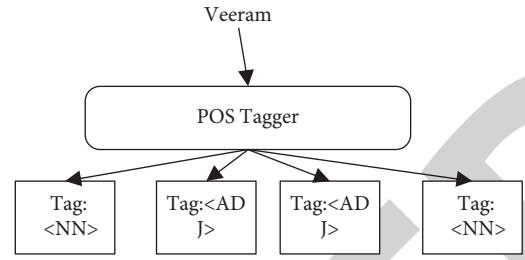


FIGURE 4: Input tagging.

The literature shows, however, that TF-IDF does not classify tweets as well as long tweets and does not follow grammatical styles, and general words are seldom repeated. Tweets, however, contain valuable information for the extraction of feelings. As a basic model, TF-IDF is chosen; it gives the meaning of the word in a dataset. Set of words in tweets should correspond to the subjects and the most frequently reported words should be obtained. To classify the tweets [23] into a set of data, the top N TF-IDF keyword values of each film are selected. Consider a film mi that is linked to a set of tweets $\{t1, t2, \dots, Tn\}$, where Tn is translation. There are several terms in each tweet which allow each film to be marked.

The group of words $x1, x2, \dots, xk$ are like a tweet sequence.

$$mi = \{u1, u2, \dots, un\} = \{x1, x2, \dots, xk\}. \quad (2)$$

Also, every movie can be considered as a tweet group. Their associated tweets contain a set of words, tf, mi , and idolatrous (wj, m).

The values are calculated as follows:

$$tf(x_j, n_j) = \frac{\text{input co-occurrence frequency of } x_j}{n_i}$$

$$idf(w_j, m) = \log\left(\frac{\text{overall emotional tweets of the person } (n)}{\text{input co-occurrence frequency of overall tweets}}\right). \quad (3)$$

Take a movie “Veeram” with 305 tweets as an example. Where the word is twenty times (Rasool) and the word is 18 times.

Total dataset tweets are as follows:

$$tf_{rate} = \frac{20}{305} = 0.0656,$$

$$idf_{rate} = \log\frac{305}{18} = 1.229, \quad (4)$$

$$tfidf_{rate} = tf * idf = 0.0806.$$

Similarly, the TF-IDF is calculated and the important keywords can be identified with the score for all the genres that correspond to the domain. The synonyms of the selected

keywords are also mapped with the Tamil dictionary to reduce dimensionalities and improve the performance of the category and word model [24].

4. Algorithm for TF-IDF

The sentiment categorizer model is used after data collection and preprocessing to achieve the following in the baseline model:

- (1) Divide all tweets into keywords or tokens.
- (2) Identify the occurrence and related words in the tweet for each keyword.
- (3) For every keyword selected from a user’s tweet, compute the TF-IDF score.

- (4) Multiply TF-IDF with SentiWordNet's Tamil score.
- (5) Set all user tweets with mean precision in the overall sentiment scores.
- (6) Assign a film by Step 5 to genre and polarity.

The image of user tweets [24] for the sentiment classes is shown in Figure 5 (Veeram). The result shows the proportion of polarity and gender tweets. Although Action and Trade point were verified by domain professionals, the film genre class has shown that it is categorized into a comedy genre (23, 60%) and love (23.88, 21%). The TF-IDF approach relies on a unigram or perfect keyword and categorizes a tweet only when the keyword is available. The TF-IDF model also does not take into account the user tweet context.

4.1. Algorithm for Genre Classification. The syntax parser determines a tweet's overall polarity and gives this score a perception categorizer model to determine a tweet's type class. Figure 6 shows the algorithm used to classify sentiments.

The tweets are identified with the syntax parser and POS taggers result using the above algorithm in the Categorizer Sensitivity Model [25]. The tweet is classified in an adjective way into the closely related class when the polarity of a tweet is positive. Extraction from parser-based negation produces a greater accuracy than syntax models of 47.32 percent. The rules on negation are designed to improve the analysis of feelings. The findings show that the model of grammatical negation dependence has a higher level of efficiency in sentiment analysis compared with frequency and other syntactic models [26]. The results show that the variety of evolving user-generated text needs to be dealt with throughout the grammar rule approach.

4.2. Adjective-Based Grammar Rules for Semantic Model.

This work is hypothesized by the adjectives as the principal semantic structure for the classification of film genres. Most of the Tamil grammarians speak only of substances and verbs. It is said that adjectives are not considered as separate categories in Tamil by traditional grammarians or linguists. Adjectives are used for the description or quantification of a noun object. Adjectives differ in occurrence and function in different languages. In modern Tamil, adjectives are mostly written just before the substrate. A pattern of adjectives concerning Tamil film tweets is identified by this method [24]. These adjective patterns are linked directly to the specific domain posts and thus rules for finding an impression of the specific posts have been developed in this context. Adjectives are also used to express the strength of user feelings by intensifying them.

Grammar modification is as follows: temporary values $\{+2\}$ in the application of rules 2 and 3 temporary scores are derived. Upon application of the rules, the initial results are not altered because the tweet has no other opinions or terms of denial. Final score is $\{+0.66\}$.

The endpoints are standardized between +1 and -1. The final result is calculated with two divided by three, for example, in this case.

$$\text{Final score} = \frac{\text{Adjective score of the input data}}{\text{Analysed words in the sentence of the system}} \quad (5)$$

The match calculation undergoes by an action point. The category match calculation: Action (+).

4.3. Supervised Model. The characteristics of the classifiers should be extracted for machine learning classification. Functional vectors affect the classifiers' performance [27]. Two methods of extraction, character presence and character count, are generally used specifically. Character count uses the count of frequencies (if the count of frequencies is high, the word is considered to be the word character), while the character presence uses the characteristic word's presence or absence. Although tweets are short, this work uses the function presence method for the extraction of functions. The first five (unigrams) adjectives correspond to the initial seed list for each genre [28]. The seed list synonyms and antonyms are derived from a software programme, Tamil WordNet. This process continues until all functional adjectives have been added to the functions list. Table 1 represents the kernel adjectives.

At first, 500 corpus tweets are selected manually to train the classification. Each tweet will extract the words of the function (93 adjectives in this book) from the list and other words will not be taken into account. Similarly, for instantiation of SVM classification, the NLTK library file is used. It is noticed that several tweets occur repeatedly through multiple posts of the individual user in the corpus. There are also some tweets with misleading feelings or feelings about the specific field. The performance of the classifiers will degrade if such tweets are selected as a set for training. The experiment is conducted with NLTK library files for both classifications. For both classifiers, 10-fold cross-validation is made.

Computational learning theory is behind the support vector machine (SVM) machine learning technology. SVM's main purpose is to find the most efficient classification function for categorizing the training dataset's classes. To handle linear and nonlinear classification issues like density estimation or pattern recognition, the SVM model is commonly utilized. Translate the training data into a higher dimension using nonlinear transformation, and then divide it into separate training sets using the linear method.

A kernel function K is replaced for the intermediate product (X, Y) in a nonlinear SVM classification model (X, Y) :

$$f(x) = \text{sign} \left(\sum_{j=1}^W b x_j K(x_j, x) + c \right). \quad (6)$$

In the learning process, SVM employs a two-layer structure. It is the initial layer that selects the kernel's base $K(x_i, x)$, where i is one of 1, 2, 3, 4, 5, or 6. Layer 2 is a linear

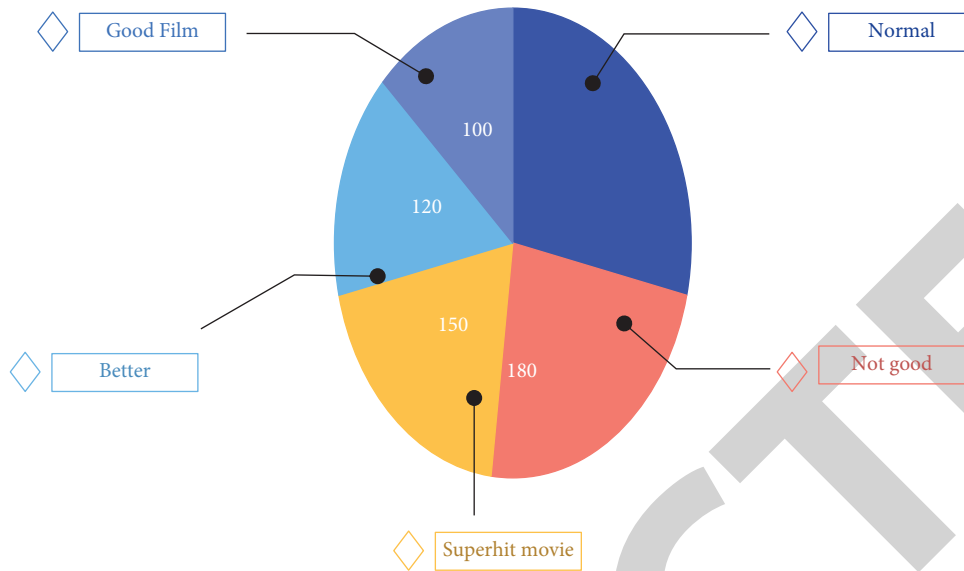


FIGURE 5: Sentiment model using TF-IDF.

```

Calculate Tweet_Category Returns Category {
    Read TweetStream;
    For each Tweet in TweetStream {
        Read the Polarity values of each adjective in a Tweet from
        ParseTree(Tweet);
        If Polarity (Tweet) = (Positive) {
            Get the Adjectives of each Tweet from POS tagger;
            then
            Category (Movie) += Category(Adjective) }
        Else {
            Category (Movie) == Category(Adjective) }
        If (Adjective is not related to any category) {
            then, discard;
            Return Category }
    }

```

FIGURE 6: Algorithm for genre classification.

TABLE 1: Instance list of early kernel adjectives and their synesis.

Initial seed list	Synonyms	Antonyms
Action	“Fight,” “action,” “veeram”	“Peace”
Love	“Love,” “romance”	“Sogam”
Commercial	“Vasool,” “mass,” “commercial”	“Ioss”
Comedy	“Comedy”	“Tragedy”
Family	“Sentiment,” “family,” “feeling”	“Aabasam”

function in the feature space formed by the first layer. Making the best hyperplane in the similar feature space is the same as it was in the previous example. It is generally accepted that hyperplanes with bigger margins are more accurate than hyperplanes with smaller margins when used to categorize feature data. The shortest distance between the hyperplane and the margins on each side is taken to be the hyperplane with the greatest margin. Hyperplane for separating planes is defined by the following equation:

$$W \cdot X + c = 0. \tag{7}$$

Margins are determined as the support vector points. The outcome of the process is the linear combination of all support vector points, and all other data points are overlooked. It comes with the notion that the complexity is not dependent on the number of features existing in the training dataset. It makes SVM very efficient for classification problems that hold a considerable number of features as compared to the number of training examples. The only drawback with SVM is that, in case of misclassified or linearly inseparable data, no separating hyperplane can be obtained. So, the SVM translated the data into higher dimensional feature space and found a suitable hyperplane. In this work, the LS-SVM Lab toolbox has been applied to classify the speech of ID from TD children. To achieve a better classification accuracy, the two regularization parameters, (γ, gam) and $\sigma^2(\text{sig}2)$, which was the squared bandwidth of RBF kernel, have been chosen optimally.

5. Experimental Results

An analysis tool was developed which incorporates all NLTK-based and Python-based algorithms. The tool shows automatically the feeling values of Tamil movie tweets both at the polarity level and at the genre level. Figure 7 shows the feelings for the Veeram film. Table 2 represents the grammar performance.

The results suggest that the general sentiment model based on the grammar rule delivers better performance compared to other syntactic models. The results also show that the precision sentimental analysis increases significantly when somaticized models in addition to normal functionality like unigrams are incorporated (TF-IDF). Compared to other feeling models, the grammar approach proposes the semantic structure of the user’s phrase within the specific domain.

The semantic grammatical model provides an average accuracy of 64.72 percent better than that shown in Figure 8

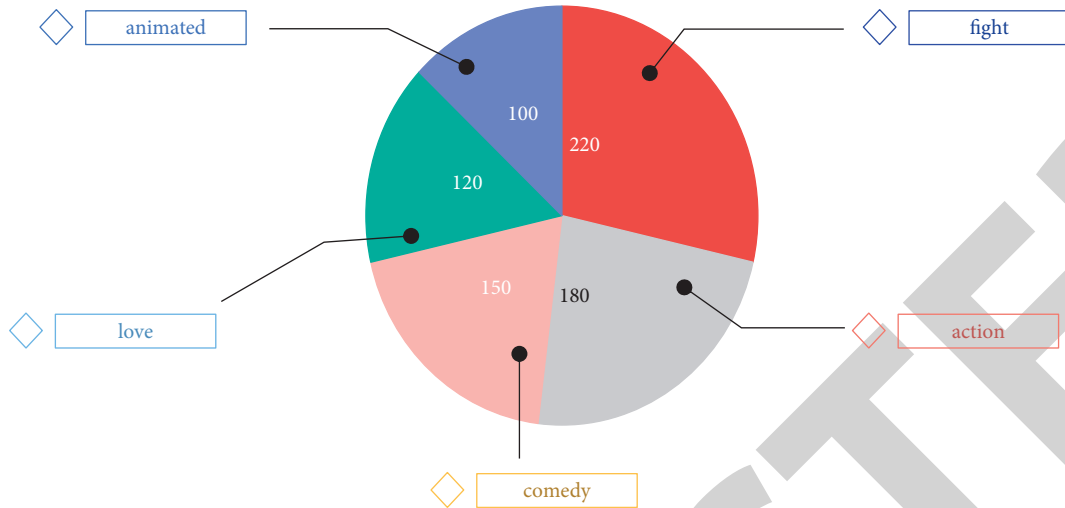


FIGURE 7: Results of feeling analysis by grammar rules.

TABLE 2: Tamil SA grammar performance.

Movie	Method	Accuracy
Veeram	TF-IDF ranking	27.13
	TF-IDF + DST	36.42
	Tweet weight	40.26
	Negation rule	47.32
	Adjective rule	60.84

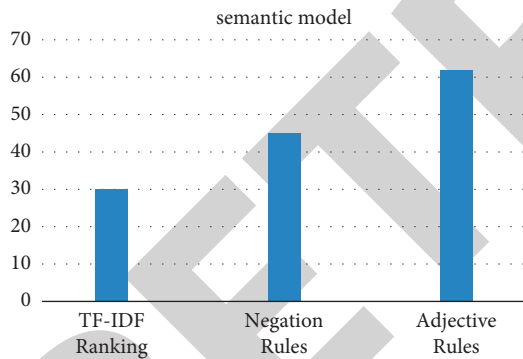


FIGURE 8: Performance comparison of grammar-based semantic models with TF-IDF.

for TF-IDF and other syntactic models. Sentiment model has analyzed tweets and found polarity, genre categories, and other algorithms using proposed grammar rules. Results demonstrate that the general grammar of negative rules and adjective rules is better because complex sentences are taken into consideration and semantic structures are better integrated. The proposed grammar rules address any sort of sentence in tweets to determine sentiments (simple, compound, and complex). The grammar rule-based model with an accuracy of 64.72 percent is the best feeling model. If the results of TF-IDF, tweet weight, and regulatory modeling are compared, the grammar-based algorithm could be found to be 20 percent better than other sentiment models. The

results show that machine learning methods alone are not good for feelings. One of the important lessons of this is that, instead of using the grammar methods, SVM is better than the syntax model. The classifier quality is only as high as the set, so all possible instances cannot be exposed to the classifying system. Therefore, to improve machine study classifier performance, adjective and negative-based grammar rules are used as a feature for classifiers to compare machine study methods with grammar rules. Table 3 shows the SVM classification performance in combination with grammar regulations.

When the grammar rules are combined with the SVM classification, it is determined to outperform all other feelings models in combination with the grammar rules method. Between the grammar rule and learning models, the accuracy changes 7%. The result highlighted again the quality and the ambiguity of Tamil grammar in the grammar-based machine learning model. A good promise is made with grammatical rules for further development. For cross-domain assessments mobile phone reviews, the proposed sentiment framework procedures are adapted. This is because the number of tweets available on this domain is the number of movie domain choices. The aim is to verify the performance of the grammar rules algorithms and methods of machine learning regardless of domain even if the size of the product field tweets is small. The domain-independent features for the training are extracted for master classification. The words that occur in all domains are domain-independent characteristics. This function is important for transferring the semantical context from one domain to another. The grammar rules are used in this research work to extract independent domain adjectives for analysis of cross-domain feelings. Table 4 shows the results of each of the three models of feeling.

The study shows that the algorithms work in a comparable way for different domains. This demonstrates the work's expansion into various areas.

TABLE 3: Performance of SVM classifier with grammar rules.

Movie	Method	Accuracy
Veeram	TF-IDF ranking	27.13
	Domain-specific tags	36.42
	Tweet weight	40.26
	Negation rule	47.32
	Grammar rule	60.84
	SVM + grammar rules	65.73

TABLE 4: Performance analysis for product domain.

Corpus	Method	Accuracy
"Mobile phone"	TF-IDF ranking	29.35
	Domain-specific tags	30.98
	Tweet weight	45.18
	Negation rule	59.97
	Grammar rule	66.82
	N-gram model	65.23
	Naïve Bayes	47.94
	SVM	58.96
	SVM + grammar rules	69.32

6. Conclusion

The present dataset has been applied to the existing algorithms like SVM and Naïve Bayes, and results were tracked. The results show that SVM model could better classify the genre of film compared to syntactic methods. The work thus suggested that both models be combined and the results traced. While the proposed algorithms with the setup of a feeling framework are successful, it is valuable to assess their performance with the system's composition in real time. The proposed model would then be tweeted in real time as part of future work.

The overall model could be changed if the work is carried out in real time. In future work, this is an important direction. If work continues, lexical resources must be developed when this research is extended to more than one area. The focus of this research has so far only been two areas (films and product), and the domain tag resources for these two areas have therefore been developed. Once the grammar models for the complex phrases are completed, the paragraphs can also extend the model. It is also essential to implement a grammar-regulative approach for handling complex and composed sentences as a cause of the error. The future will automatically focus on "generate tags" (types) from the text. The SVM, in combination with the rules of grammar, outperforms all other Tamil tweets in feeling analysis. This is an essential finding of the approach to machine learning. Two product category tweets were used, and the sentiment methods were applied to track the validity of the model in various domains. These results have been validated so that the grammatical techniques are efficient. There has been no significant improvement in outcomes when combining SVM with grammar-based techniques. The other two machine methods can also be tested in future work (Semisupervised and Unsupervised).

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] K. Dave, S. Lawrence, and D. M. Pennock, "Mining the peanut gallery: opinion extraction and semantic classification of product reviews," in *Proceedings of the 12th international conference on World Wide Web*, vol. 16, pp. 519–528, ACM, New York, USA, May 2003.
- [2] M. Taboada, J. Brooke, M. Tofiloski, K. Stede, and S. Manfred, "Lexicon-based methods for sentiment analysis," *Computational Linguistics*, vol. 37, no. 2, pp. 267–307, 2011.
- [3] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lectures on Human Language Technologies*, vol. 5, no. 1, pp. 1–167, 2012.
- [4] Y.-F. Huang and S.-H. Wang, "Movie genre classification using SVM with audio and video features," in *Active Media Technology*, vol. 4, pp. 1–10, Springer Berlin Heidelberg, Berlin, Germany, 2012.
- [5] M. G. Armentano, S. Schiaffino, I. Christensen, and F. Boato, "Movies recommendation based on opinion mining in twitter," *Advances in Artificial Intelligence and its Applications*, Springer International Publishing, vol. 23, pp. 80–91, Midtown Manhattan, New York City, 2015.
- [6] A. Shoukry and A. Rafea, "Sentence-level Arabic sentiment analysis," in *Proceedings of the International Conference on Collaboration Technologies and Systems*, vol. 10, pp. 546–550, IEEE, Denver, CO, USA, May 2012.
- [7] S. Tan, X. Cheng, Y. Wang, and H. Xu, "Adapting naive Bayes to domain adaptation for sentiment analysis, lecture notes in computer science," in *European Conference on Information Retrieval*, Springer Berlin Heidelberg, Berlin, Germany, pp. 337–349, 2009.
- [8] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Proceedings of the Language Resources and Evaluation Conference*, vol. 10, Valletta, Malta, May 2010.
- [9] R. M. Dubai and I. Qarqaz, "Arabic sentiment analysis using supervised classification," in *Proceedings of the International Conference on Future Internet of Things and Cloud*, vol. 10, pp. 579–583, IEEE, Barcelona, Spain, August 2014.
- [10] N. Amolik, A. Jane, M. Bhandari, and M. Venkatesan, "Twitter sentiment analysis of movie reviews using machine learning techniques," *International Journal of Engineering and Technology*, vol. 7, no. 6, pp. 2038–2043, 2016.
- [11] B. Allison, "Sentiment detection using lexically-based classifiers," in *Speech and Dialogue*, vol. 2, pp. 21–28, Springer, Midtown Manhattan, New York City, 2008.
- [12] X. Ding, B. Liu, and P. Yu, "A holistic lexicon-based approach to opinion mining," in *Proceedings of the international conference on Web search and web data mining*, vol. 11, pp. 231–240, Palo Alto, California, USA, February 2008.
- [13] M. A. Romero, J. Castro, and J. M. Zurita, "Lexicon-based comments oriented news sentiment analyzer system," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9166–9180, 2012.

Retraction

Retracted: Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Thilagaraj, N. Arunkumar, and P. Govindan, "Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6785707, 12 pages, 2022.

Research Article

Classification of Breast Cancer Images by Implementing Improved DCNN with Artificial Fish School Model

M Thilagaraj ¹, N. Arunkumar ², and Petchinathan Govindan ³

¹Department of Electronics and Instrumentation Engineering, Karpagam College of Engineering, Coimbatore, India

²Department of Biomedical Engineering, Rathinam Technical Campus, Coimbatore, India

³Ethiopian Technical University, Addis Ababa, Ethiopia

Correspondence should be addressed to Petchinathan Govindan; petchinathan.govindan@etu.edu.et

Received 13 December 2021; Revised 21 January 2022; Accepted 28 January 2022; Published 22 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 M Thilagaraj 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.

Breast cancer is an important factor affecting human health. This issue has various diagnosis process which were evolved such as mammography, fine needle aspirate, and surgical biopsy. These techniques use pathological breast cancer images for diagnosis. Breast cancer surgery allows the forensic doctor to histologist to access the microscopic level of breast tissues. The conventional method uses an optimized radial basis neural network using a cuckoo search algorithm. Existing radial basis neural network techniques utilized feature extraction and reduction parts separately. It is proposed that it overcomes the CNN approach for all the feature extraction and classification process to reduce time complexity. In this proposed method, a convolutional neural network is proposed based on an artificial fish school algorithm. The breast cancer image dataset is taken from cancer imaging archives. In the preprocessing step of classification, the breast cancer image is filtered with the support of a wiener filter for classification. The convolutional neural network has set the intense data of an image and is used to remove the features. After executing the extraction procedure, the reduction process is performed to speed up the train and test data processing. Here, the artificial fish school optimization algorithm is utilized to give the direct training data to the deep convolutional neural network. The extraction, reduction, and classification of features are utilized in the single deep convolutional neural network process. In this process, the optimization technique helps to decrease the error rate and increases the performance efficiency by finding the number of epochs and training images to the Deep CNN. In this system, the normal, benign, and malignant tissues are predicted. By comparing the existing RBF technique with the cuckoo search algorithm, the presented model attains the outcome in the way of sensitivity, accuracy, specificity, F1 score, and recall.

1. Introduction

The most common disease for women is breast cancer. It is cells with extra growth of mass in women's breast region. This breast tissue forms the tumor, which is classified as benign or malignant. The malignant is the most affected cancerous region, and the benign is the non-cancerous region. This disease is diagnosed by biopsy. The researchers analyze various automated diagnosis approaches to determine breast cancer. The stroma maturity of cancer in the breast is classified by the histological image [1]. The breast cancer image of the stroma is the matured result of classification. Thermograph and mammography images are used for this approach. The thermograph images are taken

from cameras, which are analyzed by infrared radiation and its intensity level. By comparing with the thermograph and mammogram, the mammogram image provides the exact result. The breast cancer images are classified using various machine learning techniques and artificial swarm optimization techniques. A capsule network model is performed in the preprocessing stage of breast cancer classifications [2].

The support vector machine learning, extreme learning machine, RBNN, DNN, and other classifiers performed the classification [3]. The optimization algorithms include ant colony optimization, honeybee optimization, PSO, genetic algorithm, and other artificial swarm intelligence techniques. The image processing methods are determined to classify the breast cancer

images. The initial stage of classification in the image processing technique is the preprocessing step, which extracts the edges, filters the image, and improves the image quality. In medical image classification, feature extraction is the major procedure, and it gives data about the filtered picture. The feature selection process is carried out with bioinspired techniques. The classification supports various machine-learning approaches and deep learning techniques [4]. It classifies the normal, benign, and malignant properties of a given medical image. Biomedical images are used for disease classification. The images are ultrasounds, X-rays, computer-aided tomography, MRI, and other images.

In this, the CS optimizer is utilized for the selection process. The RBNN performs the characterization of the given image, which classifies the various properties from the input image [5]. Computer-aided diagnosis approach is a widely used technique for classifying the breast cancer image. It provides valuable information, and it utilizes general image classification techniques. The sampling process is used in the convolutional neural network for solving the computational complexity [6]. The sampling process is carried out in the pooling layer, which is performed with the neurons in the network layer. The LSTM structure is used for data analysis in CNN. An SVM learning technique is integrated with the decision-making algorithm for big data analysis [7]. The histopathological image is utilized to perform the classification and measures various performance metrics for analyzing the classifier. KM clustering approach and the SVM classifier with softmax techniques are utilized for modeling the classifier for medical imaging. The fuzzy-based entropy segmentation of the image processing method is optimized with the fish school algorithm [8]. Various grouping techniques use the artificial fish-school optimization algorithm. The conventional methods use a weka processor for breast cancer analysis [9]. The breast cancer image is used as either mammography or a histological image. This database is analyzed [10]. Computer-aided diagnosis technology is used for detecting breast cancer from images [11]. The CAD system improves the performance of breast classifications such as a benign, lesion, and malignant properties. Here, this tool is assisted by the radiologists in the mammographic screening process. In this, the morphological and texture feature extractions are performed in the segmentation approach [12]. Digital breast tomosynthesis is the three-dimensional image processing modality, which is developed with mammography and histological images. Based on the gradient field analysis, the true mass is measured for 3D images. The lesion property of breast cancer is analyzed using CAD technology. The softmax and ReLU function performance is performed in the CNN for breast cancer classification. The breast cancer dataset is available in open source, and it studies the breast cancer classification problems. The database has 7,909 samples with different classes. In these classes, benign and malignant are the most important properties. The 1024×1024 pixels of breast cancer images are taken for

this proposed analysis. The training process in CNN uses the stochastic gradient descent technique.

The image augmentation process has different properties performed with various classes in the morphological approach. This image augmentation process performs the cropping, image resizing, enhancing, and random patches. Reference [13] has presented the DCNN for the classification model of breast cancer with the IRR approach. The pooling layer of the convolutional neural network reduces the spatial image size, and it performs downsampling. Based on the width and height of the image, the pooling layer gets a maximum pixel value of the image. The bioinspired techniques are applied in the image processing technique, especially to the classification approach. The bioinspired technique used for the application of breast cancer classification is ant colony optimization, genetic algorithm, PSO, etc. These methods improve the performance of feature selection. This optimization technique reduces the problem of overfitting and multiclass problem [14].

In the machine learning approach, the multilayer perceptron and the SVM are performed with two layers of the training process [15]. This approach uses the sequential minimal optimization algorithm, which depends on the performance of c , which means fuzzy logic. In the training process, the fuzzy rule is set for performing the feature extraction. The Kernel function is used to enhance the support vector machine learning technique. The k-means scheme is used in the feature extraction process. GLCM provides the result of energy, entropy, correlation, contrast, and homogeneity in the feature extraction process [16]. The breast cancer classification uses the machine learning technique [17]. The tissue detection, prognostic, and diagnostic are performed with different characteristics. The cross-validation technique is the important process used for partitioning the training and testing function [14]. The RBFNN is used for the flood forecasting technique, which is improved with the cuckoo search algorithm. The performance of the polyharmonic function has lesser error compared to the Gaussian approach. The deep learning technique is integrated to enhance the performance of CNN, which optimizes with the swarm intelligence approach [18]. The input layer of CNN takes the operation of preprocessing. The hidden layer analyses the feature of the given breast cancer images and the completion of pooling layer performance. The fully connected layer in the CNN will act to provide the output function. The direct preprocessed images are taken to separate the various properties of breast cancer. Based on the capsule network layer, the classification performance is worked. In this presented model, the CNN-based artificial fish school model is utilized for the classification of images in an enhanced manner. The normal, benign, and malignant are classified with optimum results. Here, this proposed method improves the accuracy, sensitivity, and specificity. The proposed image classification technique is used for various medical image processing and analyzing methods, and it gives the best result as compared with the existing work.

2. Literature Survey

Reis et al. [1] presented an automated image classification approach in the application of breast cancer analysis. The tumor tissue is classified into different classes. Here, the hematoxylin and eosin patterns are analyzed with the pathologist approach. The breast cancer prognostic method uses the normal picture, and local binary patterns are integrated with the RFC tree approach. Based on the tumor grade, the degrees are arranged with different types of breast cancer. Invasive breast carcinoma is the common breast cancer type, which mostly affects ladies. Mature, immature, and intermediate categories are analyzed in the stroma tissues. An ROI selection process is based on the input database, and it traces with randomly paired ROI. This provides the simultaneous result to observe the training and testing approach. The feature vector provides mature and immature properties. Here, the single-scale and multiscale properties are analyzed with mature and immature features.

Reference [2] illustrated the characterization of breast cancer images in the capsule network model. Here, the preprocessed breast cancer image with a histological property is used. The classification utilizes various systems such as ResNset, InceptNet, and AlexNet. These models are enhanced with the capsule network model, which classifies the breast image. It takes the direct preprocessed image for classification. In this approach, the benign, normal, in-situ, and invasive cancer type are classified. The stain normalization is the property of a patch reduction approach, which is used to avoid adverse effects on the staining. The input layer of the convolution type is created to perform the preprocessing stage. The capsule layer has primary and secondary sections for dimensional reduction. The dense layer is performed in the classification process. Here, the parameters are tuned to get the best preprocessed image for classification. The accuracy of both train and test layers is measured and compared with the conventional technique. Iteration range is increased for accuracy and loss prediction.

Reference [13] presented the convolutional neural network model for breast cancer. The inception recurrent residual network model is used in the CNN to characterize the image. Here, the DL approaches with CNN; it gives the robust execution with inception and residual networks. Here, the BreakHis dataset is used for classification. In the preprocessing stage, the original image improves the quality by resizing, removing noise, image augmentation, and patching. After that, it performs the feature extraction process. Finally, it performs the feature selection process. The limitation of this technique is the design complexity. The results are measured to analyze with various conventional methods. The accuracy, ROC, AUC, and global accuracy are calculated, and they classify different types of benign and malignant features.

Reis et al. [1] have introduced the classification technique and used the bioinspired techniques. Here, various machine learning approaches are modeled to separate breast cancer, and the optimization technique is used in the feature selection process. The image types such as the histological image, pathological, and thermography images are used for

breast cancer classifications. Hence, the SVM technique is utilized for classification, and it is compared with utilizing different bioinspired optimization techniques, that is ACO and PSO. Initially, the thermography images are applied for the preprocessing stage, which is performed with grey-scale imaging, and it reduces the noise since it improves the input image quality. The feature extraction process used the GLCM process. The feature selection technique utilizes bioinspired techniques such as ACO and PSO. Finally, it results in the classification using SVM. The accuracy, sensitivity, specificity, and recognition time are measured to obtain the resultant parameter.

Reference [3] analyzed the machine learning approach for classifying breast cancer images. Here, the machine learning scheme is used as a classifier, and it optimizes the Bayesian classifier. This method is utilized to identify and prognostic the breast cancer tissues. Here, the binary classification scheme is applied in the Bayesian classifier. In the preprocessing stage, the edges are removed to enhance the image quality. The feature extraction process is carried out with KNN, and it also performs the clustering. This approach finds different classes in the testing and training procedure. It helps to separate the features of the image. The reduction approach is employed to select different features, and it uses the cross-validation technique. The precision of both training and testing is calculated to compare with the existing work. Without using the optimization technique, it makes the reduction process inefficiency.

Reference [19] proposed the artificial fish school optimization algorithm for the grouping approach. The common method of grouping strategy is the deterministic finite automaton technique. However, it searches only the local environment since the artificial fish-school optimization approach is used, and it is used for global search. The grouping strategy uses the heuristic algorithm with intelligent optimization. The behavior of the artificial fish is used for the global optimization technique. The encoding process is grouped with the 2D array model, and it is represented with the different group expressions. The interaction rate of the artificial fish is used for the grouping algorithm. The increased amount of food in the area is searched by prey behavior. The drawback of grouping is reduced by the behavior of the swarm. The random search movement is selected by a fish interaction rate. By comparing with other optimization algorithms such as PSO, ACO, and GA, the artificial school optimization algorithm achieves the result.

Reis et al. [1] have described the optimization technique for performing the training process in the DCNN. Here, the swarm intelligence technique is used for optimizing the training process of the DCNN. Here, this optimization technique uses the artificial fish school optimization algorithm. In this approach, the text patterns are used as a dataset. This text mining process is analyzed with a random search. This text mining approach is also applicable for various statistical data science tasks. The vector optimization is performed by using FSS. The preprocessing of input text uses word2vec, and the final layer of CNN provides the statistical information. After performing training and testing, the precision of the text-mining task is calculated to

obtain the result. The technique performs the maxpooling and softmax in the feature extraction process of CNN, and it normalizes the functions by splitting the classes. The text is trained to provide the vector output, and it is optimized using FSS. Then, it performs testing to predict the accuracy.

Reference [20] presented the artificial fish school optimization technique for optimizing the functions. Here, the global optimum value is selected for function optimization. This method is compared with GA, ACO, and PSO. The function optimization task is performed by selecting the functions from search space. After selecting the functions, the AFS initializes the population, distance, max iteration, and crowding factor. The population initialization gives the exact integer space. The bulletin initialization has calculated the present state of fish, and it selects the best behavior of fish. To update the bulletin simulations, the met condition is used to select the best behavior of AF. The process is terminated after obtaining the optimum solution with optimum values. The accuracy and computation time are calculated and compared with the conventional techniques.

Reference [4] presented the breast cancer classification scheme and used the deep learning approach. The multilayer perceptron algorithm is used to calculate the best accuracy. The feature extraction and reduction are performed in the deep neural network model. Ten attributes are used in two different classes, and it is classified with the inbuilt process of the deep neural network. The 8-fold and 10-fold cross-validation technique is performed in the multilayer perceptron algorithm. In this paper, the dataset is collected from the oncology department at Medical Centre University, Yugoslavia. These datasets are not a label but a random dataset. Based on the weight of the neurons, the MLP is performed with the mathematical activation function. Weka 3.8 is utilized for the classification of breast cancer. After performing the cross-validation scheme, the precision of each scheme is calculated.

Reference [5] introduced the classification scheme for cancer images. Here, the back propagations NN and RBNN are used as a classifier. In this paper, the feature extraction process has a significant role in the classification approach, which utilizes the grey-level co-occurrence matrix. The MIAS database is utilized for the classification approach. The initial stage of classification detects either normal or abnormal. If it is abnormal, the next stage predicts either malignant or benign. The GLCM process extracts features such as correlation, entropy, homogeneity, and variance. By comparing both RBFNN and BPNN, the radial basis neural network achieves the classification results. Accuracy and mean squared error are calculated to achieve the result.

Reference [9] described breast cancer analysis. This process is carried out by performing with WEKA. The proposed work performed a sJ48 classifier, which gives the best accuracy result. During preprocessing, the missing traits are neglected. The classifiers performed in this paper are J48, REP Tree, and Naïve Bayes. Here, K-means clustering and farthest first clustering methods are employed to reduce the complexity in the feature extraction procedure. The number of classified instance with correctness and incorrectness is

measured for three classifiers. The accuracy and computation time give enhancement in the three classifiers. The WEKA programmer is used with the machine learning approach for classifying breast cancer.

Reference [8] presented the image segmentation model using the fuzzy entropy technique, which is segmented by the artificial fish school optimization algorithm. Here, the maximum entropy level is selected by the fuzzy rule, and it is used for image segmentation. The histogram images are used to process with the segmentation process. Here, the threshold technique is used to optimize the result through Otsu's threshold model. The maximum entropy is achieved by the threshold algorithm. The AFS algorithm is used to optimize the parameter from the entropy algorithm. In AFS, each fish behavior is evaluated to get the objective function. The image segmentation is performed with the fuzzy entropy model. Here, the double threshold-based image segmentation method improves the entropy level using fuzzy rules. The tire image is taken to perform the image segmentation. The fuzzy entropy level is obtained for both GA and AFS. Here, the optimization performs foraging, clustering, and tailgating.

Reference [21] presented a deep learning technique. This process is developed to improve breast cancer risk prediction. The breast cancer database is taken from a larger tertiary academic medical centre. The mammograms are filtered to analyze the function. The dataset is validated to analyze the data. Based on the patient information (attributes), the cancer set and dataset are separated by examining training, validating, and testing. The hazard ratio and position of the top/bottom decile flow are measured. The hypothesis model has mammography information to capture the cancer-risk state. The Tyrer-Cuzick version is used for hybrid deep learning. Here, the confusion matrix analyzes the density factor of the breast cancer image.

Reference [10] described about the breast cancer dataset for histopathological image classifications. The medical image classification is performed with the automated classification using a computer-aided diagnosis tool. Here, the two classes are used to compute the breast cancer region. ML scheme is modeled to utilize the classification of the medical image. The BreakeHis dataset is taken to test the classification of the breast cancer image model. Malignant and benign image sets are classified with different properties. A completed local binary pattern is utilized in the extraction process, and the texture feature is extracted by the GLCM process. The surf and sift are detected using the oriented fast and rotated brief process. Four different classifiers are used with various feature extraction approaches. Reference [14] utilized the GCNN and CNN for breast cancer classification.

3. Proposed Method

The proposed Deep CNN for classifying the breast cancer image is optimized using the artificial fish school optimization algorithm and is modeled to provide the best classification result. In this section, the proposed methodology is described with classifier and optimization technique and their application of breast cancer image analysis.

3.1. Deep CNN Using AFS Optimization for Classification.

Here, the artificial fish school technique is used to give the direct image set to the training process of the classifier. The random permutation of the AFS is selected by initializing the direct values. The maximum epochs are given in the optimizer at the stage of variable initialization. Hence, the preprocessing stage is performed with the wiener filter for improving the quality of the image. These filtered images are sent to the convolution layer, which sets the image to filter with a convolution filter for activating the features of the given breast image. The convolution 2D layer will normalize the image to perform the training process in the ReLu layer. The pooling layer selects the higher value of the image by the sampling process. The single GPU processor is used for Deep CNN to perform classification. The AFS algorithm provides a direct training image to the Deep CNN. Based on the training image, the supposition accuracy is given in the AFS. The convolution neural network having three layers with sublayers is used. The activation function enables the feature map for the regional size of each image. Then, it sends to the pooling layer for selecting the maximum value of the trained image. The proposed Deep CNN architecture is given in Figure 1.

In the convolution neural network, the above-mentioned layers performed different functions to activate the process. In this proposed method, the 90 input images are extracted for classifying breast cancer, which is given to the input layer of CNN. It is normalized in this convolution layer. The MIAS dataset is extracted for breast cancer images, and it has 1024×1024 pixels. A single GPU processor is capable to analyze the 64×64 image pixels since it is resized to perform the feature extraction. The convolution 2D layer is processed with the grey-scale image of breast cancer image to remove the features.

Here, the ReLu layer is used to provide the positive images to the CNN; otherwise, it outputs to zero. The pooling layer is the major function to activate the CNN. Here, the greater pooling layer is utilized to select the maximum value of pictures from the ReLu layer. The fully connected layer gives the output of feature extraction by classifying the breast cancer image. The extraction procedure is activated in the pooling layer. The probability of picture is applied in the multiclass functions, and it is performed with the softmax layer. In this, the artificial fish school optimization algorithm is implemented for selecting the training image with supposition accuracy. A single Deep CNN process performs the three operations such as feature extraction, feature selection, and classification. The accuracy, sensitivity, specificity, and computation time are measured to analyze the performance.

3.1.1. Dataset. The MIAS database is utilized to separate the images of breast cancer from mammography. The database is downloaded from <http://peipa.essex.ac.uk/info/mias.html>. The mammographic picture analysis society provides the digital mammogram image for classifying the different features. The database comprises of seven columns representing image, type of tissue, abnormality, its size, location, stage, and image

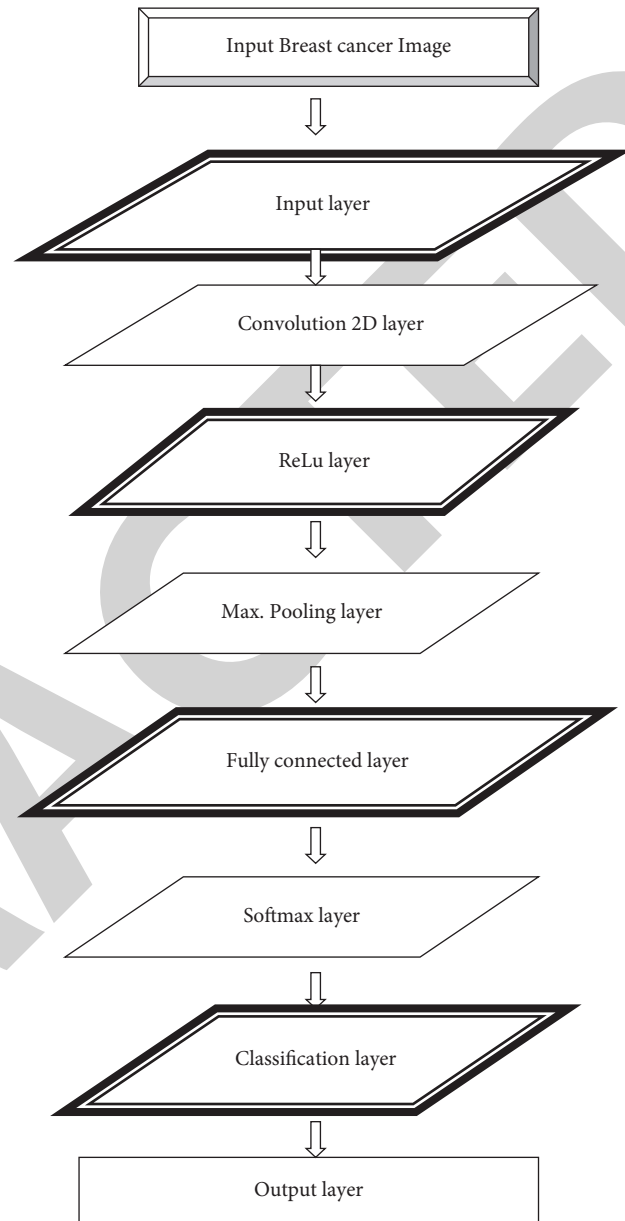


FIGURE 1: Architecture of the DCNN model.

information. This classification approach is also capable to analyze the screening mammography with a digital database. In this proposed system, the MIAS database is used, which has 200-micron pixel with the padding of 1024×1024 . The dataset containing 322 digits with 2.3 GB 8 mm tapes is used. This digital mammogram has the density of 0–3.2, and it is represented with each single pixel that has 8 bits. The MIAS database has different characters of background tissues. The abnormality of the breast cancer image is classed by masses, image distortions, asymmetry, and calcification. It classifies malignant, benign, and normal breast images.

3.1.2. Preprocessing Stage. The first stage of mammography image classification uses the preprocessing step to enhance the image quality. The input has 1024×1024 pixel

size, which is resized with 64×64 pixels. The wiener filter is used in this preprocessing stage. The 2D convolution operation removes the noise presented in the input signal. The overlapping patches, removing edges, improving the visual quality by removing the noise, resizing, and filtering are performed in this process. Here, the normal, benign, and malignant images are filtered using the wiener filter.

3.1.3. Deep CNN. The CNN is the type of a deep learning technique. The DNN model performs the same operation in the convolutional neural network. The proposed Deep CNN is modeled to separate the breast cancer image with the characteristics of normal, benign, and malignant. This process performed with various layers of convolution approach is as follows:

- (i) Convolution 2D layer: this layer is used for normalizing the input pictures to perform the feature extraction process. The 2D convolution process involves the operation as follows:

$$f1(Ps) = \max(0, \text{weight} * P + \text{BF } 1),$$

$$f[u, v] * g[u, v] = \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} f[n, m].g[u - n, v - m]. \quad (1)$$

- (ii) This process improves the image by the normalizing factor. This allows the 2D matrix images for normalizing the values. This performance is done with the grey scale image.
- (iii) Rectified linear unit layer: the ReLu layer activates the neurons presented in the convolutional neural network. It selects the maximum value in the function of $f(0, x)$. It acts as a rectifier, and it performs the differentiations by neuron points in the network. Feature representation with prefiltered images is represented as follows:

$$f2(Ps) = \max(0, \text{weight} * f1(PS) + \text{BF } 2). \quad (2)$$

$$P(v = ki|u, g, \text{weight}_i) = \frac{e^{gk(u, wk)}}{\sum_i^k e^{gk(u, wk)}}. \quad (3)$$

- (iv) The nonlinear rectification is performed to activate CNN functions. The nonlinear function activation is the major concern for various systems. The probability function with different classes is given as follows:
- (v) Max. Pooling layer: the pooling layer is employed to perform the feature extraction process. Here, it selects the maximum value of breast image to activate the pooling layer. The training image sets are selected randomly from the artificial fish-school optimization technique. It also assumes the accuracy of the each image sets.

$$f3(Ps) = \max(0, \text{weight} * f2(Ps) + \text{BF } 3). \quad (4)$$

$$L = - \sum_{i=1}^l \ln(P(v = ki|u; g; \text{weight}))$$

$$= \sum_{i=1}^l \ln \left(\sum_i^k e^{gk(u, wk)} \right) - g_{v^i}. \quad (5)$$

- (vi) The objective function with maximum value representation is given by
- (vii) The spatial size of the image is reduced for the feature reduction approach. It operates only with the feature map of each image. The maximum subsampling of the pooling layer will map the features based on the classes. The 3×3 -convolution matrix strides are used to select the maximum value. The selected training dataset is labeled with the different classes, and it performs testing. Based on the stride and padding in the matrix, the ReLu performs to activate the pooling operation.
- (viii) Fully connected layer: it is the output of feature maps, and it is performed with the AFS optimization. Each node in the Deep CNN is interconnected to perform the breast cancer image classification. It performs the feed-forward neural network model by connecting the feature reduction approaches in the convolution layer. It holds the information about the image features and provides in the vector form. The pooling layer reduces the feature map, and it is optimized with the AFS algorithm, which is connected to perform the classification. The training process selects 22 images, and the testing process is carried out with 8 images. Here, the epoch is 20 based on this; the training process will be performed to provide the classification accuracy.
- (ix) Softmax layer: the softmax layer solves the multiclass problems of integral functions. It selects the maximum value of final class predictions. The softmax activation function normalizes the input image and provides a positive result. The nodes of the fully connected layer are the softmax layer, which activates the function by selecting the multiclass classifications. The pretrained set of images is configured with the dense characteristics of a given dataset.
- (x) Classification layer: the classification layer is the output layer of the Deep CNN, and it provides the classified result of breast cancer images. It has the functionality to know about the image features. Based on stride representation, image pixels are shifted to the matrix. Then, it performs the feature extraction and reduction using Deep CNN layers. The Deep CNN layers form the array in the order of the above-mentioned layers.

$$\text{output} = \text{sigmoid}(L). \quad (6)$$

3.1.4. Artificial Fish School Optimization Algorithm. The optimization of image classification is an important process. In this proposed method, the artificial fish school optimization algorithm is employed to provide the direct training files to the classifier. It also gives the approximate accuracy for each training image. Initially, the population size, crowding factor, maximum iteration, and distance are given in the parameter initialization process. Then, the population is initialized to produce the weight of the fish. The bulletin initialization process evaluates the present state of fishes. If the state of best fish and its position is found, it terminates the process; otherwise, it again performs the operation to select the best AF.

Requires Deep CNN Output Vector of feature map

- (1) For $i = 1$ to *Number of Epoch* Do
- (2) Start
- (3) Initialize *location*
- (4) Initialize the *crowd of the fish*
- (5) Evaluate *fitness*
- (6) While
- (7) *no of epoch* is not obtained
- (8) Do
- (9) Analyze the *unique movement*
- (10) Calculate the *fitness*
- (11) Enhance the *new location*
- (12) Analyze the *fitness difference*
- (13) Analyze the *fish weight*
- (14) Analyze the *variance of the weights*
- (15) Analyze the *instinctive movement*
- (16) Enhance the *new location*
- (17) Analyze the *barycentre*
- (18) Analyze the *volitive movement*
- (19) Evaluate the *fitness*
- (20) Enhance the *individual step*
- (21) Update the *volitive step*
- (22) End For

In the presented model, the AFS algorithm is implemented to select the training image, and it provides directly to the image classifier. It also provides the respective accuracy to the images. This optimal solution is used to train the image, and it is tested. These images are classified based on the features. The approximate accuracy is used for the classification result. The training accuracy is approximately given to the Deep CNN. However, it is not an exact accuracy result of the classifier. The objective function of AFS is given by

$$X_a = X_b + \text{training data}(n). \quad (7)$$

The artificial fish performs swarming that is given in the following expression:

$$X_b^{m+1} = X_b^m + \frac{X - X_b^m}{\|X - X_b^m\|} \cdot \text{train data}(n). \quad (8)$$

If the artificial fish follows the neighbor partners for food search, the behavioral description represents the fitness value. The forward companion of AF while getting the fitness value is given by

$$X_a^{m+1} = X_a^m + \frac{X_b - X_a^m}{\|X_b - X_a^m\|} \cdot \text{step} \cdot \text{train}(n). \quad (9)$$

The movement of the AF selects the path randomly, and its prey behavior is given by

$$X_a^{(m+1)} = X_a^m + \text{visual} \cdot \text{random}(n). \quad (10)$$

The best possible image is selected by the following equations:

if (Best AF(k) – Best AF(l)) < iteration,

$$X_{\text{some}}^{m+1} = X_{\text{some}}^m + \beta \cdot \text{visual} \cdot \text{random}(n),$$

$$X_a^{m+1} = X_a^m + \left(\frac{X_b - X_a^m}{\|X_b - X_a^m\|} + \frac{X_{\text{bestAF}} - X_a^m}{\|X_{\text{bestAF}} - X_a^m\|} \right) * \text{AF_step} * \text{random}(n),$$

$$X_a^{m+1} = X_a^m + \text{AF_visual} * \text{random}(n). \quad (11)$$

The global search of the AFS algorithm is to optimize the training image to perform classification. The convergence speed shows the search optimization result of AF. Artificial fish school optimizer finds better classification accuracy. In this proposed image classification method, the convolution layers performs the extraction and reduction procedure, which is optimized by selecting the direct training files to the classifier. This optimization algorithm performs the behavior of the artificial fish to select the food in the sea. This scheme is applied to the CNN for direct selection of training images. The sensitivity, accuracy, and specificity of the proposed scheme are analyzed to compare with the existing work. By this performance analysis, the accuracy, sensitivity, and specificity are achieved for the presented model.

4. Results and Discussion

Thus, the classification of breast cancer images using DCNN and AFS optimization algorithms is modeled, and it obtains the best outcome as compared to the previous work. The MIAS dataset helps to classify the images of breast cancer with three stages, that is, normal, benign, and malignant. These images have 1024×1024 pixels in size. The input data of normal breast cancer images are illustrated in Figure 2. The total input breast images in the information base are 90. For each property, it selects 30 sets of images to classify the normal, benign, and malignant tissues. For the first set, the Deep CNN takes 22 images to train the network. The testing process utilizes 8 images. Here, the input image of benign and malignant is shown in Figures 3 and 4, respectively. Initially, the preprocessing stage filters the input image using the wiener filter. A single GPU processor does not accept large size images since it is resized to 64×64 pixels.

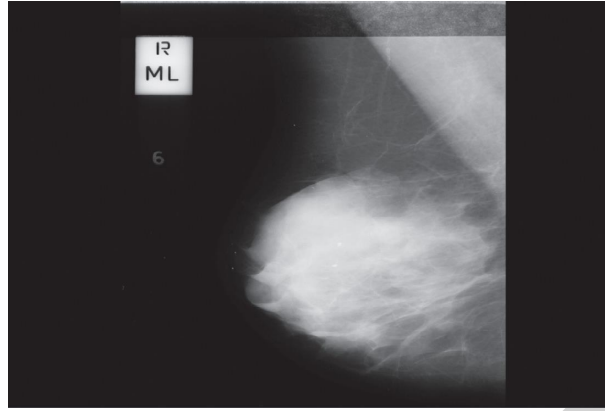


FIGURE 2: Input image (normal).

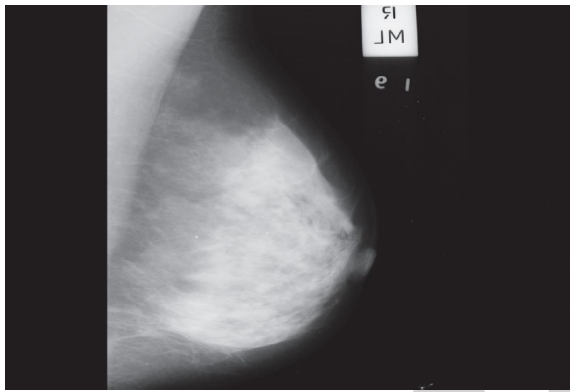


FIGURE 3: Input image (benign).

The preprocessing stage of classification uses the wiener filter for resizing and filtering the breast image input. Here, 64×64 images are applied to DCNN for classification. The properties of the breast image are filtered with the help of the wiener filter, which is shown in Figure 5.

Here, the AFS optimization algorithm is used to select the training images while performing the feature extraction process in Deep CNN. The optimal result is found in the fifth position of the training set, and its average accuracy is provided. Figure 6 shows the simulated result of the artificial fish school algorithm, which provides the best possible

training set with respective average accuracy. The CNN layers perform various operations with respective functionalities. A single processor performs three sets of functionalities as extraction, reduction, and classifications. Hence, the result of CNN is shown in Figure 7. Here, various performances in the CNN result in the classifier efficiency.

The optimal training image position and its corresponding accuracy from AFS optimizer is shown in Figure 6.

The training process in Deep CNN uses the direct selection of the training file, which uses the artificial fish-school optimization process. Here, 20 iterations are performed to obtain the best result, and they also give the reduced losses of both training and testing. The training progress is shown in Figure 8.

In this, the proposed Deep CNN using AFS optimizer provides the best result, and its performance is evaluated (see Figure 9) based on the different classes assigned in the network. The confusion matrix arranges the TP, TN, FP, and FN ranges to analyze the accuracy, sensitivity, and specificity.

Performance evaluation parameters are calculated by obtaining the result of accuracy, recall, and F1 score. In this, accuracy, sensitivity, and specificity are calculated to achieve the result of overall performance of the Deep CNN using the AFS algorithm.

$$\% \text{ of accuracy} = \frac{\text{True Neg} + \text{True Pos}}{\text{True Pos} + \text{False Pos} + \text{False Neg} + \text{True Neg}} \times 100,$$

$$\text{sensitivity} = \frac{\text{True Pos}}{\text{True Pos} + \text{True Neg}}, \quad (12)$$

$$\text{specificity} = \frac{\text{True Neg}}{\text{True Neg} + \text{False Pos}}.$$

By this measurement, the performance is evaluated for both training and testing processes. The comparison result of proposed and existing methods is given in Table 1

Time complexity is the time required for running the process. In the proposed method, the time complexity is reduced with the minimal training images. In the

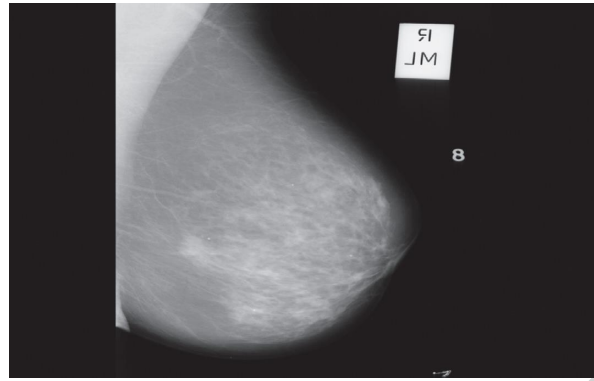


FIGURE 4: Input image (malignant).

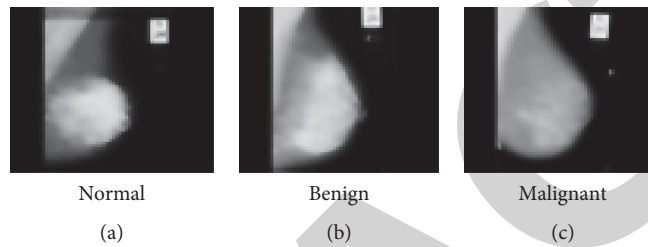


FIGURE 5: Filtered images. (a) Normal, (b) benign, and (c) malignant.

```

Command Window
Warning: Image is too big to fit on screen; displaying at 33%
> In images.internal.initState (line 71)
  In imshow (line 336)
  In main (line 8)
>> main_fish
The optimal position for the best training*5 The average accuracy 84.8
>> main_fish
The optimal position for the best training*5 The average accuracy 86.2
fx >>
  
```

FIGURE 6: Result of AFS optimizer.

presented work, the image classification is modeled based on the deep learning approach of the convolution neural network. This technique provides the classified result by optimizing the training process in the classifier, which uses the behavioral approach of the artificial fish-school optimization technique. The accuracy, sensitivity, specificity, and computation time of both training and testing

processes are evaluated and compared with the conventional approach. The proposed Deep CNN using the artificial fish school model achieves the best result for breast image classifications. Therefore, the analysis report shows that the proposed Deep CNN using the artificial fish school is a better image classifier than the conventional methods.

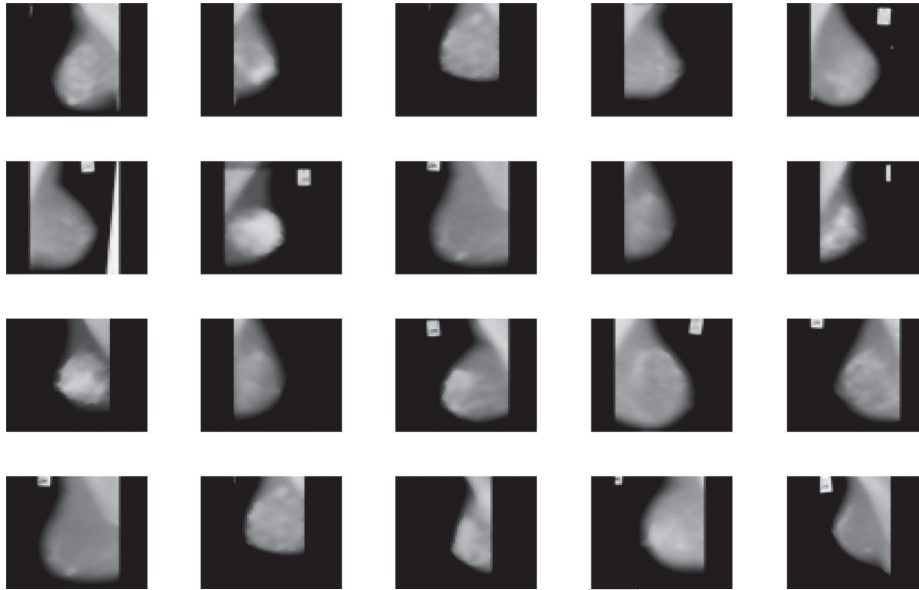


FIGURE 7: Simulated result of deep convolutional neural network.

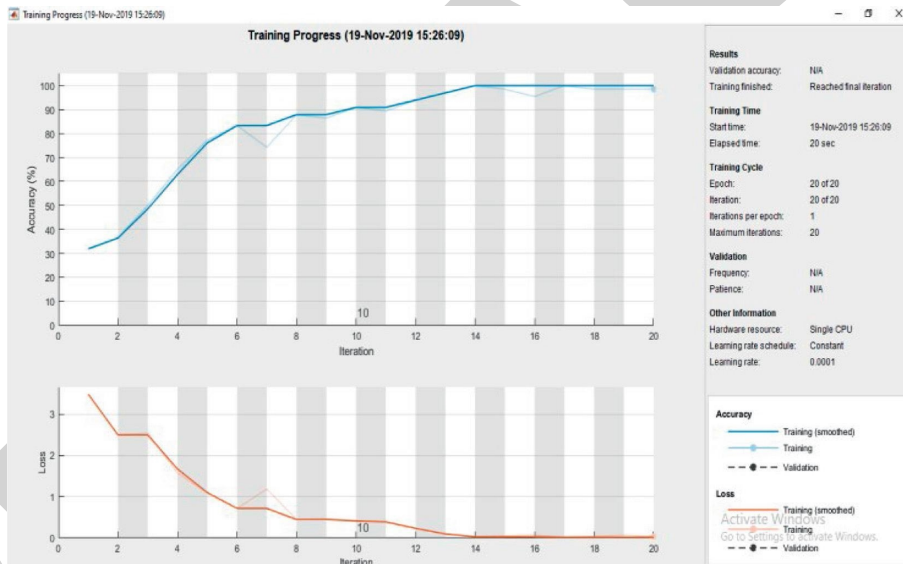


FIGURE 8: Training progress in Deep CNN using AFS optimizer.

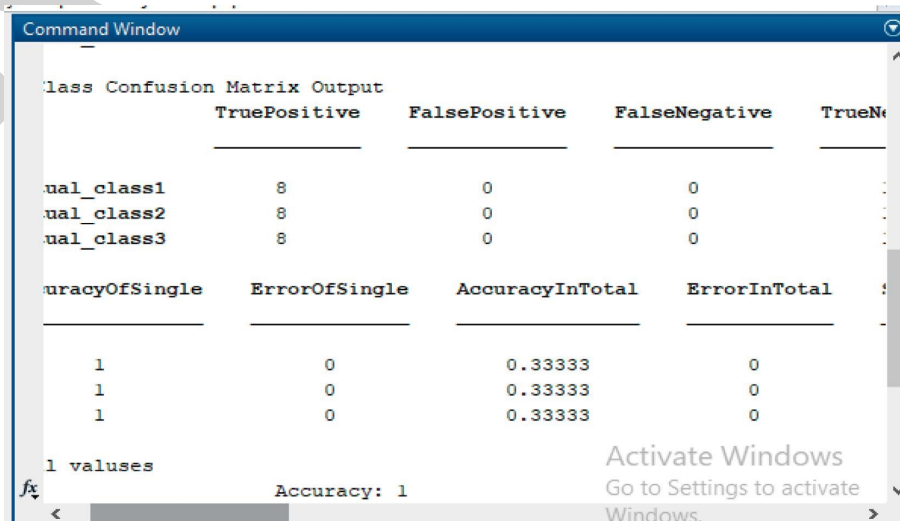


FIGURE 9: Performance evaluation of the proposed method.

TABLE 1: Comparison result of performance evaluation.

Parameters	Existing work [22]		RBFNN using CSO		Deep CNN using AFS	
	Training	Testing	Training	Testing	Training	Testing
% of accuracy	97.1	97.0	100	98.3	100	98.66
Sensitivity	0.96	0.96	100	0.98	100	0.991
Specificity	0.97	0.96	100	0.98	100	0.988
Computation time	55sec	30sec	52sec	30sec	20sec	16sec

5. Conclusion

The artificial fish-school optimization technique is used for selecting the number of epochs in the design of the deep convolutional neural network model, and it is utilized for classifying the mammography images from the MIAS database which is presented. It can classify the malignant, normal, and benign characteristics of breast cancer images. Here, the single CNN mode performs the three operations like preprocessing and feature extraction, and also it enhances the design model. The Deep CNN is optimized using AFS in the training process. Here, the artificial fish-school optimization technique selects the training images by directly assigning them to the classifier. Here, this optimizer also provides the average accuracy level of the training images. This enhancement makes the system efficient by a reduced computation time. Here, the proposed CNN model performs feature extraction, feature reduction, and classification. By comparing the performance analysis of the proposed method versus conventional methods, the proposed Deep CNN using AFS optimizes the result by providing better accuracy, sensitivity, and specificity. Thus, the system concludes that the breast image classification is performed efficiently using Deep CNN based on the AFS algorithm. In the future, the optimization technique is improved for utilizing the operations in both training and testing processes. The CNN model is modified to provide better efficiency than this system. These numbers of properties are extracted to characterize the various properties in breast cancer images. The proposed work is concentrated only on the training image and epoch. It can be improved by including other hyperparameters. [23–31].

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 no conflicts of interest.

References

- [1] S. Reis, P. Gazinska, J. H. Hipwell et al., "Automated classification of breast cancer stroma maturity from histological images," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2344–2352, 2017.
- [2] M. A. Anupama, V. Sowmya, and K. P. Soman, "Breast cancer classification using capsule network with preprocessed histology," in *Proceedings of the international conference on communication and signal processing*, pp. 0143–0147, IEEE, Chennai, India, 4 April 2019.
- [3] M. Amrane, S. O. Gagaoua, and T. Ensari, "Breast cancer classification using machine learning," in *Proceedings of the Electric Electronics, Computer science, Biomedical engineerings' Meeting*, 18–19 April 2018.
- [4] S. N. Jasmir, R. Firsandaya, Dodo Z. Abidin, Z. Ahmed, Y. N. Kunang, and Firdaus, "Breast cancer classification using deep learning," in *Proceedings of the international conference on electrical engineering and computer science*, pp. 237–241, IEEE, Indonesia, 2 October 2018.
- [5] M. Pratiwi, J. Alexander, J. Harefa, and S. Nanda, "Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network," *Procedia Computer Science*, vol. 59, pp. 83–91, 2015.
- [6] K. kumar and A. C. Sekhara Rao, "Breast cancer classification of image using convolutional neural network," in *Proceedings of the international conference on recent advances in information technology*, 15 March 2018.
- [7] D. Ribli, A. Horváth, Z. Unger, P. Pollner, and I. Csabai, "Detecting and classifying lesions in mammograms with Deep Learning," *Scientific Reports*, vol. 8, p. 4165, 2018.
- [8] L. Qin, K. Sun, and S. Li, "Maximum fuzzy entropy image segmentation based on artificial fish school algorithm," in *Proceedings of the International conference on Intelligent Human-Machine systems and cybernetics*, 27 August 2016.
- [9] T. Padhi and P. Kumar, "Breast cancer analysis using WEKA," in *Proceedings of the international conference on cloud computing, Data science and engineering*, 10 January 2019.
- [10] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, 2016.
- [11] X. Zhang, W. Liu, M. Dunder, S. Badve, and S. Zhang, "Towards large-scale histopathological image analysis: hashing-based image retrieval," *IEEE Transactions on Medical Imaging*, vol. 34, no. 2, pp. 496–506, 2015.
- [12] M. H. Waseem, M. S. A. Nadeem, A. Abbas et al., "On the feature selection methods and reject option classifiers for robust cancer prediction," *IEEE Access*, vol. 7, pp. 141072–141082, 2019.
- [13] Md Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network," in *Proceedings of the Computer vision and pattern recognition*, IEEE, Salt Lake City, Utah, 18 June 2018, <https://arxiv.org/abs/1811.04241>.
- [14] Y. D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S. H. Wang, "Improved breast cancer classification through combining graph convolutional network and convolutional neural network," *Information Processing & Management*, vol. 58, no. 2, Article ID 102439, 2021.
- [15] K. Suksut, R. Chanklan, N. Kaoungku, K. Chaiyakhan, K. Nittaya, and K. Kittisak, "Parameter optimization for

Retraction

Retracted: Evolving Long Short-Term Memory Network-Based Text Classification

Computational Intelligence and Neuroscience

Received 13 September 2023; Accepted 13 September 2023; Published 14 September 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] A. Singh, S. K. Dargar, A. Gupta et al., "Evolving Long Short-Term Memory Network-Based Text Classification," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4725639, 11 pages, 2022.

Research Article

Evolving Long Short-Term Memory Network-Based Text Classification

Arjun Singh ¹, Shashi Kant Dargar ², Amit Gupta ³, Ashish Kumar ⁴,
Atul Kumar Srivastava ⁵, Mitali Srivastava ⁵, Pradeep Kumar Tiwari ⁶,
and Mohammad Aman Ullah ⁷

¹Computer and Communication Engineering, School of Computing and IT, Manipal University Jaipur, Jaipur, India

²Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Virudhunagar, Tamilnadu, India

³Department of Electronics and Communication Engineering, Narasaraopeta Engineering College, Narasaraopeta, Andhra Pradesh, India

⁴Department of Computer Science and Engineering, School of Computing and IT, Manipal University Jaipur, Jaipur, India

⁵School of Computing, DIT University, Dehradun, India

⁶Manipal University Jaipur, Jaipur, India

⁷Department of Computer Science and Engineering, International Islamic University Chittagong, Chittagong, Bangladesh

Correspondence should be addressed to Mohammad Aman Ullah; aman_cse@iiuc.ac.bd

Received 5 November 2021; Revised 10 December 2021; Accepted 12 January 2022; Published 21 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Arjun Singh 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.

Recently, long short-term memory (LSTM) networks are extensively utilized for text classification. Compared to feed-forward neural networks, it has feedback connections, and thus, it has the ability to learn long-term dependencies. However, the LSTM networks suffer from the parameter tuning problem. Generally, initial and control parameters of LSTM are selected on a trial and error basis. Therefore, in this paper, an evolving LSTM (ELSTM) network is proposed. A multiobjective genetic algorithm (MOGA) is used to optimize the architecture and weights of LSTM. The proposed model is tested on a well-known factory reports dataset. Extensive analyses are performed to evaluate the performance of the proposed ELSTM network. From the comparative analysis, it is found that the LSTM network outperforms the competitive models.

1. Introduction

With exponential growth in text documents available on Internet, the manual labeling of textual contents in digital form into various classes is extremely challenging to realize. Therefore, many automatic text classification models have been developed such as hierarchical multi-label text classification (HMLTC) [1] and coattention model with label embedding (CMLE) [2]. These models are trained on historical datasets and processed according to a group of labeled data. These models require efficient text encoding models which decompose the text to sequence vectors [1]. The existing text classification models extract a highly discriminative text representation. But

these models are generally computationally extensive in nature [2].

Recently, multilabel text classification models were designed. These models are complex compared to single-label classification models [3]. Many researchers have utilized deep learning models for text classification such as recurrent neural network (RNN) and long short-term memory (LSTM). But these models are unable to handle data imbalanced problems [4].

Recently, many researchers have designed label space dimension reduction to classify text with multiple classes. However, the majority of the models have ignored the sequential details of texts and label correlation with the original label space. Thus, labels were assumed to be

meaningless vectors [5]. Also, for the classification of long text, there were a lot of redundant details in textual data. This redundant detail may contain some sort of knowledge too. Thus, the classification of long text requires an efficient model [6].

Mostly, the text details are available in unstructured form. Therefore, the extraction of required details from a huge number of documents becomes a challenging problem [7]. In [8], a bidirectional gated temporal convolutional attention network (BGTCAN) was designed. During the extraction of features, this model has utilized a BGTCAN to obtain the bidirectional temporal features. The attention process was also used to distinguish the significance of various while preserving the maximum text features. In [9], an efficient text classification model was proposed. It has integrated the context-relevant features with a multistage attention model by considering TCN and CNN.

In [10], an efficient hybrid feature selection model was designed. Binary poor and rich optimization (HBPRO) was utilized to compute the significant subset of required features. A Naive Bayes classifier was then used for classification. HBPRO is based on people's wealth such as rich and poor in the world. The rich group tries to widen their group gap by computing from those in the poor group. Every solution in the poor group moves towards the global optimal solution in the search space by learning from the rich group. In [11], an in-memory processor for Bayesian text classification was designed by considering a memristive crossbar model. Memristive switches were utilized to hold the required details for the text classification. In [12], a hybrid model was proposed. It has integrated the gated attention-based BLSTM and the regular expression-based classifier. BLSTM and an attention layer were utilized to weigh tokens according to their perceived significance and focus on critical fractions of a string.

In [13], a backdoor keyword identification model was proposed to overcome the backdoor attacks with LSTM-based models. In [14], label-based attention for hierarchical multilabel text classification neural network was proposed. An efficient label-based attention module was proposed to obtain significant details from the text using labels from various hierarchy levels. In [15], support vector machines (SVM) were utilized to recognize text and documents.

From the existing literature, it is found that the LSTM network suffers from the parameter tuning problem. Generally, initial and control parameters of LSTM are selected on a trial and error basis. It means the parameters of LSTM models are selected by manually selecting some possible values. Whichever combination shows better performance is followed as control parameters of LSTM. Parameter tuning deals with the optimization of the control parameters of the LSTM model. It can improve the performance of LSTM, but it comes up with additional computations during the model building time. Therefore, in this paper, an evolving LSTM (ELSTM) network is proposed. The key contributions of this paper are as follows:

- (1) An evolving long short-term memory (LSTM) (ELSTM) network is proposed for text classification.

- (2) Multiobjective genetic algorithm (MOGA) is used to optimize the architecture and weights of LSTM.
- (3) The proposed model is tested on a well-known factory reports dataset. Extensive analyses are performed to evaluate the performance of the proposed ELSTM network.

The remaining paper is organized as follows. Section 2 discusses the related work. Section 3 presents the proposed ELSTM network for text classification. Section 4 presents the performance analysis of the proposed ELSTM network on a well-known factory reports dataset. Section 5 concludes the paper.

2. Related Work

In [16], a bidirectional LSTM (BiLSTM) was proposed for text classification. The word embedding vectors and BiLSTM were utilized to obtain both the succeeding and preceding context information. Softmax was also utilized to obtain classification results. In [17], an attention LSTM (ALSTM) network was proposed for text data classification. The ALSTM has shown significant performance in terms of generalization. In [18], deep contextualized attentional bidirectional LSTM (DCABLSTM) was proposed. By utilizing the contextual attention mechanism, DCABLSTM has the ability of learning to attend to the valuable knowledge in a string. In [19], two hidden layers-based LSTM model (THLSTM) was proposed. The first layer was utilized to learn the strings to demonstrate the semantics of tokens with LSTM. The second layer has encoded the relations of tokens. In [20], a recurrent attention LSTM (RALSTM) was proposed to iteratively evaluate an attention region considering the key sentiment words. Attention and number of tokens were minimized in an efficient manner. The TSLSTM leveraged the coefficients of tokens for classification. A joint loss operator was also used to highlight significant attention regions and keywords. In [21], CNN and LSTM were combined for better performance. It has been found that the integrated model can outperform many competitive models. In [22], LSTM fully convolutional network (LSTMFCN) and attention LSTM-FCN (ALSTMFCN) were designed. The fully convolutional block with a squeeze-and-excitation block was used to improve the performance. These models require significantly lesser preprocessing. In [23], convolutional LSTM (CLSTM) network was designed. CLSTM has been found to be adaptable in evaluating big data, keeping scalability. Additionally, CLSTM was free from any specific domain. However, [16–23] are sensitive to its initial parameters.

To overcome parameter sensitivity issues with LSTM variants, in [24], particle swarm optimization (PSO) was utilized to optimize the LSTM model. PSO was utilized to tune the initial and control parameters of the LSTM network. It has been found that the PSO-based LSTM achieves remarkable results. In [25], a genetic algorithm was utilized to optimize the LSTM. This model can automatically learn the features from sequential data. In [26], a genetic algorithm was utilized to compute the epoch size, number of

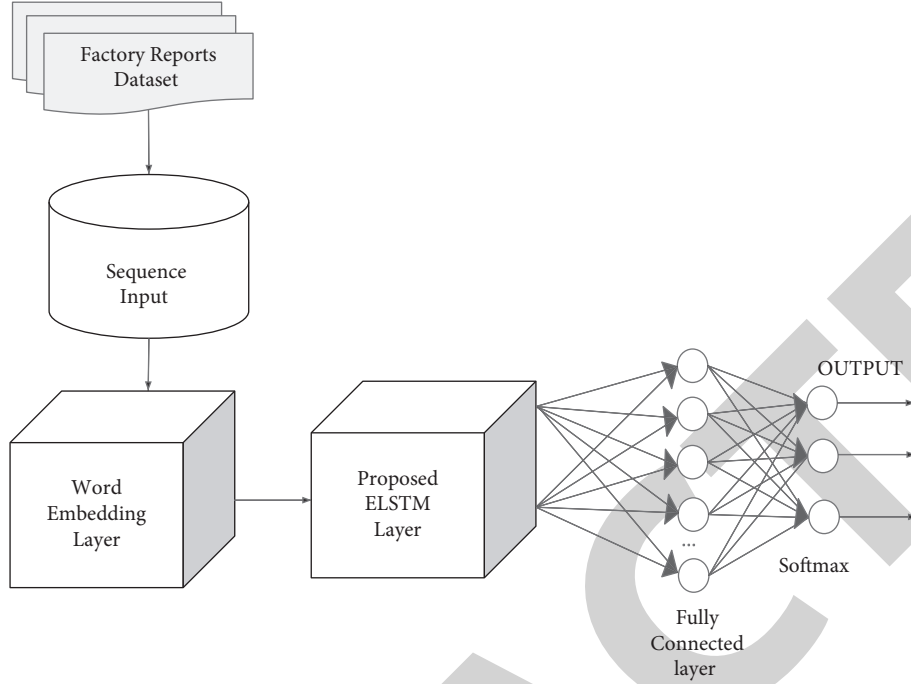


FIGURE 1: Histogram distribution of the target classes.

layers, units size in every layer, and time window size. However, [24–27] suffer from the stuck in local optima and poor convergence speed issues.

It is found that the LSTM network suffers from the parameter tuning problem. The initial and control parameters of LSTM are generally selected on a trial and error basis. Therefore, in this paper, an ELSTM network is proposed.

3. Proposed Methodology

This section discusses the proposed ELSTM model. Initially, LSTM is discussed. Thereafter, MOGA is presented. Finally, MOGA-based LSTM, i.e., ELSTM is discussed. Figure 1 shows the diagrammatic flow of the proposed model. Initially, the dataset is loaded, and preprocessing operation is applied to it.

Since the data is textual in nature, therefore, word encoding is used to convert the strings to numeric sequences. Finally, the proposed ELSTM is trained on the dataset by using a word embedding layer.

3.1. LSTM Network. LSTM is a special kind of variant of recurrent neural network (RNN). It was proposed to overcome the long dependency period problem with RNN. Thus, it can preserve information for a longer period.

Consider a sequence input $S = s_1, s_2, s_3, \dots, s_p, s_1, s_2, s_3, \dots, s_p$ showing each token in the textual data. Mathematically, LSTM can be computed as follows:

$$f_k = \sigma(W_f \cdot i_{k-1} + W_f \cdot s_k + b_f), \quad (1)$$

where σ shows a sigmoid function. W and s represent the weight matrices and bias vector attributes. i_k is the hidden state and can be computed as

$$i_k = \sigma(W_i \cdot i_{k-1} + W_i \cdot s_k + b_i). \quad (2)$$

The current layer's memory (\tilde{M}) can be computed as

$$\tilde{M} = \tanh(W_c \cdot i_{k-1} + W_c \cdot s_k + b_c). \quad (3)$$

For k^{th} token, the memory cell (C_k) block can be computed as

$$C_k = f_k \cdot C_{k-1} + i_k \cdot \tilde{M}. \quad (4)$$

The activation vector (o_k) of the output gate can be computed as

$$o_k = \sigma(W_o \cdot i_{k-1} + W_o \cdot s_k + b_o). \quad (5)$$

The output vector so-called hidden state vector (i_k) can be computed as

$$i_k = o_k \cdot \tanh(C_k). \quad (6)$$

3.2. Fitness Function. The main objective of this paper is to optimize the architecture in such a way that it achieves better performance with less number of hidden layers for the LSTM network [28–30]. Therefore, a multiobjective fitness function is designed by using validation accuracy (A_c) and the number of hidden nodes of LSTM. The fitness function ($f(t)$) can be defined as

$$f(t) = \begin{cases} \text{Maximize}(V_A), \\ \text{Minimize}(H_N). \end{cases} \quad (7)$$

```

(i) Output:  $P_f$ /*  $P_F$  shows a Pareto front.
(ii) Input: factory reports dataset ( $FR D$ ), LSTM, and initial population ( $I_P$ ).
(iii) begin
(iv) Use  $I_P$  as initial parameters of LSTM
(v) Implement LSTM on training fraction of  $FR D$ ;
(vi) Validate  $FR D$  on validation fraction of  $FR D$ ;

(vii) Compute  $V_A$  and  $H_N$ ;
(viii) Evaluate  $P_f = \{\max.V_A, \min.H_N\}$ ;
(ix)  $N_s \leftarrow$  Apply nondominated sorting on  $P_f$ ;
(x) return  $P_f$ 
(xi) end

```

ALGORITHM 1: Optimization of LSTM.

```

(i) Output: optimized population
(ii) Input: {initial parameters of MOGA}
(iii) begin
(iv) Obtain initial random solutions  $I_P$  ;
(v)  $f_t \leftarrow$  Call Algorithm 1 by considering  $I_P$ ;
(vi)  $S_P \leftarrow$  Sort  $I_P$  according to  $f_t$ ;
(vii) /* Selection operator */
(viii)  $F_P \leftarrow S_P$  ;
        While  $c_e \neq 0 \parallel l_G == 1$  do
(ix) /*  $l_G$  and  $c_e$  indicate final generation and children elimination */
(x) Randomly select  $c$ /*  $c$  mutation point */
(xi)  $c_e \leftarrow 0$  ;
(xii) for  $c$  do
(xiii) Evaluate fitness of  $c$ 
(xiv) if  $f_c \leq f_{F_P[1]}$ , then
(xv) remove  $c$ ;
(xvi)  $c_e \leftarrow c_e + 1$  ;
(xvii) else
(xviii)  $F_P \leftarrow c$  ;
(xix) end
(xx) end
/* Mutation */
for crossover do
Consider two solutions randomly as  $c_1$  and  $c_2$ /*  $c_1, c_2$ , and  $c_3$  are children */
 $c_3 \leftarrow c_1 \oplus c_2$  ;
Computer fitness of  $c_3$ 
if  $f_{c_3} \leq f_{c_1} \parallel f_{c_2}$  then
remove  $c_3$ ;
else
remove  $c_1, c_2$ ;
end
end
end
/* Ranking */  $F_P \leftarrow$  Apply nondominated sorting on  $F_P$ ;
return  $F_P[1]$ 
end

```

ALGORITHM 2: Evolving long short-term memory networks.

Description			
"Items are occasionally getting stuck in the scanner spools."			
"Loud rattling and banging sounds are coming from assembler pistons."			
"There are cuts to the power when starting the plant."			
"Fried capacitors in the assembler."			
"Mixer tripped the fuses."			
"Burst pipe in the constructing agent is spraying coolant."			
"A fuse is blown in the mixer."			
"Things continue to tumble off of the belt."			
Category	Urgency	Resolution	Cost
"Mechanical Failure"	"Medium"	"Readjust Machine"	45
"Mechanical Failure"	"Medium"	"Readjust Machine"	35
"Electronic Failure"	"High"	"Full Replacement"	16200
"Electronic Failure"	"High"	"Replace Components"	352
"Electronic Failure"	"Low"	"Add to watch List"	55
"Leak"	"High"	"Replace Components"	371
"Electronic Failure"	"Low"	"Replace Components"	441
"Mechanical Failure"	"Low"	"Readjust Machine"	38

FIGURE 2: Screenshot of the factory reports dataset.

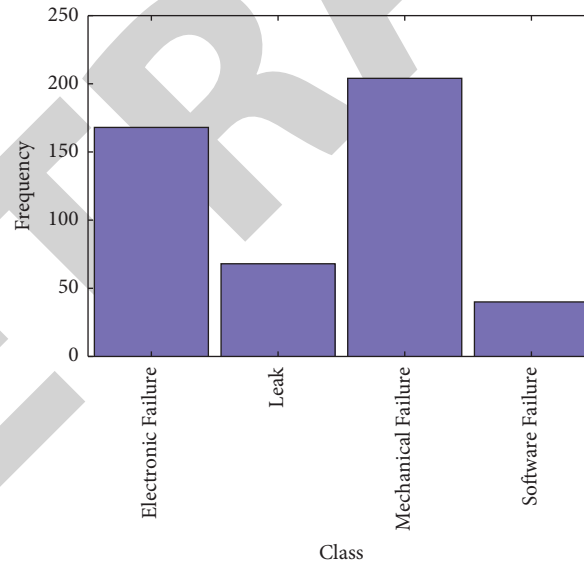


FIGURE 3: Histogram distribution of the target classes.

Here, V_A shows the validation accuracy. H_N shows the number of hidden nodes used by the LSTM network.

3.3. Multiobjective Genetic Algorithm. This section discusses the MOGA-based LSTM (ELSTM) network. Since (7) is a Pareto optimal problem, Algorithm 1 shows step-by-step procedure of the optimization of LSTM.

The genetic algorithm contains a group of operators to optimize the given fitness function [31, 32]. Initially, the normal distribution is used to obtain the random

population. These random solutions act as initial parameters of the LSTM network [33–35]. Fitness function (see Eq. mop) is then used to evaluate the fitness of the computed solutions. Nondominated sorting is then used to rank the solutions. Mutation and crossover operators are then utilized to compute the child solutions [36–39]. Mutation and crossover operators are used to obtain child solutions from the parent solutions for evolving process of genetic algorithms. The nondominated solution with a better trade-off between validation accuracy and the number of hidden nodes is used as a final solution for LSTM. Algorithm 2

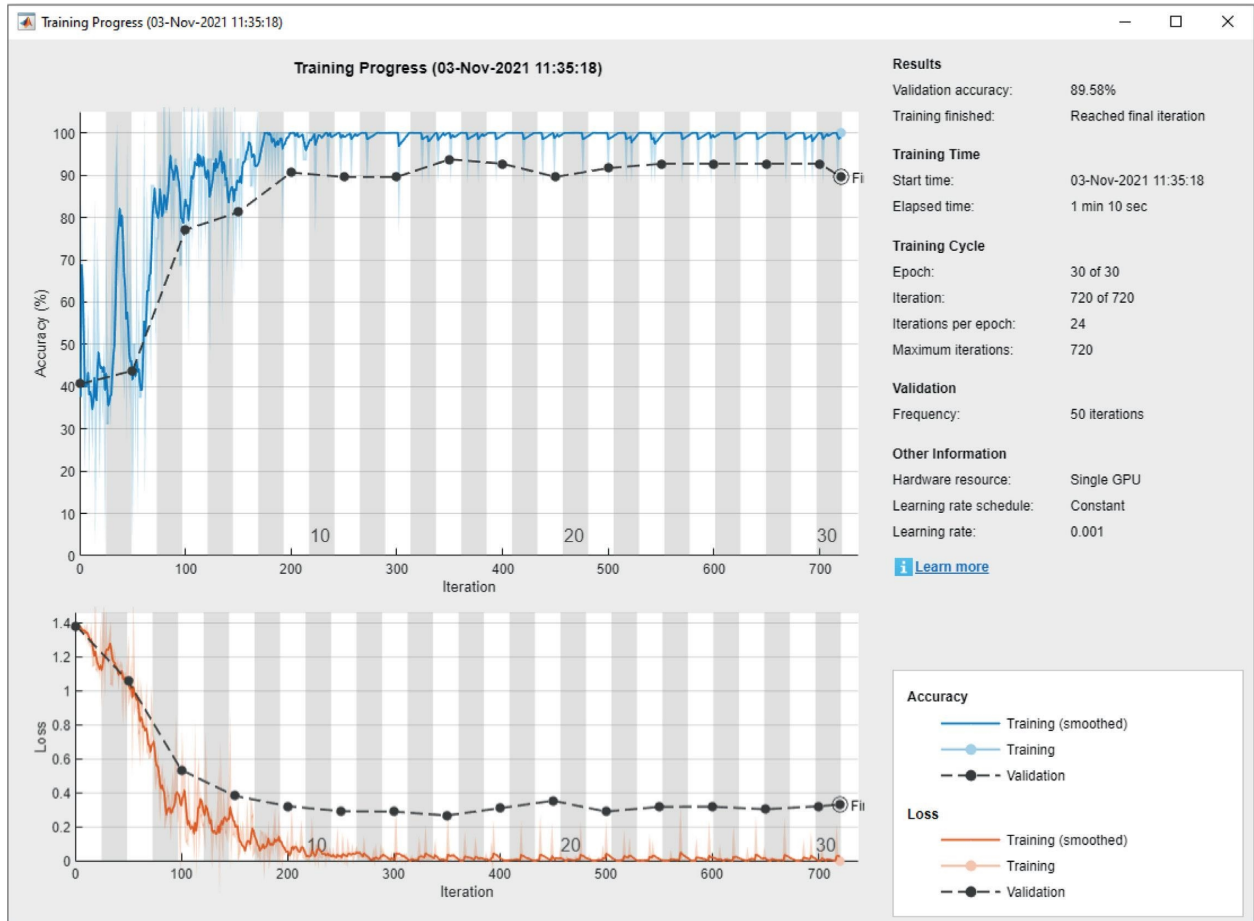


FIGURE 7: Training and validation analysis of the Adam optimizer-based LSTM network.

Epoch	Iteration	Time Elapsed (hh : mm : ss)	Mini- batch Accuracy	Validation Accuracy	Mini- batch Loss	Validation Loss	Base Learning Rate
1	1	00 : 00 : 09	37 . 50%	40 . 62%	1 . 3858	1 . 3814	0 . 0010
3	50	00 : 00 : 14	25 . 00%	43 . 75%	1 . 1859	1 . 0563	0 . 0010
5	100	00 : 00 : 19	87 . 50%	77 . 08%	0 . 3392	0 . 5337	0 . 0010
7	150	00 : 00 : 23	93 . 75%	81 . 25%	0 . 1465	0 . 3822	0 . 0010
9	200	00 : 00 : 27	100 . 00%	90 . 62%	0 . 1250	0 . 3210	0 . 0010
11	250	00 : 01 : 31	100 . 00%	89 . 58%	0 . 0126	0 . 2914	0 . 0010
13	300	00 : 00 : 36	100 . 00%	89 . 58%	0 . 0149	0 . 2888	0 . 0010
15	350	00 : 00 : 40	100 . 00%	93 . 75%	0 . 0028	0 . 2662	0 . 0010
17	400	00 : 00 : 44	100 . 00%	92 . 71%	0 . 0041	0 . 3090	0 . 0010
19	450	00 : 00 : 48	100 . 00%	89 . 58%	0 . 0020	0 . 3525	0 . 0010
21	500	00 : 00 : 52	93 . 75%	91 . 67%	0 . 1283	0 . 2893	0 . 0010
23	550	00 : 00 : 56	100 . 00%	92 . 71%	0 . 0007	0 . 3169	0 . 0010
25	600	00 : 01 : 00	100 . 00%	92 . 71%	0 . 0032	0 . 3179	0 . 0010
28	650	00 : 01 : 04	100 . 00%	92 . 71%	0 . 0017	0 . 3030	0 . 0010
30	700	00 : 01 : 08	100 . 00%	92 . 71%	0 . 0007	0 . 3205	0 . 0010
30	720	00 : 01 : 10	100 . 00%	89 . 58%	0 . 0006	0 . 3307	0 . 0010

FIGURE 8: Epoch and iteration-wise mini-batch training and validation accuracy analysis of the Adam optimizer-based LSTM network along with respective losses and base learning rate.

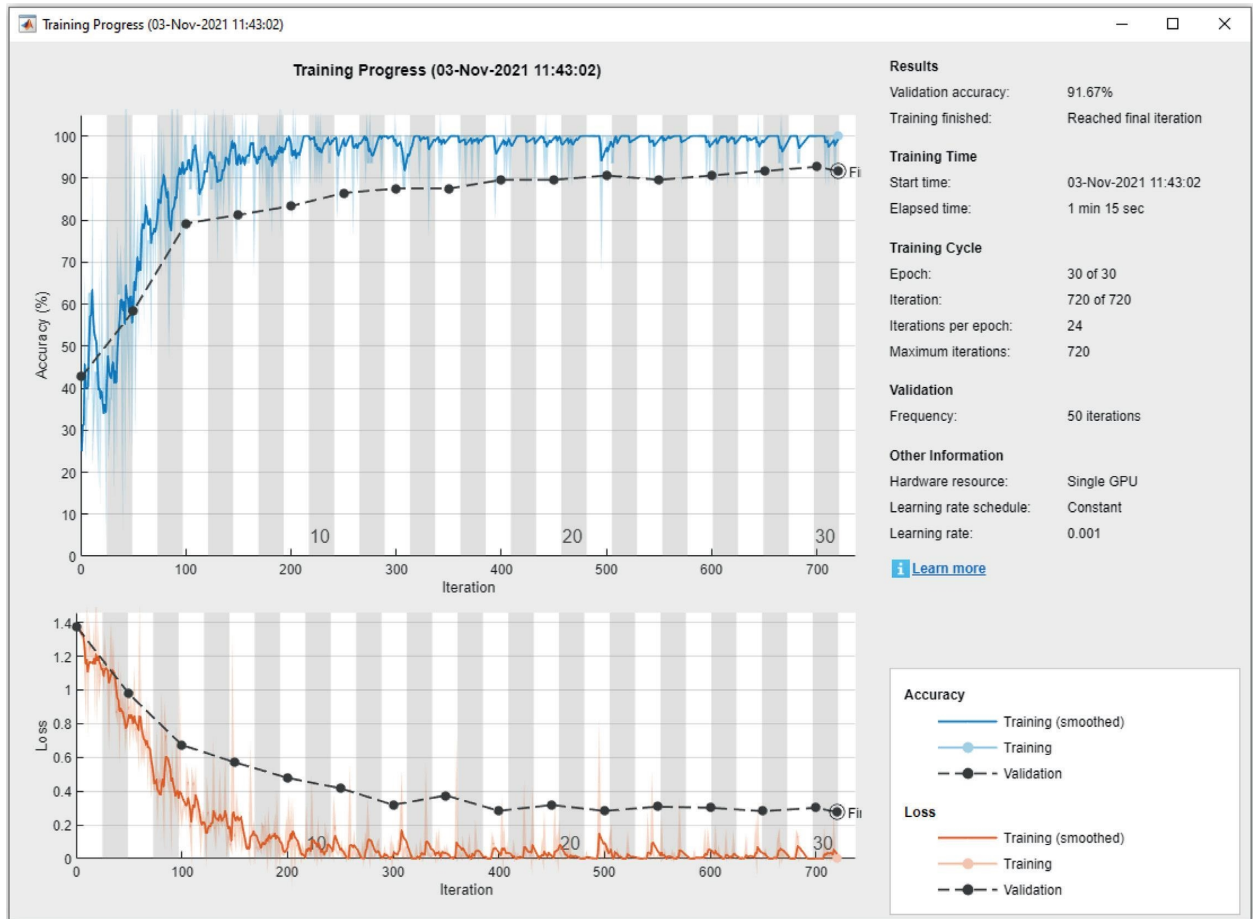


FIGURE 9: Training and validation analysis of the RMSprop optimizer-based LSTM network.

Epoch	Iteration	Time Elapsed (hh : mm : ss)	Mini- batch Accuracy	Validation Accuracy	Mini- batch Loss	Validation Loss	Base Learning Rate
1	1	00 : 00 : 09	25 . 00%	42 . 71%	1 . 3862	1 . 3707	0 . 0010
3	50	00 : 00 : 15	68 . 75%	58 . 33%	0 . 9001	1 . 9768	0 . 0010
5	100	00 : 00 : 19	100 . 00%	79 . 17%	0 . 2310	0 . 6740	0 . 0010
7	150	00 : 00 : 24	100 . 00%	81 . 25%	0 . 1316	0 . 5711	0 . 0010
9	200	00 : 00 : 28	93 . 75%	83 . 33%	0 . 1009	0 . 4787	0 . 0010
11	250	00 : 00 : 35	100 . 00%	86 . 46%	0 . 0383	0 . 4170	0 . 0010
13	300	00 : 00 : 39	93 . 75%	87 . 50%	0 . 0661	0 . 3189	0 . 0010
15	350	00 : 00 : 43	100 . 00%	87 . 50%	0 . 0164	0 . 3744	0 . 0010
17	400	00 : 00 : 48	100 . 00%	89 . 58%	0 . 0003	0 . 2837	0 . 0010
19	450	00 : 00 : 52	100 . 00%	89 . 58%	0 . 0006	0 . 3176	0 . 0010
21	500	00 : 00 : 56	100 . 00%	90 . 62%	0 . 0800	0 . 2833	0 . 0010
23	550	00 : 00 : 01	100 . 00%	89 . 58%	0 . 0026	0 . 3078	0 . 0010
25	600	00 : 01 : 05	100 . 00%	90 . 62%	0 . 0109	0 . 3017	0 . 0010
28	650	00 : 01 : 09	100 . 00%	91 . 67%	0 . 0002	0 . 2817	0 . 0010
30	700	00 : 01 : 13	100 . 00%	92 . 71%	0 . 0030	0 . 3013	0 . 0010
30	720	00 : 01 : 15	100 . 00%	91 . 67%	0 . 0008	0 . 2761	0 . 0010

FIGURE 10: Epoch and iteration-wise mini-batch training and validation accuracy analysis of the RMSprop optimizer-based LSTM network along with respective losses and base learning rate.

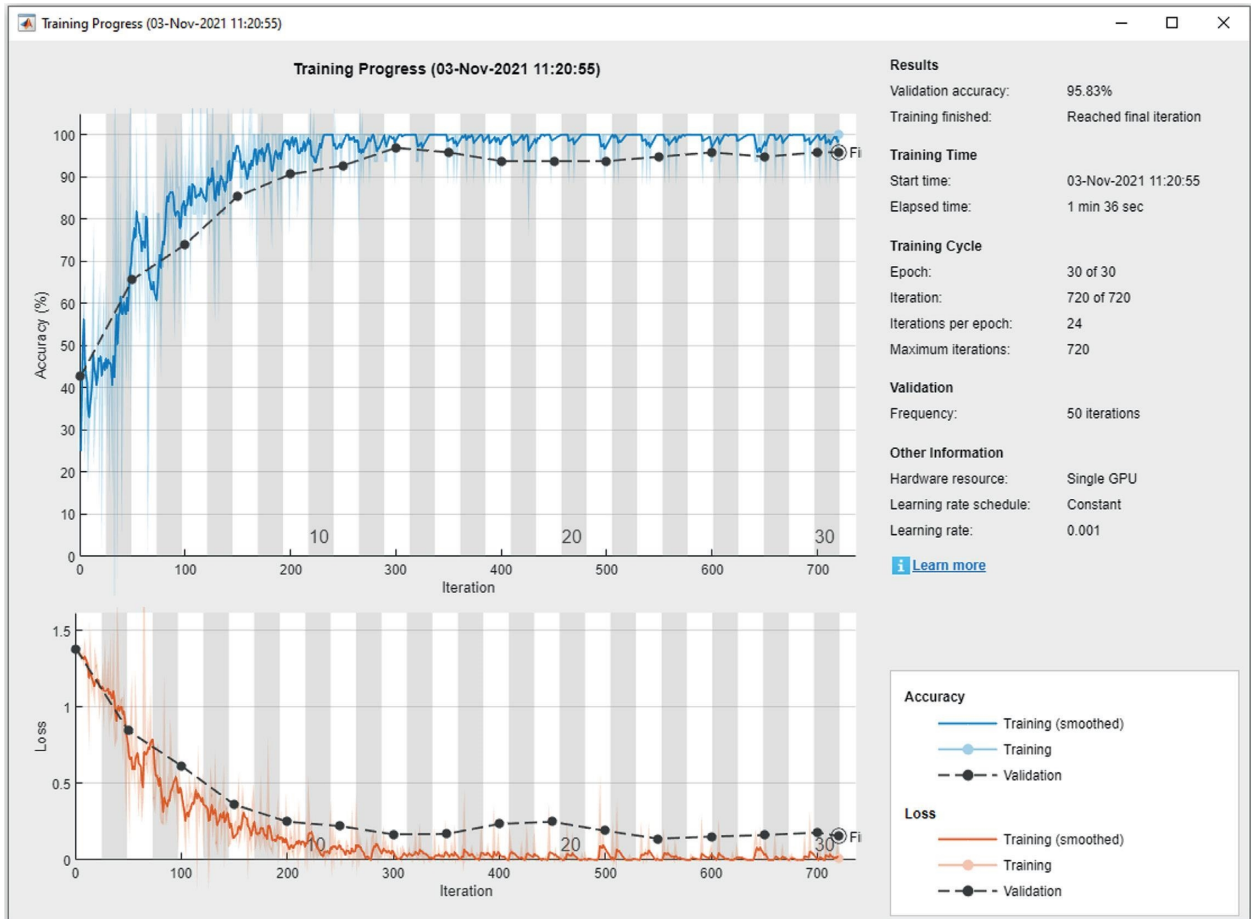


FIGURE 11: Training and validation analysis of the proposed ELSTM network.

Epoch	Iteration	Time Elapsed (hh : mm : ss)	Mini- batch Accuracy	Validation Accuracy	Mini- batch Loss	Validation Loss	Base Learning Rate
1	1	00 : 00 : 14	25 . 00%	42 . 71%	1 . 3857	1 . 3772	0 . 0010
3	50	00 : 00 : 22	75 . 00%	65 . 62%	0 . 7280	0 . 8456	0 . 0010
5	100	00 : 00 : 27	68 . 75%	73 . 96%	0 . 4225	0 . 6157	0 . 0010
7	150	00 : 00 : 33	93 . 75%	85 . 42%	0 . 2246	0 . 3623	0 . 0010
9	200	00 : 00 : 38	93 . 75%	90 . 62%	0 . 0442	0 . 2503	0 . 0010
11	250	00 : 00 : 43	100 . 00%	92 . 71%	0 . 0556	0 . 2214	0 . 0010
13	300	00 : 00 : 51	100 . 00%	96 . 88%	0 . 0248	0 . 1673	0 . 0010
15	350	00 : 00 : 57	100 . 00%	95 . 83%	0 . 0299	0 . 1696	0 . 0010
17	400	00 : 01 : 02	100 . 00%	93 . 75%	0 . 0025	0 . 2358	0 . 0010
19	450	00 : 01 : 07	100 . 00%	93 . 75%	0 . 0071	0 . 2497	0 . 0010
21	500	00 : 01 : 12	100 . 00%	93 . 75%	0 . 0022	0 . 1910	0 . 0010
23	550	00 : 01 : 18	100 . 00%	94 . 79%	0 . 0007	0 . 1394	0 . 0010
25	600	00 : 01 : 24	93 . 75%	95 . 83%	0 . 0809	0 . 1528	0 . 0010
28	650	00 : 01 : 29	100 . 00%	94 . 79%	0 . 0004	0 . 1628	0 . 0010
30	700	00 : 01 : 34	100 . 00%	95 . 83%	0 . 0049	0 . 1793	0 . 0010
30	720	00 : 01 : 36	100 . 00%	95 . 83%	0 . 0109	0 . 1564	0 . 0010

FIGURE 12: Epoch and iteration-wise mini-batch training and validation accuracy analysis of the proposed ELSTM network along with respective losses and base learning rate.

5. Conclusion

From the extensive review, it has been found that the LSTM network suffers from the parameter tuning problem. Initial and control parameters of LSTM have been selected purely on a trial and error basis. To overcome this issue, an ELSTM network has been proposed. MOGA was utilized to optimize the architecture and weights of LSTM. The proposed model has been tested on a well-known factory reports dataset. Extensive analyses have been performed to evaluate the performance of the proposed ELSTM network. From the comparative analysis, it has been found that the LSTM network outperforms the competitive models. Compared to the LSTM variants, the proposed ELSTM network achieves approximately 4.2389% validation accuracy.

Data Availability

The data collected during the data collection phase are available from the corresponding author upon request.

Conflicts of Interest

The authors would like to confirm that there are no conflicts of interest regarding the study.

References

- [1] Y. Ma, X. Liu, L. Zhao, Y. Liang, P. Zhang, and B. Jin, "Hybrid embedding-based text representation for hierarchical multi-label text classification," *Expert Systems with Applications*, vol. 187, Article ID 115905, 2022.
- [2] M. Liu, L. Liu, J. Cao, and Q. Du, "Co-attention network with label embedding for text classification," *Neurocomputing*, vol. 471, pp. 61–69, 2021.
- [3] Y. Xiao, Y. Li, J. Yuan, S. Guo, Y. Xiao, and Z. Li, "History-based attention in seq2seq model for multi-label text classification," *Knowledge-Based Systems*, vol. 224, Article ID 107094, 2021.
- [4] J. Jang, Y. Kim, K. Choi, and S. Suh, "Sequential targeting: a continual learning approach for data imbalance in text classification," *Expert Systems with Applications*, vol. 179, Article ID 115067, 2021.
- [5] H. Liu, G. Chen, P. Li, P. Zhao, and X. Wu, "Multi-label text classification via joint learning from label embedding and label correlation," *Neurocomputing*, vol. 460, pp. 385–398, 2021.
- [6] J. Deng, L. Cheng, and Z. Wang, "Attention-based bilstm fused cnn with gating mechanism model for Chinese long text classification," *Computer Speech & Language*, vol. 68, Article ID 101182, 2021.
- [7] D. Tian, M. Li, J. Shi, Y. Shen, and S. Han, "On-site text classification and knowledge mining for large-scale projects construction by integrated intelligent approach," *Advanced Engineering Informatics*, vol. 49, Article ID 101355, 2021.
- [8] J. Ren, W. Wu, G. Liu, Z. Chen, and R. Wang, "Bidirectional gated temporal convolution with attention for text classification," *Neurocomputing*, vol. 455, pp. 265–273, 2021.
- [9] Y. Liu, P. Li, and X. Hu, "Combining context-relevant features with multi-stage attention network for short text classification," *Computer Speech & Language*, vol. 71, Article ID 101268, 2022.
- [10] K. Thirumorthy and K. Muneeswaran, "Feature selection using hybrid poor and rich optimization algorithm for text classification," *Pattern Recognition Letters*, vol. 147, pp. 63–70, 2021.
- [11] A. Viswakumar, P. B. Ganganaik, P. M. P. Raj, B. P. Rao, and S. Kundu, "Memristor-based in-memory processor for high precision semantic text classification," *Computers & Electrical Engineering*, vol. 92, Article ID 107160, 2021.
- [12] X. Li, M. Cui, J. Li, R. Bai, Z. Lu, and U. Aickelin, "A hybrid medical text classification framework: integrating attentive rule construction and neural network," *Neurocomputing*, vol. 443, pp. 345–355, 2021.
- [13] C. Chen and J. Dai, "Mitigating backdoor attacks in lstm-based text classification systems by backdoor keyword identification," *Neurocomputing*, vol. 452, pp. 253–262, 2021.
- [14] X. Zhang, J. Xu, C. Soh, and L. Chen, "La-hcn: label-based attention for hierarchical multi-label text classification neural network," *Expert Systems with Applications*, vol. 187, Article ID 115922, 2022.
- [15] X. Luo, "Efficient English text classification using selected machine learning techniques," *Alexandria Engineering Journal*, vol. 60, no. 3, pp. 3401–3409, 2021.
- [16] G. Liu and J. Guo, "Bidirectional lstm with attention mechanism and convolutional layer for text classification," *Neurocomputing*, vol. 337, pp. 325–338, 2019.
- [17] Z. S. Ouyang, X. T. Yang, and Y. Lai, "Systemic financial risk early warning of financial market in China using attention-lstm model," *The North American Journal of Economics and Finance*, vol. 56, Article ID 101383, 2021.
- [18] L. Jiang, X. Sun, F. Mercaldo, and A. Santone, "Decab-lstm: deep contextualized attentional bidirectional lstm for cancer hallmark classification," *Knowledge-Based Systems*, vol. 210, Article ID 106486, 2020.
- [19] G. Rao, W. Huang, Z. Feng, and Q. Cong, "Lstm with sentence representations for document-level sentiment classification," *Neurocomputing*, vol. 308, pp. 49–57, 2018.
- [20] Y. Zhang, J. Wang, and X. Zhang, "Conciseness is better: recurrent attention lstm model for document-level sentiment analysis," *Neurocomputing*, vol. 462, pp. 101–112, 2021.
- [21] M. B. Er, E. Isik, and I. Isik, "Parkinson's detection based on combined CNN and LSTM using enhanced speech signals with v," *Biomedical Signal Processing and Control*, vol. 70, Article ID 103006, 2021.
- [22] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate lstm-fcns for time series classification," *Neural Networks*, vol. 116, pp. 237–245, 2019.
- [23] R. K. Behera, M. Jena, S. K. Rath, and S. Misra, "Co-lstm: convolutional lstm model for sentiment analysis in social big data," *Information Processing & Management*, vol. 58, no. 1, Article ID 102435, 2021.
- [24] T. Y. Kim and S. B. Cho, "Optimizing cnn-lstm neural networks with pso for anomalous query access control," *Neurocomputing*, vol. 456, pp. 666–677, 2021.
- [25] F. Shahid, A. Zameer, and M. Muneeb, "A novel genetic lstm model for wind power forecast," *Energy*, vol. 223, Article ID 120069, 2021.
- [26] A. Kara, "Multi-step influenza outbreak forecasting using deep lstm network and genetic algorithm," *Expert Systems with Applications*, vol. 180, Article ID 115153, 2021.
- [27] P. Kumar Shukla, P. Kumar Shukla, P. Sharma et al., "Efficient prediction of drug-drug interaction using deep learning models," *IET Systems Biology*, vol. 14, no. 4, pp. 211–216, 2020.
- [28] S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human

Retraction

Retracted: Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] Q. Li, Y. Liu, E. S. Sipahi Döngül, Y. Yang, X. Ruan, and W. Enbeyle, "Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2290644, 14 pages, 2022.

Research Article

Operating Room Planning for Emergency Surgery: Optimization in Multiobjective Modeling and Management from the Latest Developments in Computational Intelligence Techniques

Qiqian Li,¹ Yali Liu,² Esra Sipahi Döngül,³ Yufen Yang,⁴ Xiaoyuan Ruan,⁵ and Wegayehu Enbeyle ⁶

¹The Seventh Affiliated Hospital of Southern Medical University, Guangzhou, Guangdong, China

²The Sixth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China

³Aksaray University, Faculty of Health Sciences, Department of Social Work, Aksaray, Turkey

⁴Shenzhen University General Hospital, Shenzhen, Guangdong, China

⁵The Eighth Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China

⁶Department of Statistics, Mizan-Tepi University, Tepi, Ethiopia

Correspondence should be addressed to Wegayehu Enbeyle; wegayehu@mtu.edu.et

Received 7 December 2021; Revised 20 December 2021; Accepted 28 December 2021; Published 17 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Qiqian Li 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 presents an optimization approach for scheduling the operation room for emergency surgeries, considering the priority of surgeries. This optimization model aims to minimize the costs associated with elective and emergency surgeries and maximize the number of scheduled surgeries. In this study, surgeon assistants to perform each surgery are considered in order to achieve the goals. Since the time of each surgery varies according to the conditions of the patient, this parameter is considered as an uncertain one, and a robust optimization method is applied to deal with uncertainty. To demonstrate the effectiveness of the proposed method, a case study in one of the East Asian hospitals is presented and analyzed using GAMS software. Moreover, hybrid simulation and gray wolf optimization algorithm (GWO) have been implemented to solve the optimization model in different scenarios. The results show that increasing the risk parameters in the robust optimization model will increase the system costs. Moreover, in case of uncertainty, the solutions obtained from the GWO simulation method are on average 73.75% better than the solutions obtained from the GWO algorithm.

1. Introduction

Living a healthy life and looking for solutions to health problems are some of the priorities of mankind in every age. The development of health sciences is made possible by the development of environmental impacts. Although medical science is perceived as a whole in itself, it is possible to see that it actually benefits from many disciplines when looking at the subject in detail. Although basic sciences such as physics, chemistry, and biology are leading the way, the solution proposals put forward by

management and information systems, which have developed rapidly since the middle of the last century, have enabled these systems to be integrated into health sciences as well as social works. Social workers around the world serve as the primary professional leaders of an influential community centres movement [1].

The field of medical social work is mainly aimed at solving urgent crises. In hospitals with a lot of patient turnover, social workers have to perform very fast and effective interventions. Therefore, working in this field requires special skills.

In recent decades, health and treatment costs have been increased, and the operating room is one of the most important areas in the hospital that needs regular management due to its high potential for cost savings [2]. Operating rooms are one of the most expensive and fundamental resources of a hospital. Therefore, poor scheduling may lead to periods of unemployment, overwork, delay in surgery, or cancellation of surgery, which results in additional costs and a lack of income for the hospital [3].

Despite the multiplication of costs in the field of health and the increase in the number of hospitals, increasing services of health systems have not been able to meet the growing demand [4]. Moreover, the most important challenges of health systems are the increasing demand for surgical services, random service times, limited access to resources, and patient-surgeon compatibility [5, 6]. Accordingly, operating room scheduling plays a crucial role in the operating room management.

In this study, a two-objective mathematical model for elective and emergency surgeries is examined by considering the priority of emergency patients over selected patients and assigning surgeon assistants to each surgery. Since the time of each surgery cannot be predicted, the uncertainty is considered for this parameter, and a robust optimization method is used to deal with uncertainty. The optimization method, which is one of the issues of interest to management sciences, comes to the aid of health and social service management at this point.

In this research, given the above and the sensitivity of operating room scheduling, an attempt has been made to present a model with the following innovations:

Providing a two-objective mixed-integer model for scheduling the operating room, which includes (1) minimizing surgical costs, including the cost of shifting and the cost of not performing emergency surgery due to the operating room being not empty and (2) maximizing the number of surgeries performed

Simultaneous consideration of surgery of emergency patients and selected patients

Optimization of the mathematical model by considering the priority of emergency surgeries over selected surgeries

Assigning a surgeon's assistant for each surgery

Optimization of uncertainty problem by considering the surgical time parameter

Providing a robust optimization method to deal with uncertainty in the studied problem

Proposing a novel solution method based on the simulation and gray wolf optimization algorithm

The rest of the study is as follows: In Section 2, the studies conducted in this field are reviewed. In Section 3, the proposed operating room scheduling model is presented, and then, in Section 4, the uncertainty conditions for the surgical time parameter are applied and a robust optimization method to deal with the uncertainty is proposed. In Section 5, the results of solving the certain model as well as

solving the uncertain model for different levels of conservatism are presented and compared, and in Section 6, the conclusion and future directions of the research are presented.

2. Literature Review

The scheduling problem has been used in the manufacturing industry for many years but has recently come to the attention of many service industries, including the healthcare systems. The first research on operating room planning began in 1953 with Adair [7]. Surgeons believed that if enough space and assistance were available, surgeries can be increased by up to 50%. In the 1970s, Dal Ceilo [8] proposed a study of idle times in the operation room. In 1984, Charnetski [9] conducted one of the first studies of reducing idle time in healthcare systems.

Some researches in this field have examined the effect of patient preference and surgical timing on operating room planning in the context of social works. Min and Yih [2] examined a scheduling problem in which patients had different priorities and the facility capacity for a selective surgery was limited; in this way, when capacity is available, patients are selected with a higher priority than the waiting list and placed on a schedule. Abedini et al. [10] proposed a multistep approach to assigning surgeries to operating rooms and a priority scheduling rule. This rule categorizes surgeries based on their priority, type of surgery, and time, and their use reduces unemployment time, reduces the number of operating rooms required, and reduces the number of operating room startups.

Admission of emergency patients is an important issue for operating rooms, which can lead to delays throughout the system if resources are inadequate. Erdem et al. [11] proposed a model for rescheduling selected patients as soon as emergency patients arrived, intending to examine the overwork costs of the operating room and postanesthesia unit, the costs of delaying elective surgeries, and the costs of not admitting emergency patients. In a study by Lamiri et al. [12], a stochastic model for operating room planning is presented with two types of surgery requests (elective and emergency) in which emergency patient admission is random. In another study by these researchers [10], they examined an uncertain model for scheduling selective surgeries, considering the demand uncertainty for emergency surgeries.

One of the critical points in scheduling emergency surgeries is scheduling them as quickly as possible because delaying them may increase the risk. Essen et al. [13] examined emergency surgeries in one of the operating rooms of choice. In this study, emergency patients undergo surgery as soon as an elective surgery is completed. Reserving operating room capacity as soon as an emergency patient arrives is a common way to maximize responsiveness. Bhattacharyya et al. investigated a case in which an operating room was reserved only for essential surgeries, and additional surgeries could be performed if an unreserved room was not available.

Equipment plays a crucial role in hospitals, mainly operating rooms, and their lack or unavailability may lead to surgical arrests as well as the death of patients. Therefore, planning and scheduling hospital equipment is essential, especially if the equipment is special or of great importance [12]. Therefore, some researchers have examined the issue of operating room scheduling by considering the required resources and equipment.

Beds are a scarce resource in the intensive care unit, and patients' random demand for these beds, as well as incidental service time, makes resource management a difficult issue. In a study using a simulation model, Seung-Chul et al. [14] evaluated different bed reservation schemes because they considered reserving some beds for the exclusive use of selected patients as one way to minimize canceled surgeries.

Roland et al. conducted a study considering the limitations of human resources and considering their preferences and the limitations of the volume of reusable materials and equipment with the aim of minimizing the costs due to overtime and operating room use. According to Latorre-Núñez et al. [3], in operating room scheduling, each surgery requires a specific combination of human resources, equipment and materials, and a bed to resuscitate the patient, and the timing should be such that in addition to scheduled surgeries, emergency surgeries that may occur during the day should also be considered. Silva et al. [15] modeled surgical scheduling using specialized human resources and social workers based on their skills as well as staff time window with the aim of maximizing the use of operating rooms. Molina-Pariente et al. [16] presented a model for solving the problem of operating room planning and scheduling in which the effect of the number of surgeons on surgical teams and the dependence of surgical time on their experience and skills were examined.

Xiang [17] considered that operating room scheduling is generally a multiobjective problem. Next, they provided a meta-heuristic approach to solve the proposed multi-objective scheduling problems. Saadouli et al. [18] proposed a stochastic optimization and simulation approach for scheduling operating rooms and resuscitation beds with the aim of minimizing the maximum operating time and the total waiting time of each operating room. The effect of allowing the patient to be resuscitated in the operating room when the resuscitation bed is not available was investigated by Augusto et al.

The problem of resource uncertainty often coincides with the uncertainty of patient admission. For example, the arrival of emergency patients may be accompanied by complaints of the surgeon's absence for surgery and the unavailability of an operating room, which can lead to delays in elective surgeries [16]. Therefore, some researchers have examined operating room scheduling issues with uncertainty in mind. Persson and Persson [17] examined operating room planning, taking into account uncertainties in inpatient admission requests and surgical timing. Addis et al. [19] examined the issue of selecting a set of patients from the waiting list of selected patients and allocating them to

available operating rooms. For each patient, the maximum waiting time and uncertainty during surgery are considered, and new patients' entry in their research is examined.

Liu et al. [20] studied surgical scheduling by simultaneously examining surgery priorities and operating room scheduling decisions. Considering the uncertainty at the time of surgery, they presented a robust model for minimizing surgical time, which reduced the surgical time by considering scenarios compared to definitive scheduling methods. Kroer et al. [21] examined the planning of operating rooms for a hospital in Denmark, in which each surgery must be assigned to a specific operating room, as well as surgery for a specific time in order to minimize overtime and use capacity. Nouaouri et al. [22] considered the uncertainty of inpatient admission and surgery time. They solved the problem by assuming a limited number of surgeons and operating rooms and with the aim of maximizing the number of surgeries performed in operating rooms. Khaniyev et al. [23] have dealt with operating room scheduling in case of surgery duration uncertainty. In this study, in which scheduling is done for only one day, first, a mathematical model is developed, and then, an innovative algorithm is introduced. Rahimi and Gandomi [24] provide a comprehensive overview of operating room scheduling. In this study, various optimization models and solution methods are introduced, and then, the importance of uncertainty in this issue is emphasized. Lin and Li [25] have solved the problem of operating room scheduling with meta-heuristic methods. In this study, a modified mathematical model is provided to optimize operating room costs as well as increase operating room utilization. The artificial bee colony (ABC) algorithm has also been used to optimize this mathematical model. The results show that for a number of surgeries between 40 and 100, the ABC algorithm can provide the optimal solution in a short time.

As the literature shows, in many studies, selective surgeries have been examined in the operating room scheduling problem, and the arrival of emergency patients has been ignored. However, the arrival of emergency patients is an unavoidable issue, and not paying attention to emergency surgery cases may have life-threatening risks for patients. Uncertainty in surgical timing is also one of the issues that have been investigated in quantitative research if the timing of surgeries is not predictable and can only be estimated by examining similar previous surgeries, which is based on the specific characteristics of each surgery cannot be invoked. Moreover, having a surgeon's assistant with the surgeon for each surgery can lead to improvements in surgery time, unemployment time, overtime time and thus reduce surgery costs, as well as increase the quality of surgery and patient satisfaction. This issue, despite its importance, needs further study and research.

Considering the mentioned cases, despite their importance, has not been studied in a comprehensive study in previous researches. Accordingly, the purpose of this study is to provide a model to examine these research gaps in order to increase patient satisfaction.

3. Problem Statement

In this research, the efficiency of the proposed model of one of the Asian hospitals is studied. According to operating room officials, in hospitals, in addition to specialized operating rooms located in each ward, a number of operating rooms for surgeries that do not require specialized equipment exist in each ward. In this study, the number of operating rooms in the hospital is four, and the different types of surgeries performed in these rooms are 4.

In this research, to solve the multiobjective model, the LP-metric method is used in order to minimize the relative deviations of the goals from their optimal value. Moreover, the time of surgery is unpredictable due to the fact that it is affected by several factors. It is considered as an uncertainty parameter, and a robust optimization method is used to deal with uncertainty.

The notations and symbols used in the proposed mathematical programming model are defined in the following.

3.1. Indices.

- i : selected patient index
- h : emergency patient index
- j : index of type of surgery
- r : index of operating room
- t : index of time period
- a : index of collection of surgeon assistants

3.2. Parameters.

- Bp : capacity of the postoperative care unit (resuscitation beds) along the scheduling horizon
- Brt : capacity of operating room
- Bat : assistant surgeon a working time
- $C_{jt'}$: the cost of postponing an elective surgery of type j scheduled for time t' to time t
- git : binary parameter and equal to 1 if the patient's surgery i is scheduled for period t ; 0 otherwise
- H : number of emergency patients requesting surgery
- I : the number of selected patients whose surgery is scheduled on the time horizon
- J : the number of types of surgeries that can be performed with facilities
- m_{ij} : 1 if the selected patient i requests a type j surgery; 0 otherwise
- m'_{hj} : 1 if the emergency patient h requests a type j surgery; 0 otherwise
- R : number of operating rooms

- O_j : the expected time for a type j surgery
- r_j : the cost of not performing type j emergency surgery on time
- S_j : the length of stay in the postoperative care unit for a patient who has type j surgery
- ts : time to start surgery
- T : number of time periods on the scheduling horizon
- U^p : the upper limit of resuscitation beds in the postoperative care unit
- λ_t : number of beds occupied from the previous scheduling cycle in the postoperative care unit at time t
- A : number of surgeon assistants available to perform surgery
- Ω : maximum allowable delay in surgery
- \hat{O}_j : the amount of time error expected for a type j surgery (in case of uncertainty)
- Γ : stable level value (in case of uncertainty)

3.3. Decision Variables.

- Op : number of extra beds in the postoperative care unit
- S_{it} : binary variable and equal to 1 if the patient i chooses a bed in the PACU at time t ; 0 otherwise
- S'_{ht} : binary variable and equal to 1 if the emergency patient h occupies a bed in the PACU at time t ; 0 otherwise
- X_{irt} : binary variable and equal to 1 if surgery is started at the beginning of time t for the selected patient i in operating room r ; 0 otherwise
- X'_{hrt} : binary variable and equal to 1 if surgery is started at the beginning of time t for emergency patient h in operating room r ; 0 otherwise
- y_{irt} : binary variable and equal to 1 if the selected patient i has a surgery at time t in the operating room r ; 0 otherwise
- y'_{hrt} : binary variable and equal to 1 if the emergency patient h has a surgery at time t in the operating room r ; 0 otherwise
- Q_{iart} : binary variable and equal to 1 if elective surgery i is assigned to surgeon a , operating room r and time t ; 0 otherwise
- Q'_{hart} : binary variable and equal to 1 if emergency surgery h is assigned to surgeon a , operating room r , and time t ; 0 otherwise
- Fa : end time of last surgery of assistant surgeon a
- λ : the dual variable corresponding to the uncertainty protection function

3.4. Mathematical Modeling of the Problem

3.4.1. Objective Function.

$$\begin{aligned} \text{Min} Z1 = & \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R \sum_{t=t_s}^{T+1} \sum_{t'=t_s}^{T+1} (1 - g_{it}) * C_{jit'} * m_{ij} * X_{irt}, \\ & + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R \sum_{t_s=1}^{T+1} r_j * m'_{hj} * (1 - X'_{hrt_s}). \end{aligned} \quad (1)$$

Objective function (1) minimizes the cost of delay in elective surgeries and the cost of not accepting emergency surgeries.

$$\text{Max } Z2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R \sum_{t_s=1}^{T+1} m_{ij} * X_{irt} + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R \sum_{t_s=1}^{T+1} m'_{hj} * X'_{hrt}. \quad (2)$$

Objective function (2) maximizes the number of scheduled surgeries.

3.4.2. Constraints.

$$\sum_{i=1}^I X_{irt} + \sum_{h=1}^H X'_{hrt} \leq 1; \forall r, t, \quad (3)$$

$$\sum_{i=1}^I \sum_{j=1}^J m_{ij} * X_{irt} * O_j + \sum_{h=1}^H \sum_{j=1}^J m'_{hj} * X'_{hrt} * O_j \leq B_{rt}; \forall r, t, \quad (4)$$

$$\begin{aligned} & \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R m_{ij} * O_j * Q_{iart} + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R m'_{hj} * O_j * Q'_{hart} \\ & \leq B_{at}; \forall a, t, \end{aligned} \quad (5)$$

$$\sum_{r=1}^R \sum_{t=t_s}^T X_{irt} = 1; \forall i, \quad (6)$$

$$\sum_{r=1}^R X_{irt} \leq g_{it}; \forall i, t, \quad (7)$$

$$\sum_{r=1}^R \sum_{t=t_s+1}^T X'_{hrt} = 0; \forall h, \quad (8)$$

$$\sum_{i=1}^I \sum_{r=1}^R X_{irt} + \sum_{h=1}^H \sum_{r=1}^R X'_{hrt} + \sum_{i=1}^I \sum_{r=1}^R y_{irt} + \sum_{h=1}^H \sum_{r=1}^R y'_{hrt} \leq R; \forall t, \quad (9)$$

$$\sum_{i=1}^I \sum_{r=1}^R X_{irt} + \sum_{h=1}^H \sum_{r=1}^R X'_{hrt} + \sum_{i=1}^I \sum_{r=1}^R y_{irt} + \sum_{h=1}^H \sum_{r=1}^R y'_{hrt} \leq A; \forall t, \quad (10)$$

$$y_{irt'} \geq m_{ij} * X_{irt}; \forall i, j, r, t, t', \quad (11)$$

$$y'_{hrt} \geq m'_{hj} * X'_{hrt}; \forall h, j, r, t, t', \quad (12)$$

$$s_{it'} \geq \sum_{r=1}^R m_{ij} * X_{irt}; \forall i, r, t, t', \quad (13)$$

$$s'_{ht'} \geq \sum_{r=1}^R m'_{hj} * X'_{hrt}; \forall h, r, t, t', \quad (14)$$

$$\sum_{a=1}^A Q_{iart} \geq m_{ij} * X_{irt}; \forall i, j, r, t, \quad (15)$$

$$\sum_{a=1}^A Q'_{hart} \geq m'_{hj} * X'_{hrt}; \forall h, j, r, t, \quad (16)$$

$$\sum_{i=1}^I s_{it} + \sum_{h=1}^H s'_{ht} + \lambda_t \leq B^P + O^P; \forall t, \quad (17)$$

$$O^P \leq U^P, \quad (18)$$

$$\begin{aligned} F_a \geq & \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R m_{ij} * O_j * Q_{iart} \\ & + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R m'_{hj} * O_j * Q'_{hart}; \forall a, t, \end{aligned} \quad (19)$$

$$X_{irt} \leq \sum_{t'=t_s}^T X'_{hrt'}; \forall i, j, r, t, \quad (20)$$

$$\sum_{a=1}^A \sum_{r=1}^R \sum_{t=t_s}^T Q_{iart} = 1; \forall i, \quad (21)$$

$$\sum_{a=1}^A \sum_{r=1}^R \sum_{t=t_s}^T Q'_{hart} = 1; \forall h, \quad (22)$$

$$\sum_{i=1}^I Q_{iart} + \sum_{h=1}^H Q'_{hart} \leq 1; \forall a, r, t, \quad (23)$$

$$S_{it}, X_{irt}, y_{irt} \in \{0, 1\}; \forall i, r, t, \quad (24)$$

$$S'_{ht}, X'_{hrt}, y'_{hrt} \in \{0, 1\}; \forall h, r, t, \quad (25)$$

$$O^P \geq 0 \text{ an } d \text{ integer}, \quad (26)$$

$$Q_{iart} \in \{0, 1\}, Q'_{hart} \in \{0, 1\}. \quad (27)$$

Equation (3) indicates that a maximum of one surgery can be performed in each operating room at any one time. Equation (4) states that the total time for elective and emergency surgeries should not exceed the total working time of each room. Equation (5) states that the total scheduled time for selective and emergency surgeries by each

assistant surgeon should not exceed the total time available to the surgeon's assistant. Equation (6) allocates the start time to each surgery. Equation (7) prevents scheduling of selected surgeries between hours Ω from the current time t_s that have been rescheduled in the time window $[t_s, t_s + \Omega - 1]$.

Equation (8) ensures that emergency surgeries are performed promptly. Equation (9) states that the total number of newly started and ongoing surgeries in each time period should not exceed the number of operating rooms. Limitation (10) states that the total number of surgeries performed in each time period should not exceed the number of surgeon assistants. Equations (11) and (12) determine the timing of each elective and emergency surgery. Equations (13) and (14) occupy beds in the PACU for a specified period of time indicate the patient of choice and emergency, respectively, until recovery from surgery. Equations (15) and (16) determine each elective and emergency surgery performance by the surgeon's assistant, respectively.

Equation (17) relates to PACU capacity. Equation (18) refers to the number of extra beds that can be added to the PACU according to rules, regulations, and space restrictions. Equation (19) determines the end time of the last surgery for each assistant surgeon. Equation (20) indicates the priority of surgery. Equations (21) and (22) relate to the appointment of an assistant surgeon. Equation (23) states that at any time, the number of surgeries performed simultaneously by each assistant surgeon cannot be more than one, and finally, equations (24)–(27) indicate the type of decision variables.

3.5. Robust Optimization Model. Based on the available historical data, operating room managers provide an estimate of the time of each surgery according to its type, but since this parameter is affected by several factors, it is not predictable and cannot be considered definite values. Therefore, the model presented in this dissertation has uncertainty conditions in the parameter of expected time for surgery.

In this study, we discuss the function of Bertsimas and Sim [26] in the field of linear programming and discrete programming zero and one. In the model he presented for linear optimization, we maintained the benefits of the linear framework proposed by Soyester [27], and the model he presented would provide an answer whose number of indefinite elements is determined by the analyst's discretion.

For example, linear programming with a Bertsimas and Sim's function becomes a robust linear programming model. In this function, one can have complete control over the degree of conservatism of any constraint using the Γ parameter.

We define the initial uncertain model assuming that only the elements of matrix A have uncertainty as follows:

$$\begin{aligned} & \text{Min } C^T X, \\ & \text{s.t: } \sum_j a_{ij} x_j - \max_{|S_i|=\Gamma_i} \left\{ \sum_{j \in S_i} \hat{a}_{ij} x_j \right\} \geq b_i ; \forall i, \\ & x \geq 0. \end{aligned} \quad (28)$$

The assumptions and definitions of the model components are as follows:

S_i : a subset of coefficients in line i that have uncertainty.
 \hat{a}_{ij} : the value of the coefficient in line i where the j th coefficient has uncertainty ($j \in J_i$) $\in S_i$, so that it changes in the interval.

Γ_i : adjustment of the robustness is the proposed method against the level of problem-solving conservatism. Each row has its own Γ_i parameter, which is not necessarily an integer and changes in the $[0, |J_i|]$ range.

In this case, the answer will be justified with a certain probability, which is described in detail in Bertsimas and Sim's dissertation. By obtaining the protection function and its linearization (the proof of relations is omitted), we will obtain the following robust model:

$$\begin{aligned} & \text{Min } C^T X, \\ & \text{s.t: } \sum_j a_{ij} x_j - \sum_{j \in J_i} p_{ij} - \lambda_i \Gamma_i \geq b_i ; \forall i, \\ & \lambda_i + p_{ij} \geq \hat{a}_{ij} x_j ; \forall i, j \in J_i, \\ & p_{ij} \geq 0 ; \forall j \in J_i, \\ & \lambda_i \geq 0 ; \forall i, \\ & x \geq 0. \end{aligned} \quad (29)$$

In interpreting the different values of gamma (degree of conservatism of the objective function) we will have: if the degree of conservatism of each row is equal to zero, then the model acts as a definite model, and if it takes its maximum value, it acts like the sweater approach with the maximum amount of conservatism. This parameter can also have values between these two.

The robust optimization model will look like this after applying changes to limits that have an uncertainty parameter:

$$\begin{aligned} \text{Min } Z1 &= \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R \sum_{t=t_s}^{T+1} \sum_{t'=t_s}^{T+1} (1 - g_{it}) * C_{jtt'} * m_{ij} * X_{irt}, \\ &+ \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R \sum_{t_s=1}^{T+1} r_j * m'_{hj} * (1 - X_{hrt'_s}), \\ \text{min } Z2 &= \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R \sum_{t=t_s}^{T+1} m_{ij} * X_{irt} + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R \sum_{t=t_s}^{T+1} m'_{hj} * X'_{hrt}. \end{aligned} \quad (30)$$

Subject to Equations (1), (6)–(18), (20)–(27), and

$$\begin{aligned}
& \sum_{i=1}^I \sum_{j=1}^J m_{ij} * X_{irt} * \bar{O}_j + \sum_{j=1}^J q1_j + (\lambda1 * \Gamma1), \\
& + \sum_{h=1}^H \sum_{j=1}^J m'_{hj} * X'_{hrt} * \bar{O}_j + \sum_{j=1}^J q2_j \\
& + (\lambda2 * \Gamma2) \leq B_{rt}; \forall r, t, \\
q1_j + \lambda1 & \geq \sum_{i=1}^I m_{ij} * X_{irt} * \hat{O}_j; \forall j = 1, \dots, J; \forall r, t, \\
q2_j + \lambda2 & \geq \sum_{h=1}^H m'_{hj} * X'_{hrt} * \hat{O}_j; \forall j, t, \\
q1_j, \lambda1, q2_j, \lambda2 & \geq 0, \\
& \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R m_{ij} * \bar{O}_j * Q_{iart} + \sum_{j=1}^J q3_j + (\lambda3 * \Gamma3) \\
& + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R m'_{hj} * \bar{O}_j * Q'_{hart} + \sum_{j=1}^J q4_j \\
& + (\lambda4 * \Gamma4) \geq B_{at}; \forall a, t, \\
q3_j + \lambda3 & \geq \sum_{i=1}^I \sum_{r=1}^R m_{ij} * \hat{O}_j * Q_{iart}; \forall j, a, t, \\
q4_j + \lambda4 & \geq \sum_{h=1}^H \sum_{r=1}^R m'_{hj} * \hat{O}_j * Q'_{hart}; \forall j, a, t, \\
q3_j, \lambda3, q4_j, \lambda4 & \geq 0, \\
F_a & \geq \sum_{i=1}^I \sum_{j=1}^J \sum_{r=1}^R m_{ij} * \bar{O}_j * Q_{iart} + \sum_{j=1}^J q5_j + (\lambda5 * \Gamma5), \\
& + \sum_{h=1}^H \sum_{j=1}^J \sum_{r=1}^R m'_{hj} * \bar{O}_j * Q'_{hart} + \sum_{j=1}^J q6_j \\
& + (\lambda6 * \Gamma6); \forall a, t, \\
q5_j + \lambda5 & \geq \sum_{i=1}^I \sum_{r=1}^R m_{ij} * \hat{O}_j * Q_{iart}; \forall j, a, t, \\
q6_j + \lambda6 & \geq \sum_{h=1}^H \sum_{r=1}^R m'_{hj} * \hat{O}_j * Q'_{hart}; \forall j, a, t, \\
q5_j, \lambda5, q6_j, \lambda6 & \geq 0.
\end{aligned} \tag{31}$$

4. Gray Wolf Meta-Heuristic Algorithm

The gray wolf optimizer (GWO) algorithm is a nature-inspired meta-innovative algorithm that mimics the behavior of gray wolves and their leadership hierarchy and their hunting method. The gray wolf algorithm was proposed by Mirjalili et al. [1] based on their group hunting. The gray wolf

is a member of the Canadian wolf family. Gray wolves are at the top of the food chain and prefer to live in groups. On average, their groups are 5–12. Interestingly, they have much stricter social governance. In this way, the alpha wolf is also called the ruling wolf in the group because his instructions must be followed by the group. Alphas are basically responsible for deciding on where to sleep, when to move, and so on.

The second level of grading of gray wolves is beta. Beta is alpha-controlled wolves that assist alpha in decision-making and other group activities. Beta wolf is probably the best candidate for an alpha, playing the role of a deputy for alpha and a moderator for the band. The lowest floor is the omega gray wolf. Omega wolves play the role of victims for other members of the group. They are the last wolves allowed to eat. If the wolf is not alpha, beta, or omega, he is called obedient (or delta). Delta wolves follow alpha and beta and rule occasionally. To mathematically model the social governance of wolves, the most appropriate solution is called wolves [23, 24].

As a result, the second and third better solutions were named wolf β and δ , respectively. The remaining solutions are assumed to be ω . Therefore, in the GWO algorithm, optimization is guided by δ , β and α and ω wolves follow these three categories. Gray wolves bypass prey during hunting. For mathematical modeling, the circumvention of equations (32) and (33) has been proposed [25].

$$\vec{D} = |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)|, \tag{32}$$

$$\vec{X}(t+1) = X_p(t) - \vec{A} \cdot \vec{D}, \tag{33}$$

where t represents the flow repetition, A and C represent the vector coefficient and the hunting position vector, and X represents the position vector of a gray wolf. Vectors A and C are calculated according to the following equations:

$$\vec{A} = 2 \vec{a} \cdot \vec{r}_1 - \vec{a}, \tag{34}$$

$$\vec{C} = 2 \cdot \vec{r}_2. \tag{35}$$

We have that the \vec{a} elements are reduced linearly from 2 to 0 under the iteration path, and $r1$ and $r2$ are the random vectors in the range [0, 1] [28].

Gray wolves have the ability to detect the position of prey and bypass them. Hunting is usually guided by alpha. Beta and delta may also occasionally participate in hunting. Therefore, in an absolute search space, we have no solution for the optimal position (hunting). To mathematically simulate the hunting behavior of gray wolves, we assume that alpha (candidate best solution), beta, and delta are sufficiently aware of the potential hunting position. We need to force the other search agents to update their position according to the position of the best search agents. This operation is performed according to the following equations:

$$\begin{aligned}\vec{D}_\alpha &= |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}|, \\ \vec{D}_\beta &= |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}|, \\ \vec{D}_\delta &= |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}|,\end{aligned}\quad (36)$$

$$\begin{aligned}\vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1 \cdot (\vec{D}_\alpha), \\ \vec{X}_2 &= \vec{X}_\beta - \vec{A}_1 \cdot (\vec{D}_\beta), \\ \vec{X}_3 &= \vec{X}_\delta - \vec{A}_1 \cdot (\vec{D}_\delta),\end{aligned}\quad (37)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3}.\quad (38)$$

In short, in the GWO algorithm, the search process begins with the creation of a random population of gray wolves (candidate solutions). During the iteration period, alpha, beta, and delta wolves estimate the probable hunting position. Each candidate solution updates its distance with the bait. Parameter A is reduced from 2 to 0 to enhance the detection and attack process. When $|A| > 1$, the candidate solutions diverge, and when $|A| < 1$, the candidate solutions converge.

5. Results

This section analyzes the proposed model and its results. To evaluate the efficiency of the proposed model, one of the Asian hospitals is studied. In order to validate the proposed model for the definite and solid problem, GAMS 24.0 software and CPLEX method are used, and the solution results are analyzed in the certain and robust model, and then, we will provide the sensitivity analysis of the basic parameters of the model.

According to Tables 1 and 2, emergency surgeries are scheduled for the first time due to their high priority over other surgeries, and selective surgeries are scheduled for other time periods. The planning of the operating room where each surgery should be performed is also shown in Table 1. In Table 2, in addition to scheduling surgeries, the surgical assistant who must be present to perform each surgery is also scheduled. As it is known, due to the limitation of the number of operating rooms, it is possible to perform four surgeries at the same time, so in each part of the time, four surgeons are working.

Figures 1 and 2 show the first objective function considering the uncertainty coefficient of 0.05, 0.1, 0.25, and 0.5. As it turns out, the amount of costs increases with increasing uncertainty. On the other hand, by constantly considering the number of selected patients, costs decrease with an increasing number of emergency patients (Figure 1), and also by constantly considering the number of emergency patients, costs increase with an increasing number of selected patients (Figure 2). This may be due to the fact that the selected patients have more surgical transfers than the program envisaged by the management, which incurs costs.

TABLE 1: Optimal solution of operation room scheduling.

	T1	T2	T3	T4	T5	T6	T7	T8
R1	H2	I14	I3		I20	I8	I13	I11
R2	H3	I7	I2	I15	I21	I6	I23	I5
R3	H1	I9	I25	I19	I10	I22	I24	I1
R4	H4	I16	I18		I4		I12	I17

TABLE 2: Optimal solution of surgical assistant scheduling.

	T1	T2	T3	T4	T5	T6	T7	T8
A1	H3		I25					I11
A2		I14	I18					
A3					I10			
A4		I9				I6	I24	I17
A5		I7	I2				I12	I5
A6	H4	I16			I21	I22		
A7	H1			I19	I4		I23	
A8	H2		I3	I15	I20	I8	I13	I1

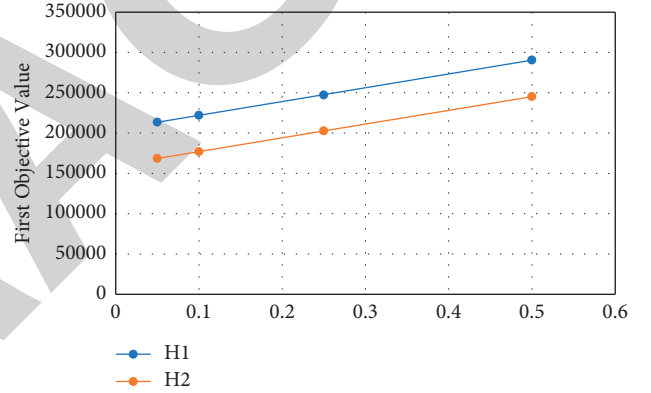


FIGURE 1: Results of the first objective function for $I=10$ and for $H=1$ and $H=2$ under uncertainty.

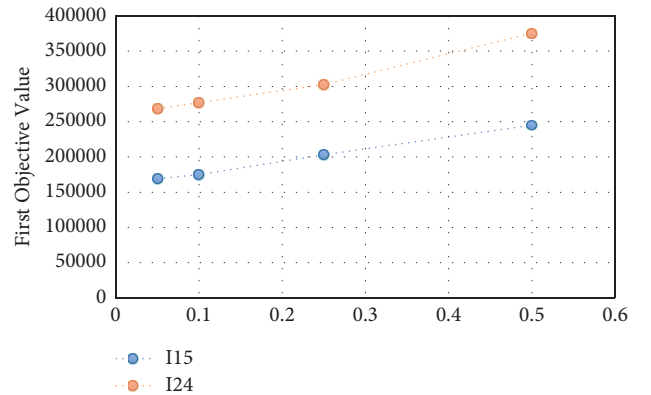


FIGURE 2: Results of the first objective function for $H=3$ and for $I=15$ and $I=24$ under uncertainty.

5.1. Evaluation of the Efficiency of the Proposed Solution Method. Endoscopic surgery is not considered in this study. There are 12 residents and 2 fellowships in this hospital, which have the ability to perform certain operations

depending on the academic year and the experience gained. In addition to the surgeon, each surgery requires a team of technicians (anesthesiologist, anesthesiologist, and operating room technician). In some surgeries, in addition to the surgeon, the surgeon's help is also needed. There is a recovery bed as well as medical equipment related to each type of operation.

In that study, data from all operations performed in operating rooms over two years were collected from a hospital database. These data include a list of patients at each date, operation code, operation name, start and end time of operation, name of the surgeon, and assistant surgeon. Some of the information was also collected from experts is as follows:

- (i) Surgeon data: the surgeries that each surgeon is able to perform, the residents' training program (including the operations they must perform in each course), the operations that each surgeon can perform
- (ii) Equipment related: types of equipment available and their inventory, type, and number of equipment used for each operation
- (iii) Surgery related: types of surgery (urology, laparoscope, etc.), surgeries that require the help of a surgeon
- (iv) Operating room related: the number of operating rooms, rooms equipped with special equipment, surgeries that can be performed in each room
- (v) Technician groups related: the number of technician groups in each shift per day

Using the collected data and frequency graphs and the duration of different operations, the discrete probability distribution of the duration of each type of operation has been approximated. The time distribution of different operations is completely different.

Actual data from the 15 days of the study hospital were used to evaluate the proposed approach. The information for each issue is information about the actual day of surgery of the hospital. This information includes a list of surgeries per day, type of surgery, possible surgeons for each surgery, number of operating rooms available, number of technicians available, number of beds in the preoperative and recovery wards, equipment needed for each surgery, the available number of equipment, and the surgeon's assistants are possible for each operation.

For the duration of the surgeries in the definite case, the average time of the surgeries is used, and for the probable case, the probability distribution function of the times of each surgery is used.

First, the problem is solved by considering the average duration of surgery in the definite case using the gray wolf algorithm. In this algorithm, the number of wolves per repetition is equal to 10, and the condition for stopping the algorithm from achieving a better answer in 100 consecutive repetitions, the position change coefficient equal to 0.03, the impact factor of the best position of each wolf equal to 1, and

the impact factor of the wolf community equal to 3 are considered. The parameters of the algorithm are set by running the algorithm on a sample of different problems and using the method of simultaneous tuning of the parameters by using self-organizing neural networks.

To select the most effective innovative information in the design of the GWO algorithm, the following seven scenarios have been designed:

Scenario 1: selection of the next surgery as well as allocation of resources to each stage of the surgery. It should be done without considering the innovative information and only based on the amount of pheromone.

Scenario 2: the first level of innovative information is based on LPT, and the second level of innovative information is based on ES.

Scenario 3: level 1 of innovative information based on LPT and level 2 of innovative information for surgeon selection based on LFM and for other sources based on ES.

Scenario 4: the first level of innovative information is based on LFJ, and the second level of innovative information is based on ES.

Scenario 5: initiative information of the first level for the first surgery is based on LFJ and for other surgeries is based on LPT, and innovative information of the second level is to select the surgeon of the first surgery based on LFM and for other surgeries based on ES.

Scenario 6: initial level 1 information for selecting the first two surgeries based on LFJ and for other surgeries based on LPT, and innovative level 2 information for selecting the surgeon for the first two surgeries based on LFM and for other surgeries based on ES.

Scenario 7: level 1 of innovative information based on LFJ and level 2 innovative information for surgeon selection based on LFM and for other sources based on ES.

The data for the 15 actual samples are given in Table 3. Each instance is solved in 10 replications, and the mean of the objective function under different scenarios is given in Table 3.

In this research, the performance of the GWO algorithm for optimizing the robust model with the optimal solution of the mathematical model and the answer obtained from the algorithm of Xiang et al. The authors of [5] designed to test the performance of the GWO algorithm, the answer obtained from this algorithm in the definite state in 15 real problem samples with the optimal definite problem solution obtained from solving the mathematical programming model in GAMS software and the answer obtained are compared by the algorithm of Xiang et al.

Designed to test the performance of the GWO algorithm, the answer obtained from this algorithm in the definite state in 15 real problem samples with the optimal definite problem solution obtained from solving the mathematical programming model in GAMS software and the answer obtained are compared by the algorithm of Xiang et al. [5].

TABLE 3: Results of different scenarios of different combinations of innovative rules in GWO algorithm design.

Test problem	Number of surgeries	Number of surgeons	Number of operating rooms	Average objective value						
				Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5	Scenario 6	Scenario 7
1	16	9	6	4.32	4.11	2.05	0.68	0.22	0.00	0.00
2	23	11	7	53.87	57.83	43.11	49.56	0.22	3.31	0.00
3	27	12	7	141.16	29.11	29.16	34.53	19.87	23.47	0.00
4	22	10	7	7.19	0.92	3.87	0.00	0.49	2.37	0.00
5	16	12	7	0.00	0.26	0.00	0.00	0.00	0.00	0.00
6	27	9	7	47.16	3.48	5.68	2.97	2.47	3.83	0.00
7	25	9	7	94.18	78.15	84.19	104.35	23.19	20.04	0.00
8	18	11	7	2.41	0.00	0.00	0.00	0.00	0.00	0.00
9	18	9	7	57.19	20.19	74.83	34.86	0.00	0.25	0.00
10	22	11	7	84.19	32.87	37.89	39.96	34.18	32.73	27.92
11	21	8	7	16.79	1.98	7.88	2.64	2.54	1.44	0.00
12	17	9	7	4.53	0.00	0.00	0.81	0.00	0.00	0.00
13	20	16	8	5.79	16.23	6.75	0.00	0.00	0.00	0.00
14	22	12	7	3.72	0.25	0.00	0.00	0.00	0.00	0.00
15	19	9	8	26.93	20.19	37.14	18.92	3.19	2.19	0.00
		Average		36.63	17.70	22.17	19.29	5.76	5.98	1.86

Table 4 shows the impact of using innovative rules in the design of the GWO algorithm. The numbers in the first column of this table indicate the state of nonuse of innovative information, which is significantly higher than the numbers in the other columns. The following columns each show the different combinations of considered innovative rules. The results of Table 4 show that Scenario 7 has the lowest rates of overtime and unemployment; Scenario 7 is therefore used in the design of the GWO algorithm. It also compares the best solution obtained from the GWO method with the optimal solution obtained from the mathematical programming model of the problem in the definite case and the solution obtained from the algorithm of Xiang et al. [5].

Comparison of GWO results with optimal solutions proves the efficiency of the proposed algorithm. Out of 15 cases, the algorithm has reached the optimal solution in 13 cases and is slightly different from the optimal solution in 2 cases; also, the results of comparing the proposed algorithm with the algorithm of Xiang et al. [5] show the superiority of the designed GWO algorithm. In all cases, both algorithms have the same solutions, or the solution of the algorithm designed in this research is significantly better.

To better show the results of different methods in the definite case, a sample problem with 19 surgeries and 4 operating rooms was solved with different methods and Gantt diagram of the results of the GWO algorithm of Xiang et al. [5], GWO was designed in this study, and the results of the solid mathematical model are depicted in Figures 3 and 4. The numbers shown in each rectangle in these diagrams represent the code for each surgery.

According to Figure 3, in solving this sample problem with Xiang et al.'s [5] algorithm, operating room 3 has two unemployment time blocks between surgeries, which is due to sharing the surgeon's source in surgeries 121 and 122. This means that both surgeries 121 and 122 are assigned to surgeon number 1. On the other hand, as Figure 3 shows, due to the unbalanced distribution of operations in different rooms, overwork occurred in operating room 2. In operating room 3,

however, there is an unused time period. Figure 4 shows the Gantt diagram of the results of the GWO algorithm proposed in this study and the results of the certain optimal model.

Figure 4 shows that although the two proposed GWO methods and the definitive model lead to different answers, they do not differ in the value of the objective function, which shows the efficiency of the proposed GWO algorithm.

After confirming the performance of the designed GWO in the final state, this section evaluates the simulation-based optimization (GWO simulation); for this purpose, 15 real problems have been solved using this algorithm in terms of the probability of the duration of surgery. Table 4 shows the results of solving each problem in 100 iterations of the simulation.

For each problem, the answers obtained (including the sequence of surgeries and the allocation of resources to each stage) are tested on the following scenarios, and the results are presented in Table 5:

Scenario 1: longest surgery scenario; in this scenario, it is assumed that all surgeries are at their maximum time, and the goal is to obtain the objective function in the worst case.

Scenario 2: average value scenario; in this scenario, the time of surgery is equal to the average value.

Scenario 3: the shortest surgical time scenario; in this scenario, it is assumed that all surgeries are in their shortest time. The goal is to obtain the objective function at its best.

Table 5 shows the worst, best, and mediocre results that can be obtained for the answer obtained from the GWO in each of the problems. In order to validate the GWO simulation method, the criterion of using the probabilistic model (VSS) 1 has been used. VSS is the value of considering uncertainty by determining the value of the optimal objective function of the probability problem and the expected value of the objective function for the optimal solution of the problem of the mean value [1, 5, 27–34].

TABLE 4: Results of the proposed hybrid algorithm in the sample of different problems.

Test problem	Number of surgeries	Number of surgeons	Number of operating rooms	Average objective value	CPU time
1	16	9	6	0.18	2
2	23	11	7	3.85	11
3	27	12	7	25.36	37
4	22	10	7	0.00	20
5	16	12	7	0.00	1
6	27	9	7	5.08	27
7	25	9	7	5.81	16
8	18	11	7	2.63	5
9	18	9	7	0.00	1
10	22	11	7	41.04	26
11	21	8	7	0.95	18
12	17	9	7	0.00	1
13	20	16	8	0.00	1
14	22	12	7	0.09	31
15	19	9	8	0.00	17

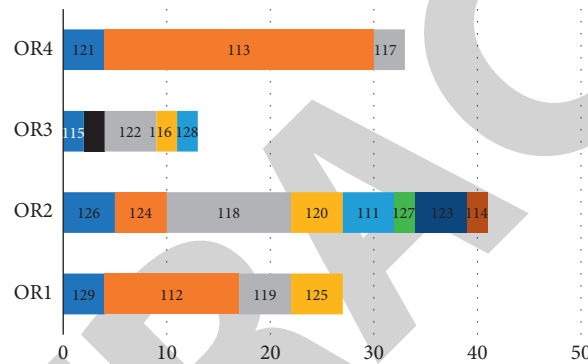


FIGURE 3: Gantt chart of Xiang et al.'s [5] results in solving a sample problem.

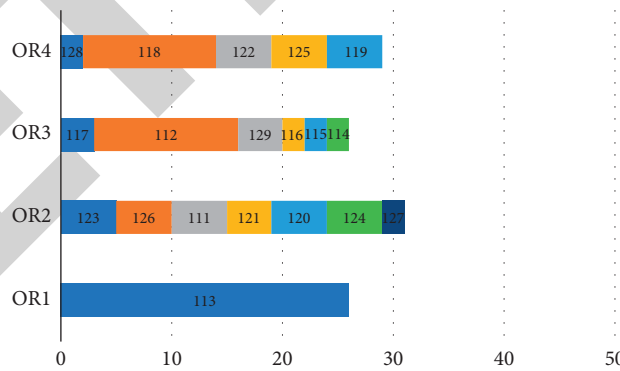


FIGURE 4: Gantt chart of the results of the GWO algorithm proposed in this study in solving the sample problem.

The mean value problem is the definite problem equivalent to the problem considering the average duration of each action. In other words, this value indicates the value of using the probabilistic model instead of the definitive model. To determine the value of the GWO simulation method in comparison with the GWO algorithm, the solution of the combined simulation-based optimization method with the answer obtained from the GWO algorithm for the mean value of the operation time is compared in

conditions of uncertainty. For this purpose, for each sample of the problem, the answer to the problem is first obtained by considering the average duration of surgery with the GWO algorithm. In these samples, the average duration of the operation is considered for each surgery. In this experiment, examples of problems are considered that the GWO algorithm has finally achieved the optimal solution. Each answer includes the sequence of surgeries, and how resources are allocated to them; then, for the same example of problems,

TABLE 5: Evaluation of GWO simulation response on different scenarios.

Number of solutions	Number of surgeries	Worst scenario			Average value scenario			Best scenario		
		Overtime	Idle time	Objective value	Overtime	Idle time	Objective value	Overtime	Idle time	Objective value
1	14	0	4	9.12	0	0	0	0	0	0
2	19	6	0	26.46	0	0	0	0	0	0
3	223	19	5	11.28	2	2	9.38	0	0	0
4	21	19	0	7.29	0	0	0	0	0	0
5	14	0	0	0	0	0	0	0	0	0
6	25	11	2	42.59	3	0	14.75	0	0	0
7	23	9	0	37.25	3	5	16.86	0	0	0
8	16	2	0	5.39	0	0	0	0	0	0
9	19	0	0	0	0	0	0	0	0	0
10	21	42	0	18.93	12	0	47.39	0	0	0
11	19	0	0	0	0	0	0	0	0	0
12	15	0	0	0	0	0	0	0	0	0
13	18	0	0	0	0	0	0	0	0	0
14	24	9	5	26.43	0	0	0	0	0	0
15	17	0	3	5.69	0	0	0	0	0	0

TABLE 6: Comparison of the proposed hybrid method (GWO simulation) and pure GWO algorithm.

Test problem	Number of surgeries	Number of surgeons	Number of operating rooms	Average GWO objective function in scenarios	Mean GWO simulation objective function in scenarios	Percentage of improvement in average responses (VSS)
1	15	6	5	21.452	5.492	73.3
2	20	9	5	6.957	0	100
3	25	10	5	12.79	3.482	74.28
4	20	8	5	3.752	0	100
5	15	10	5	59.73	4.483	95.79
6	25	7	5	9.672	1.572	78.27
7	24	7	5	0.728	0.947	0
8	18	9	5	2.783	0	100
9	16	7	5	1.466	0.792	46.37
10	20	9	4	2.886	0.591	78.42
11	19	6	4	8.394	1.924	76.92
12	15	5	4	1.067	0.219	97.68
Average improvement (percentage)						73.75

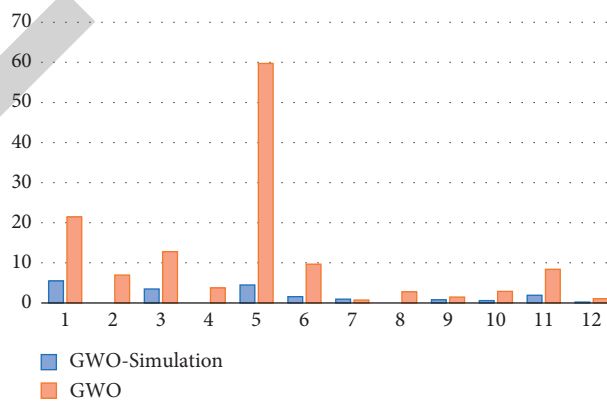


FIGURE 5: Comparison of the GWO with the GWO simulation.

the solution of the problem is obtained by implementing a hybrid optimization algorithm based on simulation and considering uncertain conditions.

In order to compare the results of these two methods, for each problem sample, 50 test scenarios with random numbers were created from the probability function of the

duration of surgery; each scenario shows a possible combination of the duration of the surgery. The solutions of the two-level GWO algorithm (including the sequence of surgeries and resource allocation) and the answers of the GWO algorithm based on simulation (GWO simulation) are related to 12 problems, run with these scenarios, and for each test problem, the average value of the objective function is 50. The scenario is calculated. The results of comparing the mean of the objective function in the scenarios for the answers of the GWO algorithm and the hybrid algorithm are shown in Table 6.

The results presented in Table 6 show that in conditions of uncertainty, the answers obtained by the GWO simulation method are in all cases better than the answers obtained by the GWO algorithm. And only in one case with a slight difference is the GWO answer better. The analysis of the results shows that in conditions of uncertainty, the answers obtained by the GWO simulation method are on average 73.75% better than the answers obtained by the GWO algorithm. The comparison results of the GWO simulation method and GWO algorithm are shown in Figure 5.

Figure 5 clearly shows the superiority of the proposed hybrid method compared to the GWO algorithm. Although GWO simulation has been tested on problems the size of the actual problems of the studied hospital, this algorithm is also capable of solving larger test problems.

6. Conclusion

The field of health has received much attention in recent years, and given the importance of operating rooms and the costs of this part of the hospital, the timing of operating rooms is very important. Therefore, in this article, the issue of operating room scheduling has been studied by considering some existing conditions and limitations.

Considering the operating room separately does not fully address the actual conditions and features of the operating room planning and scheduling issue. Therefore, in this article, several operating rooms and types of surgeries are considered. Also, the operating room and resuscitation sections are considered as two sections that affect each other and are closely related to each other.

Patients who apply for surgery have different priorities according to the type of surgery and degree of sensitivity. The main effect of considering the priority of patients in the context of social services is seen during the time each patient waits for surgery, and patients with more sensitive surgery have a higher priority for surgery. In addition to the operating room time restrictions, the capacity of the resuscitation ward is also limited in terms of the number of beds available and the duration of the patient's stay in the unit.

One of the most important factors in the complexity of patients' planning and scheduling issues is the uncertainty in the duration of surgeries because this factor is influenced by many factors and even in similar surgeries due to certain patient conditions, and the need to use special tools and the duration of surgery are not the same. Also, the presence of a surgeon's assistant in any surgery can improve the outcome

of surgery in cases such as reducing the time of surgery and reducing the possibility of errors and mistakes in surgery.

According to the hypotheses and cases mentioned, a multiobjective mathematical planning model for the operating room scheduling problem is presented, and a case study has been conducted in one of the hospitals in East Asia. The LP-metric method for solving the selected model, as well as the uncertainty in the surgical time parameter, has been investigated, and the model solution results for the problem with different dimensions are presented. The GWO algorithm is also used to solve the model in large dimensions. Due to the uncertainty in the problem under study, this algorithm was combined with the simulation approach, and a new method called GWO simulation was proposed. The results show that with increasing the stability factor in the problem state, the value of the objective function also increases, which indicates that as the model stability increases, the operating room will incur more costs. In other words, the more resistant we are to the answer to uncertainty, the more we have to pay to find the optimal answer. The proposed meta-heuristic algorithm can also be used as a new and efficient tool in operating room scheduling in emergencies.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Esra Sipahi Döngül and Yufen Yang contributed to conceptualization. Yali Liu and Qiqian Li were involved in methodology. Esra Sipahi Döngül provided software. Xiaoyuan Ruan, Yufen Yang, Yali Liu and Qiqian Li were responsible for validation. Esra Sipahi Döngül and Xiaoyuan Ruan carried out formal analysis. Esra Sipahi Döngül took part in data curation. Yufen Yang wrote the original draft. Qiqian Li performed supervision.

References

- [1] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
- [2] D. Min and Y. Yih, "An elective surgery scheduling problem considering patient priority," *Computers & Operations Research*, vol. 37, no. 6, pp. 1091–1099, 2010.
- [3] G. Latorre-Núñez, A. Lüer-Villagra, V. Marianov, C. Obreque, F. Ramis, and L. Neriz, "Scheduling operating rooms with consideration of all resources, post anesthesia beds and emergency surgeries," *Computers & Industrial Engineering*, vol. 97, pp. 248–257, 2016.
- [4] G. Xiao, W. van Jaarsveld, M. Dong, and J. van de Klundert, "Stochastic programming analysis and solutions to schedule overcrowded operating rooms in China," *Computers & Operations Research*, vol. 74, pp. 78–91, 2016.

Retraction

Retracted: An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding

Computational Intelligence and Neuroscience

Received 9 December 2022; Accepted 9 December 2022; Published 9 February 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.

Computational Intelligence and Neuroscience has retracted the article titled “An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding” [1] due to concerns that the peer review process has been compromised.

Following an investigation conducted by the Hindawi Research Integrity team [2], significant concerns were identified with the peer reviewers assigned to this article; the investigation has concluded that the peer review process was compromised. We therefore can no longer trust the peer review process, and the article is being retracted with the agreement of the Chief Editor.

The authors do not agree to the retraction.

References

- [1] A. Alsharif, K. Aggarwal, Sonia, D. Koundal, H. Alyami, and D. Ameyed, “An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding,” *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8467349, 8 pages, 2022.
- [2] L. Ferguson, “Advancing Research Integrity Collaboratively and with Vigour,” 2022, <https://www.hindawi.com/post/advancing-research-integrity-collaboratively-and-vigour/>.

Research Article

An Automated Toxicity Classification on Social Media Using LSTM and Word Embedding

Ahmad Alsharef,¹ Karan Aggarwal,² Sonia ,¹ Deepika Koundal,³ Hashem Alyami,⁴ and Darine Ameyed ⁵

¹Yogananda School of Artificial Intelligence, Computing and Data Science, Shoolini University, Solan, Himachal Pradesh 173229, India

²Electronics and Communication Engineering Department, Maharishi Markandeshwar (Deemed to be University), Mullana, Ambala 133207, India

³Department of Systemics, School of Computer Science, University of Petroleum & Energy Studies, Dehradun, India

⁴Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁵System Engineering Department, Ecole de Technologie Supérieure, University of Quebec, Montreal, Canada

Correspondence should be addressed to Sonia; soniacsit@yahoo.com

Received 19 October 2021; Accepted 17 January 2022; Published 15 February 2022

Academic Editor: Carmen De Maio

Copyright © 2022 Ahmad Alsharef 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 automated identification of toxicity in texts is a crucial area in text analysis since the social media world is replete with unfiltered content that ranges from mildly abusive to downright hateful. Researchers have found an unintended bias and unfairness caused by training datasets, which caused an inaccurate classification of toxic words in context. In this paper, several approaches for locating toxicity in texts are assessed and presented aiming to enhance the overall quality of text classification. General unsupervised methods were used depending on the state-of-art models and external embeddings to improve the accuracy while relieving bias and enhancing F1-score. Suggested approaches used a combination of long short-term memory (LSTM) deep learning model with Glove word embeddings and LSTM with word embeddings generated by the Bidirectional Encoder Representations from Transformers (BERT), respectively. These models were trained and tested on large secondary qualitative data containing a large number of comments classified as toxic or not. Results found that acceptable accuracy of 94% and an F1-score of 0.89 were achieved using LSTM with BERT word embeddings in the binary classification of comments (toxic and nontoxic). A combination of LSTM and BERT performed better than both LSTM unaccompanied and LSTM with Glove word embedding. This paper tries to solve the problem of classifying comments with high accuracy by pertaining models with larger corpora of text (high-quality word embedding) rather than the training data solely.

1. Introduction

With the increased dependence on machine learning (ML) models for different purposes and tasks, researchers recognized the existence of unfairness in machine learning models as one of the most important challenges facing users of ML technologies, as most of these models are trained using human-generated data, which means human bias will emerge clearly in these models. In other words, ML models are biased as the humans who generated the data of training.

Machine learning models' designers must take the initiative in recognizing and relieving these biases; otherwise, the models might propagate unfairness in classification [1]. This unintended bias in the models can also be a result of the demographics of the online users, the underlying or overt biases of those doing the labelling or the selection and sampling [2].

This work aims to improve the classification accuracy of toxicity in online chat forums, but the classification methods presented here can be applied to any other classification

purpose. Toxicity is explained as anything that is insolent, uncivil, or excessive that would make someone want to leave a conversation. Machine learning models will usually learn the simplest associations to predict the corresponding labels of inputs, so any biases or incorrect associations in the training data can propagate unintended biased associations in the classification results. Trained models are known to have the ability to capture contextual dependencies. However, with insufficient data, the models might cause errors and become unable to identify the dependency model and become more probable to generalize, causing the false-positive bias in classification. Toxicity classification models specifically have been shown to capture biases that are common in society from society-generated training data and repeat these biases in classification results, for example, miss-associating frequently attacked identity groups, such as “Black” and “Muslim”, with toxicity in any context even in nontoxic contexts. The following sections will include a description of related works. Furthermore, on proposed models, a technique has been applied by embedding data to relieve the bias. Finally, metrics used for evaluating the classification accuracy in a model will demonstrate that the proposed techniques reduce bias while enhancing overall models’ quality and accuracy.

2. Related Works

Prominent researchers have worked in the area of text analysis. They have analyzed the text and put several security features for its authentication [3, 4]. Authentic data can assist in reducing text toxicity, since not everyone reveals themselves while posting unwanted data.

Many other efforts have been put forward so far to solve the problem of classification in texts [5–11]. Various recent works have studied how concepts of fairness and unintended bias are applied to machine learning models. Researchers have proposed various metrics for the evaluation of fairness in models. Kleinberg et al. [12] and Friedler et al. [13], both groups of researchers, compared different fairness metrics. These works depended on the availability of demographic data to distinguish and relieve bias. Beutel et al. [14] presented a new mitigation technique that used adversarial training techniques and only required a small amount of deceptive labelled demographic data for training. Other works have been conducted on fairness for text classification tasks. Some researchers [15] analyzed different sentiment analysis techniques on the Turkish language with supervised and unsupervised ensemble models to explore the predictive efficiency of the term weighting schemes which is a process to compute and assign a numeric value to each term. The results indicated that supervised term weighting models can outperform unsupervised models in term weighting. Blodgett et al. [16], Hovy et al. [17], and Tatman [18] discussed the impact of using unfair models on real-world tasks but did not provide solutions to adjust this impact. Paryana et al. [19] have suggested intrusion detection techniques to catch such kinds of people. However, directly how it can be applied to the present problem has not been determined. Bolukbasi

et al. [20], in 2016, demonstrated gender bias in word embeddings and provided a solution to counter it using fairer embeddings. Prominent authors [21] proposed an ensemble method for text sentiment analysis and classified it. It aggregates individual features obtained by different methods to obtain a crisp feature subset, and this proposed method outperformed the previous technique. Also, Onan [22] proposed an approach which uses TF_IDF glove embedding technique that gives better results in comparison to the conventional deep learning models in sentiment analysis.

Onan et al. [23] proposed a technique that contains a three-layer bidirectional LSTM network which showed a promising efficiency with a classification accuracy of 95.30%. Also, Onan [24] presented sentiment classification in MOOC reviews. In [25], researchers presented a machine learning-based approach to analyze sentiments with a corpus of 700 student reviews of higher educational institutions written in Turkish, and this machine learning-based approach achieved efficiency in analyzing the sentiments of these reviews.

Georgakopoulos et al. [26] compared convolutional neural networks (CNNs) against the traditional Bag-of-Words for text analysis where the frequency of each word is used as a feature for training combined with algorithms proven to be effective in text classification such as support vector machines (SVMs), Naïve Bayes (NB), K-nearest neighbours (KNNs), and linear discriminant analysis (LDA). They used the same as one of the datasets used in our experiments [27]. A CNN network pretrained with Word2Vec word embedding achieved the highest performance with respect to precision and recall and had the lowest false-positive ratio meaning that this CNNword2vec mistakenly predicted nontoxic comments as toxic the lowest number of times compared to the other models.

In [28], researchers presented an ensemble scheme based on depending on cuckoo search and k-means algorithms. The performance of the proposed model was compared to the conventional classification models and other ensemble models using 11 text benchmarks. The results indicated that the proposed classifier outperforms the conventional classification and ensemble learning model.

This paper adds to this growing effort of research into toxicity classification, an analysis of approaches to relieve bias in text classification tasks achieving high accuracy and F1-score which were the measures of classification as in [29]. Our proposed model used pretrained word embeddings to pertaining classification models instead of training them on the training dataset solely which causes vulnerability to bias.

3. Materials and Methods

This section should contain sufficient detail so that all procedures can be repeated. It may be divided into headed sections if several methods are described.

In this work, several text classifiers were built to identify toxicity in comments from public forums and social media websites. The performance of cache must be good to implement such kind of classifiers as suggested by Sonia et al.

[30]. These classifiers were trained depending on two datasets and tested depending on one dataset.

The first training dataset [31] was of 1.8 million comments, labelled by human raters as toxic and nontoxic. The target column value measures the toxicity rate and determines whether the comment is toxic or not.

The second training dataset [27] was of 223,549 comments labelled in six categories of “toxic,” “severe toxic,” “insult,” “threat,” “obscene,” and “identity hate.”

The testing dataset [32] contained 97,321 entries labelled as approved meaning nontoxic or rejected meaning toxic.

The project focused on the effect of word embeddings on LSTM model binary classification accuracy. Given an input of a comment, it returns whether this comment is toxic or nontoxic. The metrics of measuring the classification accuracy were accuracy score and F1-score. The steps followed in the experimental work are illustrated in Figure 1.

The models applied in this work are illustrated in Table 1.

3.1. Analysis of the 1st Training Dataset. The first training dataset [31] was published by the Jigsaw unit of Google [33] throughout the competition of “Jigsaw Unintended Bias in Toxicity Classification” on the Kaggle community. Each comment in this dataset had a toxicity label (target). This attribute is a fractional value that represents the judgment of human raters who estimated how much toxicity is contained in a given comment. For classification accuracy evaluation, test set examples with (target ≥ 0.5) were considered as toxic, while other comments having target < 0.5 were considered as nontoxic. Table 2 is a tiny sample of these comments and their corresponding “target” value.

From Table 2, we observe that the first two comments are not toxic having target < 0.5 , whereas the third comment is toxic having target > 0.5 .

Terms affected by the false-positive bias usually occur in comments and are usually misclassified by NLP models as toxic even in nontoxic comments especially that the training data of models is usually human generated. The disproportionate number of toxic examples containing these terms in the training dataset can lead to overfitting in the classification model. For example, in this dataset, the word “gay” appears in only 3% of toxic comments and only 0.5% of the overall comments. Biased models can make overfitting such as always linking the word “gay” with toxicity which is not always correct, and it can come in a nontoxic context.

Visualization of data is reported in the next paragraphs.

We can see a relation between the target and certain categories of toxic words. The scatter charts illustrated in Figure 2 show the relationship between some of these categories and toxicity (target value).

The occurrence of comments holding these categories such as insult and identity attack increases its potential to be classified as toxic in the training dataset.

On the contrary, some words occurrence does not usually lead to toxicity. This is concluded from the scatter charts illustrated in Figure 3 which show the relation between some categories of comments and toxicity.

Collect	Collecting Data.
Describe	Describing Data.
Preprocess	Preprocessing Data.
Build	Building Toxicity Classification Models.
Train	Training the Models with Training Data.
Test	Testing the Models with Testing Data.
Observe	Observing Evaluation Metrics.
Discuss	Discussing and Analyzing Results.

FIGURE 1: Experimental workflow.

TABLE 1: Classification models of this work.

Experiment 1		Experiment 2	
Neural network	Word embedding	Neural network	Word embedding
LSTM	Glove	LSTM	BERT

The occurrence of comments holding these words, such as black and Buddhist, does not usually increase its potential to be classified as toxic in the training dataset.

3.2. Analysis of the 2nd Training Dataset. The second training dataset [27] used in this work included 223,549 published by the Jigsaw unit of Google [33] throughout the “Toxic Comment Classification Challenge” on Kaggle. These user comments were labelled by human labellers within six labels: “toxic,” “severe toxic,” “insult,” “threat,” “obscene,” and “identity hate.” Some comments could be categorized into different labels at once. The dataset labels distribution is shown in Table 3.

Two lakh one thousand and eighty one comments were classified under the “clean” category matching none of the six categories constituting 89.9% of overall comments, whereas the other comments belonged to at least one of the other classes constituting 10.1% of overall comments. The comments collected were mostly written in English with some outliers of comments from different languages, e.g., in Arabic and Chinese. The comment was considered as “toxic” if it was classified under any of the six categories and as “nontoxic” otherwise (not categorized under any of the six categories).

3.3. Training Data Preprocessing. The text data preprocessing techniques followed before processing and modeling the data are as follows.

Punctuation removal: removing punctuation is a necessary step in cleaning the text data before performing

TABLE 2: Training dataset 1 sample.

Comment_text	Target
This is so cool. It's like, "would you want your mother to read this?"	0
Thank you!! This would make my life a lot less anxiety-inducing.	0
Haha, you guys are a bunch of losers.	0.8936

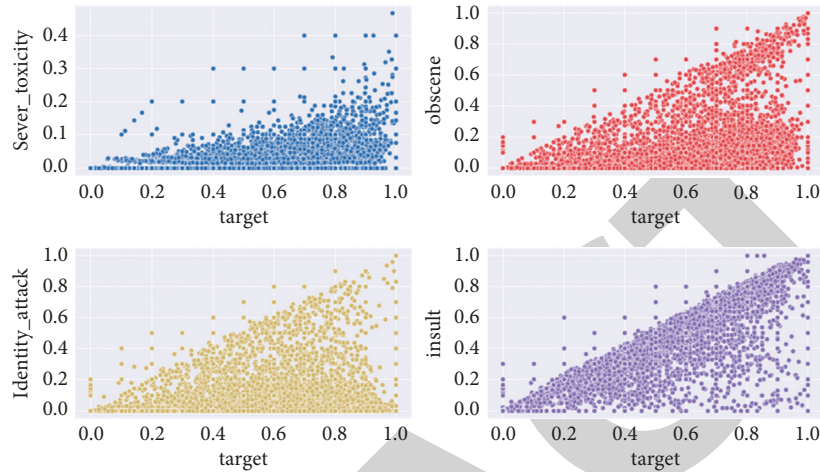
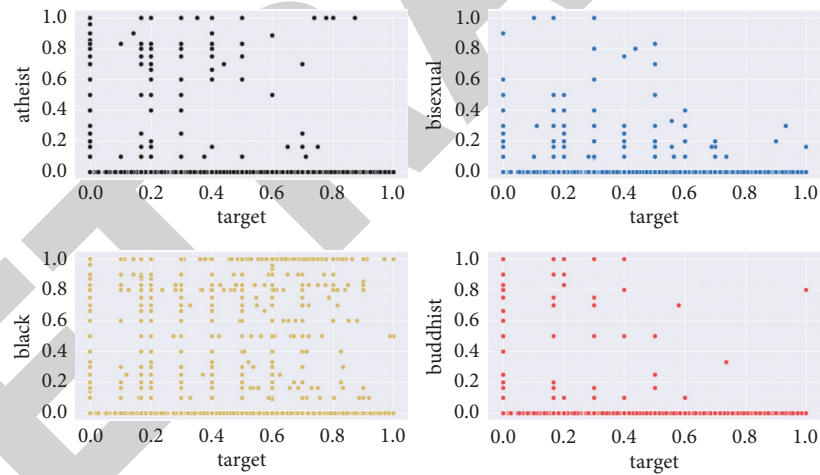
FIGURE 2: The relation between some features of comments and toxicity in the 1st training dataset.

FIGURE 3: The relation between some features of comments and toxicity.

TABLE 3: Label distribution of the 2nd training dataset.

Class	No. of occurrences
Clean	201,081
Toxic	21,384
Obscene	12,140
Insult	11,304
Identity hate	2,117
Severe toxic	1,962
Threat	689

analytics. In this work, all punctuation marks in all comments were removed.

Lemmatization: lemmatization is the process of grouping together the inflected forms of a word so they can be

analyzed as a single term. In this work, lemmatization was performed for every comment.

Stop words' removal: stop words are words that do not contain any significance in a context. Usually, these words are filtered out from text blocks because they have unnecessary information such as the, be, are, and a.

3.4. Testing Dataset. The test dataset used for evaluation in this work was downloaded from the Kaggle competition of "Jigsaw Unintended Bias in Toxicity Classification" [32]. It contained 97321 entries labelled as approved (nontoxic) or rejected (toxic). A sample of the testing dataset is given in Table 4:

TABLE 4: Testing dataset sample.

Comment_text	Rating
Sorry, you missed high school. Eisenhower sent troops to Vietnam after the French withdrew in 1954	Approved
Our oils read; President IS taking different tactics to deal with a corrupt malignant, hypocritical . . .	Rejected
Why would 90% of articles print fake news to discredit Trump? Where are you getting your new” . . .	Approved

3.5. *LSTM Model*. Initially, LSTM [34, 35] was created where the information flows through cell states. In this way, LSTMs can selectively remember or forget information. This study worked on using LSTM and word embeddings for toxicity classification. The design of the LSTM neural networks used in this work is shown in Figure 4.

The designed fine-tuned LSTM of this work takes a sequence of words as an input.

A word embeddings’ layer that provides a representation of words and their relative meanings was added. This embedding layer transforms encoded words into a vector representation.

Then, a spatial dropout layer that masks 10% of the word embeddings’ layer output makes the neural network more robust and less vulnerable for overfitting.

Then, to process the resulted sequence, an LSTM layer with 128 units was used as well as another 10% dropout layer.

After all, a dense output layer was used to output the multilabel classification.

3.6. *Word Embedding*. Word embedding is a concept used for representing words for text analysis, generally in a form of a vector of real values that encodes the meaning of the word in such a way where the words that are closer in the vector space are expected to have related meanings [36]. Word embeddings can be obtained using different techniques where words from the vocabulary are mapped to vectors of real numbers. Each word is mapped to one vector. Figure 5 illustrates the different types of word embeddings.

In this work, Glove static (context-independent) word embeddings and a contextualized word embeddings generated by BERT were used for pretraining the classification models before training them on the training datasets. The word embedding this work used is as follows.

Glove: it is a learning algorithm for calculating vector representations of words regardless of sentence context. Training in glove is performed on aggregated global word occurrence statistics from a large corpus [37, 38]. The Glove word embeddings this work used to pretrain the models are as follows:

Wikipedia 2014: 400 thousand word vectors trained on a largeWikipedia-2014 corpus [39].

Gigaword 5: Gigaword Fifth Edition archive of newswire text [40].

Twitter: 1.2 million word vectors trained on large Twitter corpora [4].

BERT: it is an encoder was proposed in a paper published by Google AI in 2018 [34, 41]. Its main innovation is to apply bidirectional training to the transformer, which is a

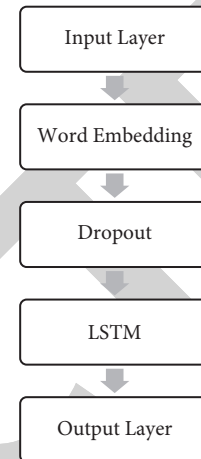


FIGURE 4: LSTM model layers’ design.

well-known attention model in language modeling. Results predict that a bidirectionally trained language model can sense more deeply in context of language in comparison to the single directional language model. Bidirectional LSTM can also be trained on both sides that are left to right for detecting the next word of sentence and vice versa to find out the previous word. That means this will use both forward and backward LSTMs. However, none of the techniques considered both ways simultaneously like taken in BERT [19]. BERT also can generate various context-dependent word embeddings of a word dynamically informed by words around it [42].

4. Results and Discussion

The evaluation metrics used to evaluate the efficiency of models were accuracy and F1-score. The following paragraph will describe these metrics:

- (i) Accuracy describes the accuracy achieved on the testing set. The formula for accuracy is

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}. \quad (1)$$

- (ii) Precision is defined as the ratio of correctly predicted positive observations to the total predicted positive observations. The formula for precision is

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

- (iii) Recall is defined as the proportion of correctly identified positives. The formula for recall is

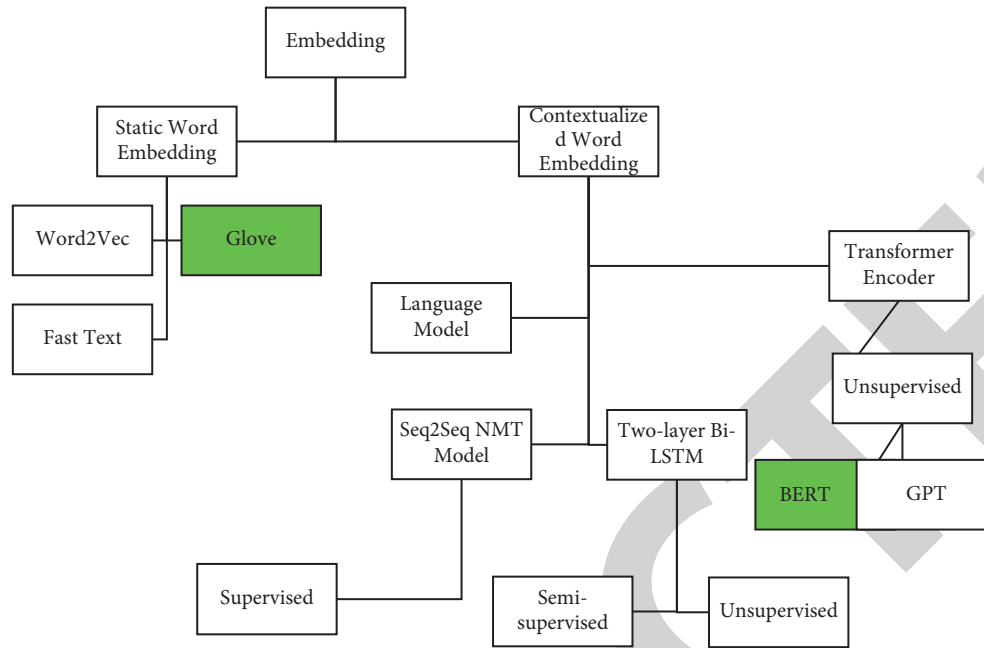


FIGURE 5: Word embedding types.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

(iv) F1-score is the harmonic mean of precision and recall. The formula for F1-score is

$$\begin{aligned} F1 &= \frac{2}{\text{precision}^{-1} + \text{recall}^{-1}} \\ &= 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \\ &= \frac{2TP}{2TP + FP + FN} \end{aligned} \quad (4)$$

The experiments applied the LSTM model by pertaining it with different word embeddings each time. The LSTM model itself is known for its memory that can keep long sequences of words and its suitability for word classification. After adding the Glove word-embedding layer and applying the LSTM model, we obtained a high accuracy of 93% and a high F1-score of 0.84 on the previously mentioned training and testing datasets. However, in LSTM, according to Singh [19], the language models built on word embeddings do not accurately capture the nuances and meanings of the sentences. This made the added word embeddings not highly effective for language modeling. Using bidirectional word embeddings solved the problem where combining LSTM with BERT and applying the same settings as in the previous model gave a higher classification accuracy of 94% and a higher F1-score of 0.89, in classifying toxic comments, on the previously mentioned training and testing datasets. The summary of the results are represented in Table 5.

From the results, we could find that using word embeddings could improve the efficiency of classification.

TABLE 5: Accuracy and F1-score of LSTM with different word embeddings in classifying toxic words.

	Accuracy (%)	F1-score
LSTM + Glove	93	0.841
LSTM + BERT	94	0.894

Words embedding generated by the BERT model was proved to be more efficient than static Glove word embeddings when used with LSTM since it trains in both directions allowing higher efficiency, and because BERT analyzes every sentence with no specific direction, it does a better job at understanding the meaning of homonyms than previous NLP methodologies, such as Glove embedding methods.

Word embeddings trained on a large corpus such as Glove trained on Wikipedia, Gigword, and Twitter were also found effective to enhance the accuracy of classification but less effective than BERT (in classifying toxicity in text documents).

5. Conclusions

Many former research works have recognized unfairness in ML models for toxicity classification causing inaccurate classification as a concern to relieve. This can be observed obviously in toxicity classification in public talk pages and online discussion forums. In this paper, various machine learning and natural language processing models for toxicity classification were proposed, implemented, and illustrated. It was found that many errors in toxicity identification occur due to the lack of consistent quality of data. By adding word embeddings, the accuracy of classification increased notably. Finally, an accuracy of 94% and an F1-score of 0.89 were achieved using a hybrid BERT and LSTM classification

model. This work can be further extended by exploring the potential of subword embeddings [43] which can further enhance the accuracy of classification. A more robust model can be developed by applying AutoNLP and AutoML techniques on the same datasets where in order to obtain better results and accurate classifications these techniques automatically find the models that fit data the best.

Data Availability

The data presented in this study are openly available in Kaggle competition of “Jigsaw Unintended Bias in Toxicity Classification.”

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

The project has been funded by Taif University, Kingdom of Saudi Arabia, under grant no. TURSP-2020/306.

References

- [1] B. Van Aken, J. Risch, A. Krestel, and A. Löser, “Challenges for toxic comment classification: An in-depth error analysis,” 2018, <https://arxiv.org/abs/1809.07572>.
- [2] D. Borkan, L. Dixon, J. Sorensen, N. Thain, and L. Vasserman, “Nuanced metrics for measuring unintended bias with real data for text classification,” in *Proceedings of the Companion Proceedings of the 2019 World Wide Web Conference*, pp. 491–500, San Francisco, CA, USA, May 2019.
- [3] F. Ahmadi, Sonia, G. Gupta, S. R. Zahra, P. Baglat, and P. Thakur, “Multi-factor biometric authentication approach for fog computing to ensure security perspective,” in *Proceedings of the 2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 172–176, IEEE, New Delhi, India, March 2021.
- [4] Data.world, “Twitter,” <https://data.world/marcusyyy/twitter>.
- [5] A. Onan, “Topic-enriched word embeddings for sarcasm identification,” *Advances in Intelligent Systems and Computing*, Springer, New York, NY, USA, pp. 293–304, 2019.
- [6] H. Bulut, S. Korukoğlu, and A. Onan, “Ensemble of keyword extraction methods and classifiers in text classification,” *Expert Systems with Applications*, vol. 57, pp. 232–247, 2016.
- [7] A. Onan, “An ensemble scheme based on language function analysis and feature engineering for text genre classification,” *Journal of Information Science*, vol. 44, no. 1, pp. 28–47, 2018.
- [8] M. A. Toçoğlu and A. Onan, “Satire identification in Turkish news articles based on an ensemble of classifiers,” *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 28, no. 2, pp. 1086–1106, 2020.
- [9] H. Bulut, S. Korukoğlu, and A. Onan, “A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification,” *Information Processing & Management*, vol. 53, no. 4, pp. 814–833, 2017.
- [10] S. Korukoğlu and A. Onan, “Exploring the performance of instance selection methods in text sentiment classification,” in *Artificial Intelligence Perspectives in Intelligent Systems*, pp. 167–179, Springer, Cham, New York, NY, USA, 2016.
- [11] A. Onan, “A fuzzy-rough nearest neighbor classifier combined with consistency-based subset evaluation and instance selection for automated diagnosis of breast cancer,” *Expert Systems with Applications*, vol. 42, no. 20, pp. 6844–6852, 2015.
- [12] J. Kleinberg, S. Mullainathan, and M. Raghavan, “Inherent trade-offs in the fair determination of risk scores,” 2016, <https://arxiv.org/abs/1609.05807>.
- [13] A. Friedler, J. Scheidegger, and V. S. Cii, “On the (im) possibility of fairness,” 2016, <https://arxiv.org/abs/1609.07236>.
- [14] A. Beutel, Z. Z. Chen, and E. H. Chi, “Data decisions and theoretical implications when adversarially learning fair representations,” 2017, <https://arxiv.org/abs/1707.00075>.
- [15] A. Onan, “Ensemble of classifiers and term weighting schemes for sentiment analysis in Turkish,” *Scientific Research Communications*, vol. 1, no. 1, pp. 1–12, 2021.
- [16] S. L. Blodgett and B. O’Connor, “Racial disparity in natural language processing: a case study of social media african-american English,” 2017, <https://arxiv.org/abs/1707.00061>.
- [17] D. Hovy and Spruit C151, “The social impact of natural language processing,” in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, vol. 2, pp. 591–598, Berlin, Germany, August 2016.
- [18] R. Tatman, “Gender and dialect bias in YouTube’s automatic captions,” in *Proceedings of the First ACL Workshop on Ethics in Natural Language Processing*, pp. 53–59, 2017.
- [19] A. Singh, “Building state-of-the-art language models with BERT, medium,” 2019, <https://medium.com/saarthi-ai/bert-how-to-build-state-of-the-art-language-models-59ddd9ac5d>.
- [20] T. Bolukbasi, K. W. Chang, J. Y. Zou, V. Saligrama, and A. T. Kalai, “Man is to computer programmer as woman is to homemaker? debiasing word embeddings,” *Advances in Neural Information Processing Systems*, vol. 29, pp. 4349–4357, 2016.
- [21] S. Korukoğlu, “A feature selection model based on genetic rank aggregation for text sentiment classification,” *Journal of Information Science*, vol. 43, no. 1, pp. 25–38, 2016.
- [22] A. Onan, “Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks,” *Concurrency and Computation: Practice and Experience*, vol. 35, 2020.
- [23] M. Tocoglu and A. Onan, “A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification,” *IEEE Access*, vol. 9, pp. 7701–7722, 2021.
- [24] 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, 2020.
- [25] M. A. Toçoğlu and A. Onan, “Sentiment analysis on students’ evaluation of higher educational institutions,” in *Proceedings of the International Conference on Intelligent and Fuzzy Systems*, pp. 1693–1700, Springer Cham, Istanbul, Turkey, 2020 July.
- [26] S. V. Georgakopoulos, S. K. Tasoulis, and G. Vrahatis, “Convolutional neural networks for toxic comment classification,” in *Proceedings of the 10th hellenic conference on artificial intelligence*, pp. 1–6, Patras, Greece, July 2018.
- [27] A. I. Jigsaw/Conversation, “Toxic comment classification challenge | Kaggle (Train.csv),” Kaggle.com,” 2018, <https://www.kaggle.com/c/jigsaw-toxic-comment-classification-challenge/overview>.

Retraction

Retracted: Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] D. Yadav, N. Lalit, R. Kaushik et al., "Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 3411881, 14 pages, 2022.

Research Article

Qualitative Analysis of Text Summarization Techniques and Its Applications in Health Domain

Divakar Yadav ¹, Naman Lalit ¹, Riya Kaushik ¹, Yogendra Singh ¹, Mohit ¹, Dinesh ¹,
Arun Kr. Yadav ¹, Kishor V. Bhadane ², Adarsh Kumar ³, and Baseem Khan ⁴

¹Department of Computer Science and Engineering, NIT Hamirpur (HP), Hamirpur, India

²Amrutvahini College of Engineering Sangamner, Ghulewadi, Maharashtra, India

³Department of Systemics, School of Computer Sciences, UPES, Dehradun, India

⁴Department of Electrical and Computer Engineering, Institute of Technology, Hawassa University, Hawassa, Ethiopia

Correspondence should be addressed to Baseem Khan; baseem.khan04@gmail.com

Received 21 October 2021; Accepted 20 January 2022; Published 9 February 2022

Academic Editor: Lerina Aversano

Copyright © 2022 Divakar Yadav 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.

For the better utilization of the enormous amount of data available to us on the Internet and in different archives, summarization is a valuable method. Manual summarization by experts is an almost impossible and time-consuming activity. People could not access, read, or use such a big pile of information for their needs. Therefore, summary generation is essential and beneficial in the current scenario. This paper presents an efficient qualitative analysis of the different algorithms used for text summarization. We implemented five different algorithms, namely, term frequency-inverse document frequency (TF-IDF), LexRank, TextRank, BertSum, and PEGASUS, for a summary generation. These algorithms are chosen based on various factors. After reviewing the state-of-the-art literature, it generates good summaries results. The performance of these algorithms is compared on two different datasets, i.e., Reddit-TIFU and MultiNews, and their results are measured using Recall-Oriented Understudy for Gisting Evaluation (ROUGE) measure to perform analysis to decide the best algorithm among these and generate the summary. After performing a qualitative analysis of the above algorithms, we observe that for both the datasets, i.e., Reddit-TIFU and MultiNews, PEGASUS had the best average F-score for abstractive text summarization and TextRank algorithms for extractive text summarization, with a better average F-score.

1. Introduction

Summarizing textual information requires understanding and analyzing the linguistic, conceptual, and semantic attributes of the given information. In addition, a summary generated should succeed in incorporating the essential details and the main ideas of the given text. Extractive summarization techniques can extensively analyze the given text semantically, i.e., on sentences, words, keywords, etc., identified by the algorithm. Extractive summarization techniques [1, 2] are also computationally more feasible to implement since they require fewer resources, computation power, and time to assess and generate a summary since they are statistically oriented. However, the techniques generate a summary by identifying the imperative sentences. The

keywords are identified in a given text based on the frequency of their occurrences. The technique might not efficiently incorporate the information given or might leave out some crucial details [3].

On the other hand, abstractive summarization techniques [4] analyze the data using a natural language processing approach and generate a summary by reforming the given information concisely around the vital idea of the information. A summary generated by abstractive summarization methods is more comparable to a human-generated summary, a criterion that a summary generated by extractive summarization techniques (EST) might not satisfy. Abstractive summarization methods require efficient implementation of various machine learning techniques with large datasets with good variety and conditional

aspects. Since abstractive summarization techniques require the implementation of machine learning algorithms, it is computationally expensive and requires time to be implemented efficiently. The cost of implementation grows exponentially with the size of data being summarized.

Abstractive techniques can be understood as the way humans analyze any textual document. It selects words that

Information record \rightarrow get setting \rightarrow semantics \rightarrow make own rundown. (1)

Extractive summarization techniques focus on summarizing a textual document by selecting words or sentences that are important to the context or appear more frequently [5]. The extractive summarization techniques score or assign loads to words or sentences and use pieces of significance or

are semantically appropriate for the content. The summary generated might include words that were not even included in the given data since abstractive summarization deciphers and examines the content using natural language processing techniques and creates concise data that constitutes the most basic idea and key contents of the textual data given for summarization.

equivalent for a summary generation. Various methods and mathematical calculations are used to assign loads or scores for the words/sentences, which are further used to position the sentences/words according to their significance and comparability [6].

Information record \rightarrow sentences closeness \rightarrow score sentences \rightarrow selection of sentences with higher significance. (2)

The abstractive strategies require a proficient understanding of the textual data as compared to the extractive strategies. The simplistic statistical and mathematical approach of extractive strategies is often more efficient and successful at summarization than the complex and sophisticated approach of abstractive summarization techniques (AST), which considers several factors like inference and attributes, semantic presentation, language, etc., which are more complex than statistic driven ideologies, for example, sentence/word extraction. We have used the ROUGE metric to evaluate and compare the performance of different methods and techniques in this work.

The following are the main contributions in this article:

- (i) Five different algorithms for text summarization: TF-IDF, LexRank, TextRank, BertSum, and PEGASUS have been implemented on two different datasets: Reddit-TIFU and MultiNews
- (ii) An exhaustive, detailed qualitative analysis is performed to evaluate the algorithms on three ROUGE parameters, i.e., Rough-1, Rough-2, and Rough-L, and finally, F-score is computed and found promising results for EST and AST, respectively

The work in this article is arranged in sections as follows. The next section discusses the related works for ESTs and ASTs. Section 3 discusses the methodology. Further, Section 4 discusses the datasets and implementation. The result analysis is discussed in Section 5, followed by a conclusion and references.

2. Related Works

Various researches have been done to analyze different summarization algorithms, and hence, several research articles for the purpose mentioned above have been published.

We aimed to gather optimal knowledge from research on summarization techniques [7] and efficiently implement and optimize our models for assessing its performance and concluding with concrete results. We learned various summarization techniques for single and multidocuments [8]. We read about some of the most widely used methods such as frequency-driven methods, topic representation approaches, and graph-based and machine learning techniques [9] through this paper.

A thorough study provided insight into recent trends and advancements in automatic summarization techniques [10] that describe the state of the art in this research area. Generally, there are two types of summarization techniques. Here is some previous research work in the following fields.

2.1. Extractive Summarization. Extractive summarization, at the most basic level, can be approached by using the sentence scoring technique that obtains the text's keyword [11]. It is done by analyzing and filtering the words which are used most frequently in the text. The sentences with a high frequency of these words are used for generating a summary of the original text by using the sentences with high scores in decreasing order of scores [5]. For better performance and efficiency, graph-based methods were introduced, making the models capable of considering more complex attributes of the textual information and presenting concise information with better accuracy.

In graph-based approaches, the words are considered nodes, and their relation with other words is based on their frequency, which is depicted as edges. The edges are weighted and are analyzed for choosing the query words for generating a summary [12]. Several algorithms like PageRank, TextRank, TexRank, etc., can be used for efficient text summarization techniques [13]. A bipartite graph is created to represent sentences and topics separately. Scores are

assigned to each sentence, and sentences in decreasing scores are added to the summary. Several techniques like Levenshtein distance, semantic similarity, and cosine similarity are used for determining the relation between sentences and words, which then pave the way for an efficient summary generation [14].

We have implemented, executed, and assessed four different extractive summarization techniques [15] in this work, namely, TF-IDF (term frequency-inverse document frequency) summarization algorithm, LexRank algorithm, TextRank algorithm, BertSum algorithm, and PEGASUS algorithm. The task required us to comprehend the fundamentals and complexities of each algorithm. Below are brief explanations about these algorithms.

In the TF-IDF algorithm, large texts are converted into sentences and then weighted term frequency, and inverse sentence frequency is calculated where the sentence frequency is defined as the number of sentences of the document, which involve these terms [16]. The vectors of the sentences are calculated and compared with the other sentences and are then scored. The product of TF and IDF calculates the TF-IDF value of a word/term, where TF (term frequency) is defined as the number of times a word occurs in a document and IDF is inverse document frequency [8]. The sentences with the highest score are considered the conclusive sentences for summary [17]. This paper provides more detailed information about the application of the TF-IDF algorithm on multidocument extractive text summarization.

LexRank algorithm is an unsupervised graph-based method for automatic text summarization (ATS) [18]. Graph method is used to compute the score of sentences. LexRank is used for computing sentence importance based on the concept of eigenvector centrality in a graph representation of sentences. In this algorithm, we have a connectivity matrix based on intrasentence cosine similarity, used as the adjacency matrix of the graph representation of sentences [19]. This sentence extraction majorly revolves around the set of sentences with the same intent; i.e., a centroid sentence is selected, which works as the mean for all other sentences in the document. Then, the sentences are ranked according to their similarities.

TextRank algorithm, for automatic text summarization, is an unsupervised graph-based ranking approach. The scoring of sentences is performed using the graph method, where each vertex is scored based on the linking of those tokens, which are considered vertex in the graph [20]. TextRank can be used for keyword extraction and sentence extraction. Here, we have used TextRank as sentence extraction with a higher score. An important aspect of TextRank is that it does not require deep linguistic knowledge, nor domain or language-specific annotated corpora, which makes it highly portable to other domains, genres, or language.

BertSum algorithm assigns scores to each sentence that represents how much value that sentence adds to the overall document [21]. Scores of each node or vertex are decided by either a “voting” or “recommendation” system, where each node or vertex votes for all others. The importance of a node/

vertex is decided based on the votes received. The value of each vote also depends on the importance of the node casting it. The sentences with the highest scores are then collected and rearranged to give the overall summary of the article.

A quantitative and qualitative assessment of 15 algorithms has been performed by Ferreira et al. [22] for sentence scoring. They evaluated these algorithms in three datasets: convolutional neural network (CNN) news dataset, Blog summarization dataset, and SUMMAC dataset. In the paper [23], the authors proposed extractive text summarization of Hindi novels and stories. They create a good corpus of the dataset of Hindi novel and perform summarization with standard evaluation parameters. Also, they evaluate the proposed model on slandered English dataset and concluded that proposed model outperforms as compared to state of art methods.

An extractive multidocument text summarization using a quantum-inspired genetic algorithm is proposed in the paper [24]. They proposed a quantum-inspiring genetic algorithm to summarize silent sentences of web-based multidocuments. The proposed model is evaluated on standard benchmark datasets DUC 2005 and DUC 2007. They concluded that the proposed model outperforms as compared to the state-of-the-art methods. Kumar et al. [25] presented an improvised extractive approach based on a thematic approach for summarization of Hindi text documents.

In the paper [26], the authors proposed a new dataset “SIGIR2018” for extractive text summarization. They evaluated the dataset on standard matrices and compare the results of other publicly available datasets like DUC 2005 and DUC 2007.

2.2. Abstractive Summarization. On the other hand, abstractive summarization does not focus on the semantic representation of data and utilizes techniques of natural language processing (NLP) and linguistic approach to concise the given information [14]. Summaries generated by abstractive summarization might not be composed of original sentences or words and might have been replaced by morphed sentences and new words. Summaries generated by abstractive summarization are more comparable to human-generated summaries [27]. They succeed in better comprehension of the context and idea of the information; however, since the algorithms require training of models and implementation of NLP models, they require high computational power and more resources than extractive summarization techniques.

PEGASUS algorithm is an abstractive summarization algorithm, which uses a sequence-to-sequence framework using encoder-decoder architectures based on recurrent neural network (RNNs) [28]. It uses pretrained sequence-to-sequence models with sentences masked and then passed to the encoder-decoder [29]. This paper gives more detailed info about sequence-to-sequence models. It is computationally expensive and needs a lot of time and resources for implementation. The masked information can be sentences,

words or collocations, etc. A study into this domain gave insights into the abstractive text summarization algorithm, which can generate a summary of texts based on the concept of extracted gap sentences. Pretraining them on different models leads to more accurate results, as the model can predict the missing sentences and then is used for the summarization of lengthy text.

An abstractive text summarization using a hierarchical human-like deep neural network is proposed in the paper [30]. The authors' main objective is to generate abstractive text summarization as much similar to a human-generated summary. The proposed model is based on a knowledge-aware hierarchical attention module, a multitask learning module, and a dual discriminator generative adversarial network. They compare the results on a standard dataset with standard evaluation matrices.

These were some of the algorithms whose literature work is mentioned above, and now in the next section, we will be explaining the algorithms in more detail, their implementation, and their detailed analysis by comparing their results from the various datasets used for text summarization.

3. Methodology

Summarization processes in extractive and abstractive-based algorithms can be tackled by focusing on semantic attributes and semantic relationships among the constituents of the given information. These relationships can be established by considering various aspects, e.g., by using different algorithms like K-means clustering, using scoring systems for words and sentences, using voting systems among words and sentences, through machine learning.

As it is known that there are many algorithms available for text summarization, each one of them has its characteristics and performs better on different datasets. Mainly, all the algorithms are classified into various categories based on their implementation.

The extractive-based algorithms are classified into three types mainly based on the different types of learning as shown in Figure 1 [31]. It provides detailed information about all the techniques used for selecting the best extractive algorithm based on various attributes.

Regarding abstractive-based text summarization, the algorithms are categorized into two main types of summarization algorithms based on approaches, i.e., semantic-based approach and structured-based approach, as shown in Figure 2. This article provides a deep understanding of these approaches and helps to identify algorithms suitable for text summarization.

All the algorithms focus on determining meaningful sentences, keywords, and words for generating the summary, which concisely conveys vital information. After gathering information about the performance of different algorithms, we have selected five algorithms, which perform better than other algorithms and are extensive in delivering better results. In particular, PEGASUS (abstractive-based algorithm) shows the state-of-the-art performance among all the other abstractive-based algorithms [28].

As discussed earlier in Section 1, the algorithms used in this paper are TF-IDF, LexRank, TextRank, BertSum, and PEGASUS that have been developed around the concepts mentioned earlier. TF-IDF, LexRank, and TextRank algorithms work by calculating word or sentence scores through their system, whereas BertSum and PEGASUS use a voting system among words or sentences, whichever is more optimal, and use machine learning and RNN techniques, respectively. The algorithms have been discussed thoroughly, along with their results and conclusions in the following sections.

3.1. Term Frequency Algorithm. Large messages are first changed over into sentences, and afterward-weighted term frequency and inverse document frequency are determined where the sentence recurrence is characterized as the number of times these terms have appeared in the sentences of the archive [8]. The vectors of the sentences are determined, contrasted, and different sentences and are then scored.

The TF-IDF estimation of a word is determined by the result of TF (term frequency) and IDF (inverse document frequency), where TF (term frequency) is defined as the occasions a term happens in a record [32]. The sentences with high weight values are selected to be the definitive sentences for synopsis. In this technique, each word is given a value between 0 and 1, where the closer the value is to 1, the higher will be its priority. Moreover, each word is known as a term, and it helps in outlining the important terms in the document, thereby generating a better summary.

In contrast to different calculations requiring man-made consciousness and AI, this programmed rundown exploration need not bother with any AI because of the utilization of libraries currently available to us, for example, NLTK and BeautifulSoup. Utilization of the current libraries helps us focus on the most proficient method to ascertain TF-IDF and the content. The program is isolated into three primary capacities, which are preprocessing, highlight extraction, and synopsis.

We have composed the calculation in Python to produce the rundown utilizing this calculation. Figure 3 shows the flow chart of TF-IDF technique implementation.

Preprocessing capacity measures the archive with NLTK capacities like grammatical feature (POS) tagger, tokenization, stemming, and stop words [33]. After the archive is inputted into the program, the preprocessing capacity parts the content into a rundown of terms utilizing tokenization capacities. The emotional development of the Internet has led to the overpowering of individuals by the enormous measure of online textual data and reports [34]. This growing accessibility of records has demanded thorough exploration in the programmed text summary or outline. An outline is defined as "a book that is created from at least one message, that conveys significant data in the given text (s), and that is shorter or equivalent to half of the given text(s) and normally, altogether not as much as that." For instance, web crawlers produce scraps as the analysis of the given text. Different models incorporate news sites, which produce consolidated portrayals of information themes, usually as

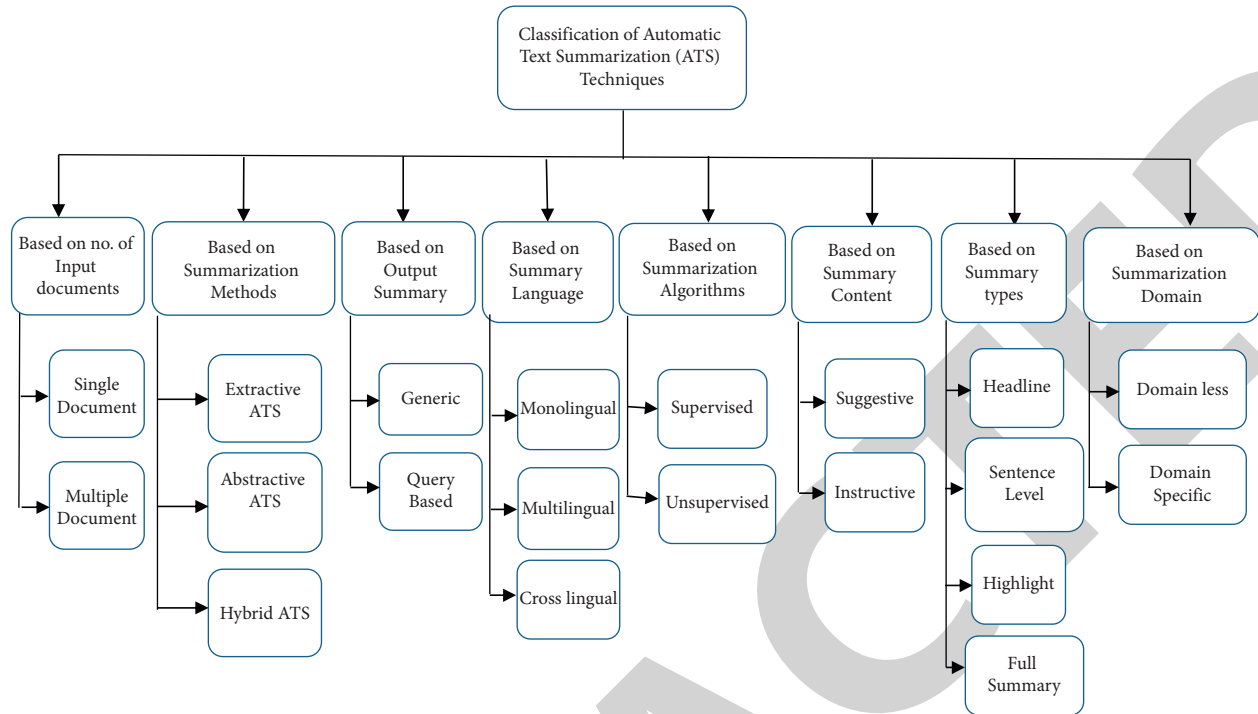


FIGURE 1: Classifications of automatic text summarization methods.

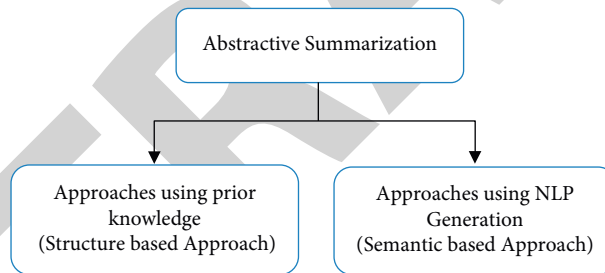


FIGURE 2: Abstractive text summarization techniques.

features to encourage examining or information extractive ideology or techniques.

We as humans summarize any given data by first reading it from top to bottom to comprehend the context and then composing our summary by featuring the main idea or concerns. Since machines cannot read or understand like humans, it has made programmed text synopsis extremely troublesome. Programmed text rundowns have been an area of interest since the 1950s. A significant amount of attention to this field was due to the summarization of logical archives. Luhn [35] set the foundation stone for programmed summarization by proposing the summarization technique by considering sentences from content utilization highlights—for example, term and sentence recurrence. The technique required assigning weight to the sentences of the given text to determine words with high recur frequency and ignoring the common words with high recurring frequency.

From the start, we standardize the reports, and the content is changed over into lowercase, so the two words, for example, Hello and hi, are not viewed as of particular. At that

point, the cycle of tokenization happens where the sections are changed over into singular sentences. After this, the sentences are further tokenized and changed over into a rundown of words. Presently, every word in the rundown is arranged utilizing the POS tagger work to have no superfluous words. The words are characterized into various kinds, for example, DET (determiners), CONJ (conjunctions), PRT (particles or other capacity words), NUM (cardinal numbers), X (other: unfamiliar words, errors, shortenings), “.” (accentuation), VERB (action words), NOUN (things), PRON (pronouns), ADJ (modifiers), ADV (intensifiers), and ADP (adpositions). All the stop words and clitics are eliminated so that there are no ambiguities. At that point, standardization happens of the words where fastens are eliminated to ensure that the outcome is the known world in the word reference.

The TF-IDF estimation of everything and the action word would then be determined from the preprocessed rundown of words. The calculations of TF-IDF can be performed using equation (5).

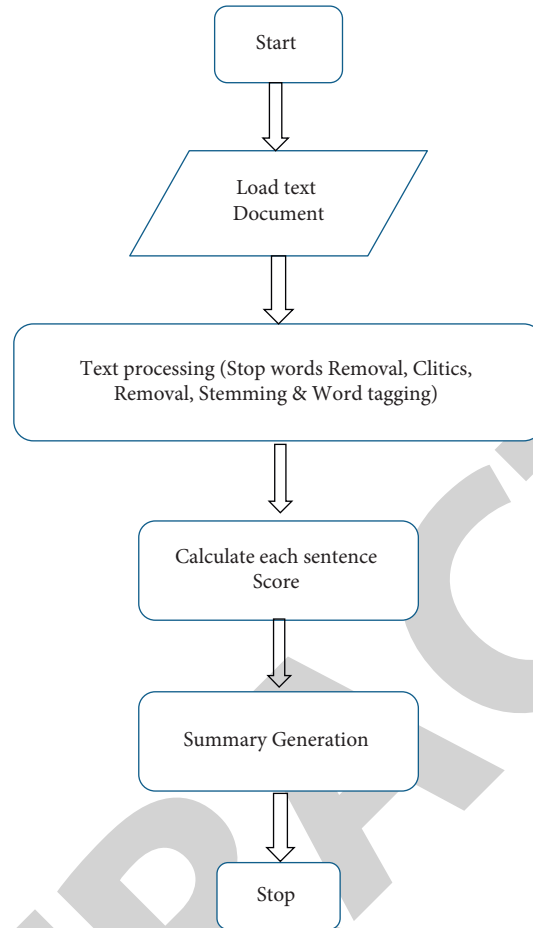


FIGURE 3: TF-IDF technique.

$$TF(t, d) = \frac{\text{Total appearance of the term } t \text{ in a document}}{\text{Total terms in the document } d}, \quad (3)$$

$$IDF(t) = \log \frac{\text{Total number of documents in a document set}}{\text{Document Frequency of the term } t}, \quad (4)$$

$$TF-IDF(t, d) = TF * IDF. \quad (5)$$

The estimation of TF-IDF goes from zero to one with ten-digit accuracy. After being determined, these words are arranged in sliding requests by their worth. At that point, it is incorporated into the new word reference of words, and they are worth it. This arrangement is imperative to break down the position of TF-IDF esteem from the entirety of the words to check the yield rundown. In the wake of knowing the TF-IDF estimation of each word, it can compute the significance estimation of a sentence. The significance estimation is an amount of the estimation of each thing and action word in the sentence. Each sentence in the archive is arranged in a diving request.

Finally, five sentences with the highest TF-IDF esteem are picked. The number of sentences in the last synopsis may change contingent upon the pressured pace of the program

picked by the client. As TF-IDF is an extraction technique, the sentences in the outline are equivalent to the first report. These picked last sentences are arranged as per their appearance in the first archive. For the multirecord outline, the sentences are arranged comparatively with a single report synopsis [36]. The thing that matters is that it begins from the archive, which has the minimal absolute of TF-IDF. The TF-IDF algorithm works by this means.

3.2. LexRank Algorithm. LexRank is an extractive technique used for text synopsis. LexRank method for text summarization where another baby method used is the PageRank method with a sibling TextRank. This learning technique is based on the unsupervised graph. The scoring of sentences is

finished utilizing the diagram strategy. LexRank is utilized for figuring sentence significance dependent on the idea of eigenvector centrality in a chart portrayal of sentences. Under this algorithm, if one sentence is similar to many of the other sentences, it is assumed that it is more important in the document.

This model has a network framework dependent on intrasentence cosine likeness, which is utilized as the continuous grid of the diagram portrayal of sentences [18]. This sentence extraction significantly rotates around the arrangement of sentences with the same plan. For example, a centroid sentence is chosen, filling in the mean for any remaining sentences in the record. Later, the sentences are arranged as per their similarities.

3.2.1. Components of LexRank Algorithm. LexRank algorithm consists of various components, which are discussed as follows:

- (a) Sentences and cosine similarity scores are represented by the graph's node and edges, respectively, as shown in Figure 4

Graphical Approach

- (i) It is based on eigenvector centrality
- (ii) Usually, sentences are placed at the end of vertices of the graphs
- (iii) We can calculate the weight of the edges using the cosine similarity metric

Concerning this graph, S_i are the sentences at the vertices, respectively, and W_{ij} are weights at the end of the edges

- (b) Nodes: TF-IDF vector over each term in the sentence is computed in equations (3) and (4), respectively
- (c) Edges: the similarity between two sentences is then defined by the cosine between two corresponding vectors, as shown in

$$idf - \text{modified} - \text{cosine}(x, y) = \frac{\sum_{w \in x, y} tf_{w,x} tf_{w,y} (idf)^2}{\sqrt{\sum_{x_i \in x} (tf_{x_i, x} idf_{x_i})^2} \sqrt{\sum_{y_i \in y} (tf_{y_i, y} idf_{y_i})^2}} \quad (6)$$

where $tf_{w,s}$ is the number of occurrences of the word w in the sentence s and idf is the inverse document frequency, defined in equation (4).

For generating the summary, we used the "Sumy" library in Python, which uses the LexRank algorithm for generating the summaries of lengthy text.

Methodology:

The prominent approach is an unsupervised graph.

Advantages:

- (i) Maintains redundancy
- (ii) Improves coherency

3.3. TextRank Algorithm. TextRank is used for text preprocessing to determine the keywords and relevant sentences in a given text. It is an unsupervised graph-based ranking model. Then, those sentences are used to generate the text summary. Since the TextRank algorithm is graph-based, the significance of a vertex is determined based on the complete information provided by the graph. The TextRank algorithm makes this decision based on "votes" or "recommendations" of a vertex. All the vertices except for the one being accounted for will vote for a vertex [20]. The importance or value of a vertex is calculated based on the votes received by the vertex. Also, each vertex's vote has its importance calculated by considering the value of the vertex, which is casting a vote. Once all vertices are scored or valued, the vertices with maximum scores are further chosen as important keywords. These keywords are used to determine the key context of the text and the sentences, which should be added to the summary generated.

For using the TextRank algorithm to summarize any textual information, the text must first be transformed into a graph. Various attributes of textual information can be used as vertices of a graph and can be further processed. Such attributes may include words, collocations, and entire sentences. Once the textual information has been transformed into a graph, the vertices are scored based on the above voting system.

The formula used to calculate the score of a vertex is explained as follows.

Formally, let a directed graph with the set of vertices V and set of edges E be represented as $G = (V, E)$, where E is a subset of $V \times V$. Let $\text{In}(V_i)$ be the set of vertices for a given vertex V_i that point to it, and let $\text{Out}(V_i)$ be the set of vertices to which vertex V_i points to. The score of a vertex is defined using equation (7) [20].

$$S(V_i) = (1 - d) + d * \sum_{j \in \text{In}(V_i)} \frac{1}{|\text{Out}(V_j)|} S(V_j). \quad (7)$$

Here, d is the damping factor whose value lies between 0 and 1. It integrates the probability of jumping from a given vertex to another vertex into the graph.

It is an iterative algorithm. Initially, a random value is assigned to each node. Several iterations of the algorithm are performed till convergence below a set threshold. After executing the algorithm thoroughly, each node has a score associated with them, which determines a node's importance.

In the TextRank algorithm, the initial values given to a node will not affect the results or conclusions of the algorithm. However, the number of iterations of the algorithm might affect the results.

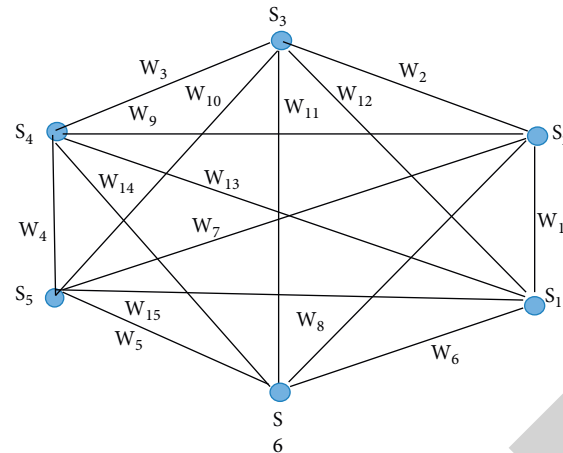


FIGURE 4: Unsupervised graph [18].

Although the TextRank algorithm is used for directed graphs, it can be used on an undirected graph as well in which the out-degree and in-degree of a vertex are equivalent. For loosely connected graphs, undirected graphs have more gradual convergence when the number of edges is proportional to the number of vertices.

3.4. BertSum Algorithm. Bertram is an abstractive summarization algorithm based on BERT (Bidirectional Encoder Representations from Transformers), an unsupervised learning architecture built on top of the Transformer architecture. The BERT architecture has successfully performed more efficiently for a wide range of tasks than the existing models in the NLP space [21].

The BERT architecture was built by Google, along with several published papers and pretrained models that can be used for transfer learning in many domains and various tasks.

BertSum algorithm generates sentence embeddings by using the tokenized textual information given. These sentence embeddings can then be incorporated with the K-means algorithm to calculate the significance of each sentence embedding. The significance of each sentence embedding is determined by calculating its distance from the centroid. Since the algorithm generates sentence embeddings and these sentences can be clustered with a size of k , the size of the summary generated can be controlled by managing the value of k . Previous frameworks and algorithms of abstractive algorithms have not been able to achieve this.

BertSum requires the textual information to be tokenized, i.e., removing too small or too large sentences or sentences or words that require more context to be included in the summary. Several tokenization models then can be used to produce tokenized text. If sentences that fall into the criteria as mentioned earlier were not removed from the data, then it was observed that these sentences/words/pieces of information were rarely used in generating the summary; also, their presence affected the centroid of the data and the

algorithm produced different results, and its performance deteriorated.

K-means algorithm is implemented on the tokenized data to select tokens of more importance and value. The importance of each token is calculated based on each token's distance from the centroid. The algorithm generates the summary based on determined keywords and important sentences. The summary size can be controlled by changing the value of k or the size of the cluster.

The BertSum algorithm has superior performance over all other NLP algorithms. The BertSum algorithm has specific pretraining objectives, it randomly masks 10% or 15% of the input, and the algorithm has to predict the masked word or sentence. In another step, the algorithm takes two sentences, namely, input sentence and candidate sentence. The algorithm has to foretell whether the input sentence correctly comprehends the candidate sentence. The pretraining of the BertSum algorithm is computationally expensive and might take days to pretrain the model even with impressive computational power and resources. Google has launched two pretrained models of the BertSum algorithm for more straightforward implementation by users, more variety of use cases, and better analysis and testing.

3.5. PEGASUS Algorithm. Pretraining with Extracted Gap sentences for Abstractive Summarization (PEGASUS) is an abstractive summarization algorithm that uses the sequence-to-sequence framework, which uses RNNs, based on encoder-decoder architectures. It uses pretrained sequence-to-sequence models with sentences masked and then passed to the encoder-decoder as shown in Figure 5.

The objective is based on predicting missing sentences from the article [28]. Google AI introduces a new state-of-the-art algorithm for doing abstractive summarization. The main contribution of the paper is the introduction of a new pretraining objective for the summarization task. The authors test their transformer-based seq-to-seq summarization model on 12 relevant datasets. The new pretraining objective leads to improved performance over baselines trained

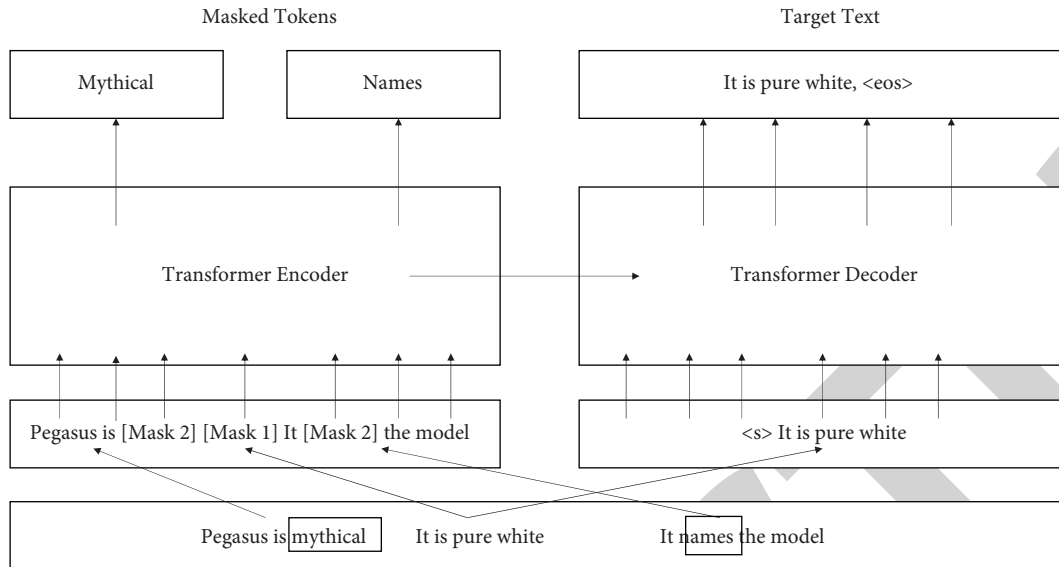


FIGURE 5: Transformer encoder-decoder model [28].

directly on downstream summarization datasets and over alternative pretraining objectives.

So, they have used a seq-to-seq transformer encoder-decoder model to generate a new algorithm known as “Gap Sentences;” here, sentences of significance are selected and masked from the input document. These sentences are altogether used to generate a single sequence from the remaining sentences, and this can be considered following the ideology of extractive technologies.

Let us suppose that we have a document and then use different strategies such as

Random Method: randomly picking “ m ” sentences

Lead Method: selecting the first “ m ” sentences

Principal Method: selecting top “ m ” sentences based on their significance in the document (i.e., picking those sentences that maximally overlap with the document as a whole based on ROUGE-1 metric)

After selecting the best strategy, the sentences are passed in the transformer, where some of them are masked, and a gap is created between them and is fed in the system, expecting that the model will train itself internally and generate the desired output that we ask for.

They have used multiple datasets to pretrain their model like C4, the Colossal, and Clean version of Common crawl, which consisted of 350 M pages, 750 GB of data. This model is run on downstream summarization datasets (12 in number) of different domains (news, science, stories, instructions, emails, etc.). Moreover, the results are prepared and generated in batches, which can be used for evaluation.

4. Datasets Used and Implementation

In this section, we mainly discuss the datasets used in our work and then compare their results using ROUGE metrics.

4.1. Datasets. The sort of data it gets as input primarily determines any algorithm’s performance. Some algorithms perform and give better results for one type but not for the other one. For this paper, we picked two of the most popular datasets from the available datasets on Tensor flow in the category of text summarization. The datasets are MultiNews Dataset [37] and Reddit-TIFU Dataset [38]. Both the datasets have different properties, and in this paper, we have compared the human-generated summaries available from them with the outputs generated by our algorithm.

We have implemented five algorithms in Python and are therefore used for comparing results. Let us discuss the libraries used for implementing the following algorithms:

- TF-IDF algorithm:* the implementation of the TF-IDF algorithm has been done by utilizing the NLTK kit for sentence tokenization. The generated results have been discussed hereinafter
- LexRank algorithm:* SUMY is a Python-based library that helps to extract summaries from HTML pages. It can be considered an automated text summarizer library that has provided the basic implementation for the LexRank algorithm
- TextRank algorithm:* SUMMA provides text summarization algorithms and resources required for the same. It is built by utilizing the GATE API. For this algorithm, we used the SUMMA package for summary generation. It contains the implementation of various algorithms. We have utilized the TextRank algorithm provided by the package
- BertSum algorithm:* BertSum algorithm is implemented and provided by the best-extractive-summarizer library. We have utilized the implementation per our requirements and according to the benchmarks utilized for evaluation and comparison
- PEGASUS algorithm:* for implementing the PEGASUS algorithm, we used inbuilt libraries sentence

piece and transformers, in which we used “google/pegasus-multi_news” and “google/pegasus-red-dit_tifu” models while generating the summaries of text

4.2. Evaluation Metrics. In general, there are three types of evaluations: coselection-based assessment (with a reference summary), document-based assessment (with the original document), and content-based assessment (without reference summary) [32]. We briefly discuss them as follows.

- (a) *Coselection-Based Evaluation Metrics.* This evaluation technique is based on keywords in the system summary, and it necessitates a comparison of reference summaries of the documents. The reference summary and system summary’s common words are chosen and assessed separately. Recall, F-score, and precision are the measurements
- (b) *Content-Based Evaluation Metrics.* This technique assesses the summarizing system in terms that are widely understood. The outline cannot get a network of thoughts, a stream of sentences, the relatedness of sentences to previous phrases, or content curiosity. Every one of these difficulties may be addressed using a content-based approach. We show some content-based assessment methodologies that take into account a text’s varied features. It just necessitates a system overview, which contains metrics like cohesiveness, nonredundancy, and readability
- (c) *Document-Based Evaluation Metrics.* When two phrases in a document have the same relevance, but neither is included in the reference summary, these evaluation metrics fail to assess the system summary properly

Regarding this paper, we have used coselection-based metrics for evaluation, especially the ROUGE framework, which is explained in more detail hereinafter.

4.2.1. ROUGE. Since the mid-2000, the ROUGE metric has been broadly utilized for programmed assessment of outlines [16]. Lin called it Recall-Oriented Understudy for Gisting Evaluation (ROUGE), and he presented various measurements that help in naturally deciding the nature of an outline by comparing it with human (reference) synopses considered mostly as the ground truth.

Different types of ROUGEs are used in comparing different sentences. The granularity of texts compared between the system summaries and reference summaries can be thought of as ROUGE-L, ROUGE-N, and ROUGE-S.

- (a) *ROUGE-N* identifies overlap between unigrams, bigrams, trigrams, and higher-order n-grams
- (b) *ROUGE-L* uses the longest common sentence (LCS) to determine the most extended corresponding sequence of terms. LCS has the benefit of demanding in-sequence matches that capture sentence-level word order rather than sequential matches. You do

not need to specify an n-gram length since it contains the longest in-sequence typical n -grams by default

- (c) *ROUGE-S* is any pair of words in the proper order of a phrase, accounting for gaps. This is referred to as skip-gram concurrence. Skip-bigram, for example, tests the overlap between word pairs with a limit of two spaces between them. For example, the skip-bigrams for the term “dog in the basket” will be “dog in, dog the dog basket, in the, in a basket, the basket”

ROUGE-1 refers to the overlap of unigrams between the device description and the reference summary regarding this study. ROUGE-2 refers to the overlap of bigrams between the method and comparison summaries. Generally, there are three metrics [39] that ROUGE generates for analyzing the results.

- (i) *Recall.* Recall is an aspect of the ROUGE metric that can be considered as the amount of original data given to the model that has been used to generate the summary.
- (ii) *F-Score.* The F-score is a numerical value derived using precision and recall. It is utilized to express the right combination of recall and precision.

$$F\text{-score} = \frac{2 * \text{recall} * \text{precision}}{\text{recall} + \text{precision}} \quad (8)$$

- (iii) *Precision.* Precision refers to the measurable amount of summary generated that was essentially needed or required for generating an efficient summary.

Both the dataset and our algorithm outputs are provided into the ROUGE function, which is used to assess the similarity of two phrases by counting the number of overlapping words and then generating a result in the form of three metrics called recall, F-score, and precision.

5. Result Analysis

Let us consider the result generated by each of the datasets discussed in the section above. The datasets have been selected from the set of datasets available at [40] Tensor flow catalog under the summarization section.

5.1. MultiNews Dataset. This dataset contains a human-generated summary of the various news articles cited on <https://newser.com> [37]. Professional editors have written these summaries and include links to the original articles cited.

For this dataset, the average summary generated by all the examples contains an average of three sentences in the resultant summary. Therefore, for better result generation, we kept a three-sentence summary as a reference. The results generated after using this dataset as a reference summary provider are depicted in Table 1.

We can see from Table 1 that on the MultiNews dataset, TextRank gives the best result out of all the algorithms on ROUGE-1 metrics, and PEGASUS delivers the best

TABLE 1: ROUGE metrics for MultiNews dataset.

No.	Algorithm	Rog-1-f	Rog-1-p	Rog-1-r	Rog-2-f	Rog-2-p	Rog-2-r	Rog-l-f	Rog-l-p	Rog-l-r
1	TF-IDF	0.2971	0.35273	0.25663	0.0821	0.0987	0.0703	0.2495	0.2849	0.222
2	LexRank	0.2941	0.4203	0.22619	0.0765	0.1077	0.0593	0.2307	0.3306	0.1772
3	BertSum	0.2584	0.42442	0.18581	0.0745	0.1325	0.0519	0.2268	0.3501	0.1678
4	TextRank	0.5948	0.60544	0.58456	0.1112	0.0736	0.2276	0.2828	0.2041	0.4605
5	PEGASUS	0.438	0.49796	0.39095	0.1998	0.2261	0.179	0.3734	0.4296	0.3303

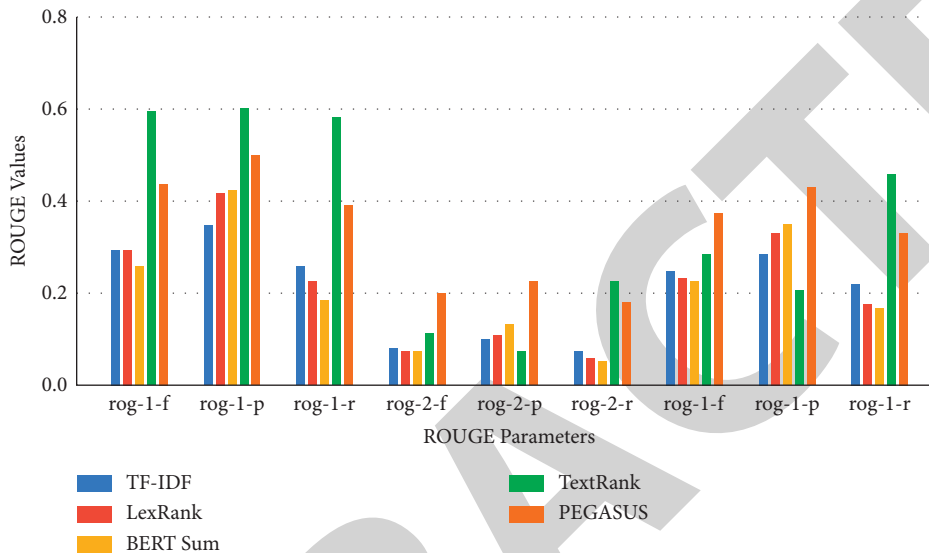


FIGURE 6: Comparison of results of summarization algorithms on MultiNews dataset.

performance for ROUGE-2 and ROUGE-L metrics out of all the compared algorithms. If we compare the overall average of F-score, then PEGASUS has the best F-score for all, and TextRank has the second best average F-score and best among the extractive-based algorithms.

Here is a visual representation of the above-gathered data, which will analyze the performance of different algorithms in Figure 6.

5.2. Reddit-TIFU Dataset. This dataset contains the samples of the Reddit dataset, and TIFU here stands for the subreddit's name [38]. It also contains handwritten summaries of the samples present in the dataset, which are used for reference. For this dataset, the average summary generated for each sample was of a 3-sentence length. Therefore, while fetching the results, we used three-sentence summaries as the generated summary from our algorithms. The results were then compared using the ROUGE library implemented in Python and are shown in Table 2 for all five algorithms.

It is visible from Table 2 that, for the Reddit dataset, the TextRank algorithm gives the best possible results out of the four extractive-based algorithms, which has the highest average of F-score and PEGASUS outperforms them all, as visible in the chart below too. Either of these algorithms can be used for generating summaries of long texts, which are similar to the samples of the Reddit-TIFU dataset.

Here is a visual representation of the above-gathered data, which will analyze the performance of different algorithms in Figure 7.

In this paper, we have compared the algorithms using the two datasets mentioned above. We also observed that, in various other research papers, these algorithms had been compared on different datasets than the ones mentioned here, and our algorithms have shown better results on both the datasets mentioned in this paper. Possibly, TF-IDF, LexRank, and TextRank showed excellent performance [23]. This paper also compares TextRank and LexRank algorithms on the Opinosis dataset and the ROUGE values generated by [41] are presented in Table 3.

It is visible that both these algorithms TextRank and LexRank give better results on Reddit-TIFU and MultiNews dataset when compared to the result generated by the Opinosis dataset.

TextRank algorithm has performed better than other extractive summarization algorithms because of various reasons. TextRank algorithm follows unsupervised learning as there is no requirement of training data set and no human-generated input which allows the algorithm to deliver better results as compared to other algorithms. TextRank algorithm is designed in such a way that due to its internal implementation of PageRank algorithm and generation of the similarity matrix, its performance is better than LexRank and BERT Algorithm.

TABLE 2: ROUGE metrics for REDDIT-TIFU dataset.

No.	Algorithm	Rog-1-f	Rog-1-p	Rog-1-r	Rog-2-f	Rog-2-p	Rog-2-r	Rog-1-f	Rog-1-p	Rog-1-r
1	TF-IDF	0.2095	0.1819	0.4208	0.1251	0.1578	0.1839	0.1282	0.1835	0.3525
2	LexRank	0.2199	0.1183	0.3312	0.1275	0.2034	0.1806	0.1442	0.1709	0.2713
3	BertSum	0.2261	0.1165	0.3887	0.1209	0.1263	0.1832	0.1356	0.1905	0.3362
4	TextRank	0.2159	0.1098	0.5056	0.1258	0.1555	0.2215	0.1258	0.1784	0.4279
5	PEGASUS	0.2376	0.2139	0.3293	0.1845	0.2766	0.2023	0.2175	0.1974	0.3846

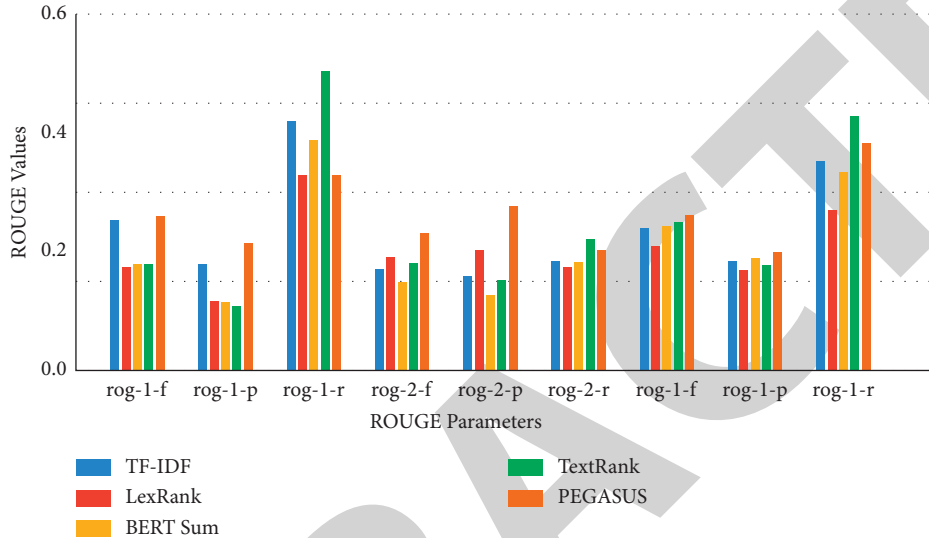


FIGURE 7: Comparison of results of summarization algorithms on REDDIT-TIFU dataset.

TABLE 3: Precision, recall, and F -measure of algorithms.

Algorithm	F -measure	Recall	Precision
TextRank	0.133	0.085	0.382
LexRank	0.19	0.148	0.331

6. Conclusion

Development inaccessibility and prominence of the web have given us a massive measure of crude and chaotic information, which can be put to great use. For simplicity of information, appraisal effective and mechanized synopsis has gotten significant, and a request will probably be filled in coming years. In this paper, we have examined, applied, and assessed diverse extractive synopsis methods, broke down their downsides, and flourished to arrive at an ideal answer for a productive outline. Even though it is not possible to explain the implementation for each algorithm in detail, we have tried to give an insight into each algorithm through our paper and depict the advancements in various techniques for summarization. We have continually focussed on improving the proficiency of rundown strategies, and it has prompted a vigorous establishment for us to work upon.

We have extensively compared the algorithms, i.e., Extractive and Abstractive, on different datasets, and they have shown excellent results and are better than their previous implementations in other papers. This paper

mainly compared them on two popularly known datasets, i.e., Reddit-TIFU and MultiNews, and suggested the best possible algorithm for text summarization out of the five available algorithms. Therefore, it is clear from the analysis that, for both the datasets, PEGASUS delivered the best results among all the algorithms with the highest average F -score and TextRank delivered the best results among all the extractive-based algorithms. Moreover, all the other algorithms used in this paper have also shown better results on both these datasets compared to other datasets used in various other papers mentioned above. This study may be useful for researchers in the future for the selection of the appropriate algorithm for different text summarization. They may directly use PEGASUS for abstractive text summarization and TextRank for extractive text summarization for other datasets.

Tough, automatic text summarization has unlimited scope in the present scenario but one of its crucial applications may be in the summarization of biomedical documents. The traditional approaches in text summarization concerning biomedical documents suffer from fundamental

issues such as its inability to capture clinical context, producing a summary of biomedical documents, and quality of shreds of evidence. The proposed text summarization techniques can be used as one of the tools to retrieve and produce meaningful information to end-users from a huge biomedical repository and thus can help people make complex clinical decisions.

Data Availability

Data will be available on request. For data, kindly contact Divakar Yadav, divakaryadav@nith.ac.in.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] N. Vanetik, M. Litvak, E. Churkin, and M. Last, "An unsupervised constrained optimization approach to compressive summarization," *Information Sciences*, vol. 509, pp. 22–35, 2020.
- [2] R. A. García-Hernández and Y. Ledeneva, "Single extractive text summarization based on a genetic algorithm," in *Mexican Conference on Pattern Recognition*, pp. 374–383, Springer, Berlin, Germany, 2013.
- [3] W. S. El-Kassas, C. R. Salama, A. A. Rafea, and H. K. Mohamed, "Automatic text summarization: a comprehensive survey," *Expert Systems with Applications*, vol. 165, Article ID 113679, 2021.
- [4] R. Nallapati, B. Zhou, C. Gulcehre, and B. Xiang, "Abstractive text summarization using sequence-to-sequence rnns and beyond," 2016, <https://arxiv.org/abs/1602.06023>.
- [5] C. Khatri, G. Singh, and N. Parikh, "Abstractive and extractive text summarization using document context vector and recurrent neural networks," 2018, <https://arxiv.org/abs/1807.08000>.
- [6] S. Singla, N. Duhan, and U. Kalkal, "A novel approach for document ranking in digital libraries using extractive summarization," *International Journal of Computer Applications*, vol. 74, no. 18, pp. 25–31, 2013.
- [7] N. M. Abdelaleem, H. A. Kader, and R. Salem, "A brief survey on text summarization techniques," *IJ of Electronics and Information Engineering*, vol. 10, no. 2, pp. 103–116, 2019.
- [8] A. Elrefaiy, A. R. Abas, and I. Elhenawy, "Review of recent techniques for extractive text summarization," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 23, pp. 7739–7759, 2018.
- [9] A. Sinha, A. Yadav, and A. Gahlot, "Extractive text summarization using neural networks," 2018, <https://arxiv.org/abs/1802.10137>.
- [10] A. Nenkova and K. McKeown, "A survey of text summarization techniques," in *Mining Text Data*, pp. 43–76, Springer, Boston, MA, USA, 2012.
- [11] J. N. Madhuri and R. G. Kumar, "Extractive text summarization using sentence ranking," in *Proceedings of the 2019 International Conference on Data Science and Communication (IconDSC)*, pp. 1–3, IEEE, Bangalore, India, 2019, March.
- [12] K. Vimal Kumar and D. Yadav, "An improvised extractive approach to Hindi text summarization," *Advances in Intelligent Systems and Computing*, Springer, vol. 339, pp. 291–300, New Delhi, 2015.
- [13] M. Allahyari, S. Pouriyeh, M. Assefi et al., "Text summarization techniques: a brief survey," 2017, <https://arxiv.org/abs/1707.02268>.
- [14] M. Dutta, A. K. Das, C. Mallick, A. Sarkar, and A. K. Das, "A graph based approach on extractive summarization," in *Emerging Technologies in Data Mining and Information Security*, pp. 179–187, Springer, Singapore, 2019.
- [15] Y. K. Meena and D. Gopalani, "Evolutionary algorithms for extractive automatic text summarization," *Procedia Computer Science*, vol. 48, pp. 244–249, 2015.
- [16] J. P. Verma and A. Patel, "Evaluation of unsupervised learning based extractive text summarization technique for large scale review and feedback data," *Indian Journal of Science and Technology*, vol. 10, p. 17, 2017.
- [17] J. M. Sanchez-Gomez, M. A. Vega-Rodriguez, and C. J. Pérez, "The impact of term-weighting schemes and similarity measures on extractive multi-document text summarization," *Expert Systems with Applications*, vol. 169, Article ID 114510, 2021.
- [18] G. Erkan and D. R. Radev, "Lexrank: graph-based lexical centrality as salience in text summarization," *Journal of Artificial Intelligence Research*, vol. 22, pp. 457–479, 2004.
- [19] W. Xiao and G. Carenini, "Extractive summarization of long documents by combining global and local context," 2019, <https://arxiv.org/abs/1909.08089>.
- [20] R. Mihalcea and P. Tarau, "Textrank: bringing order into text," in *Proceedings of the 2004 conference on empirical methods in natural language processing*, pp. 404–411, Barcelona, Spain, 2004, July.
- [21] D. Miller, "Leveraging BERT for extractive text summarization on lectures," 2019, <https://arxiv.org/abs/1906.04165>.
- [22] R. Ferreira, L. de Souza Cabral, R. D. Lins et al., "Assessing sentence scoring techniques for extractive text summarization," *Expert Systems with Applications*, vol. 40, no. 14, pp. 5755–5764, 2013.
- [23] R. Rani and D. K. Lobiyal, "An extractive text summarization approach using tagged-LDA based topic modeling," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3275–3305, 2021.
- [24] M. Mojriani and S. A. Mirroshandel, "A novel extractive multi-document text summarization system using quantum-inspired genetic algorithm: mtsqiga," *Expert Systems with Applications*, vol. 171, Article ID 114555, 2021.
- [25] K. V. Kumar, D. Yadav, and A. Sharma, "Graph based technique for Hindi text summarization," *Advances in Intelligent Systems and Computing*, Springer, vol. 339, pp. 301–310, , New Delhi, 2015.
- [26] B. Mutlu, E. A. Sezer, and M. A. Akcayol, "Candidate sentence selection for extractive text summarization," *Information Processing & Management*, vol. 57, no. 6, Article ID 102359, 2020.
- [27] N. Moratanch and S. Chitrakala, "A survey on abstractive text summarization," in *Proceedings of the 2016 International Conference on Circuit, power and computing technologies (ICCPCT)*, pp. 1–7, Nagercoil, India, March 2016.
- [28] J. Zhang, Y. Zhao, M. Saleh, and P. Liu, "Pegasus: pre-training with extracted gap-sentences for abstractive summarization," in *Proceedings of the International Conference on Machine Learning*, pp. 11328–11339, Virtual Event, 2020, November.
- [29] T. Shi, Y. Keneshloo, N. Ramakrishnan, and C. K. Reddy, "Neural abstractive text summarization with sequence-to-sequence models," *ACM/IMS Transactions on Data Science*, vol. 2, no. 1, pp. 1–37, 2021.

Retraction

Retracted: Learning Enhanced Feature Responses for Visual Object Tracking

Computational Intelligence and Neuroscience

Received 1 August 2023; Accepted 1 August 2023; Published 2 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] R. Zhang, C. Fan, and Y. Ming, "Learning Enhanced Feature Responses for Visual Object Tracking," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1241687, 15 pages, 2022.

Research Article

Learning Enhanced Feature Responses for Visual Object Tracking

Runqing Zhang , Chunxiao Fan , and Yue Ming

Beijing Key Laboratory of Work Safety and Intelligent Monitoring, School of Electronic Engineering,
Beijing University of Posts and Telecommunications, Beijing, China

Correspondence should be addressed to Chunxiao Fan; fcxg100@163.com

Received 22 December 2021; Accepted 15 January 2022; Published 8 February 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Runqing 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.

Visual object tracking is an important topic in computer vision, which has successfully utilized pretrained convolutional neural networks, such as VGG and ResNet. However, the features extracted by these pretrained models are high dimensional, and the redundant feature channels reduce target localization and scale estimation precision, leading to tracking drifting. In this paper, a novel visual object tracking method, called learning enhanced feature responses tracking (LEFRT), is proposed, which adopts the target-specific features to enhance target localization and scale estimation responses. First, a channel attention module, called target-specific network (TSNet), is presented to reduce the redundant feature channels. Secondly, the scale estimation network (SCENet) is introduced to extract spatial structural features to generate a more precise response for the scale estimation. Extensive experiments on six tracking benchmarks, including LaSOT, GOT-10k, TrackingNet, OTB-2013, OTB-2015, and TC-128, demonstrate that the proposed algorithm can effectively improve the precision and speed of visual object tracking. LEFRT achieves 90.4% precision and a 71.2% success rate on the OTB-2015 dataset, improving the tracking methods based on the pretrained features.

1. Introduction

Visual object tracking is one of the fundamental tasks in computer vision, widely used in the civil and military fields, such as image segmentation [1], intelligent transportation [2], object detection [3], and human-computer interaction [4].

Recently, pretrained deep features bring state-of-the-art performance to existing trackers, effectively separating foreground objects from the background. However, the tracking task aims to distinguish the object/foreground and the background, and the pretrained features with high dimensions are redundant. What is more, the pretrained VGG and ResNet are trained for the preset object categories, but the target in a tracking task can be an arbitrary object. As shown in Figure 1, the response generated by the pretrained features may focus on the object in the pretrained set. At the same time, tracking methods based on the correlation filters utilized the convolutional features and the a priori scale coefficients to estimate the shape of the targets. The a priori

scale coefficients are set as the discrete constant term parameters, which leads to the precision limit. Therefore, it is of great importance to exploit more compact features to represent specific targets.

In this paper, we propose a novel visual object tracking algorithm called learning enhanced feature responses tracking (LEFRT), as shown in Figure 2, including the TSNet and the SCENet. We proposed the TSNet to reduce the redundant feature channels for the tracking methods based on the correlation filters and the pretrained features. What is more, we initiated the TSNet with the target's appearance in the first frame to reverse the channels paying attention to the given targets. We proposed the SCENet to regress the target's shape. The bounding box regression was introduced to accurately predict the target's finer shape. What is more, the convolutional features lacked spatial topology information for shape estimation, and we introduced 2D-RNN layers to establish the spatial relation for the target's structure.

Specifically, we extract the features using the pretrained model firstly. Secondly, the TSNet is proposed to

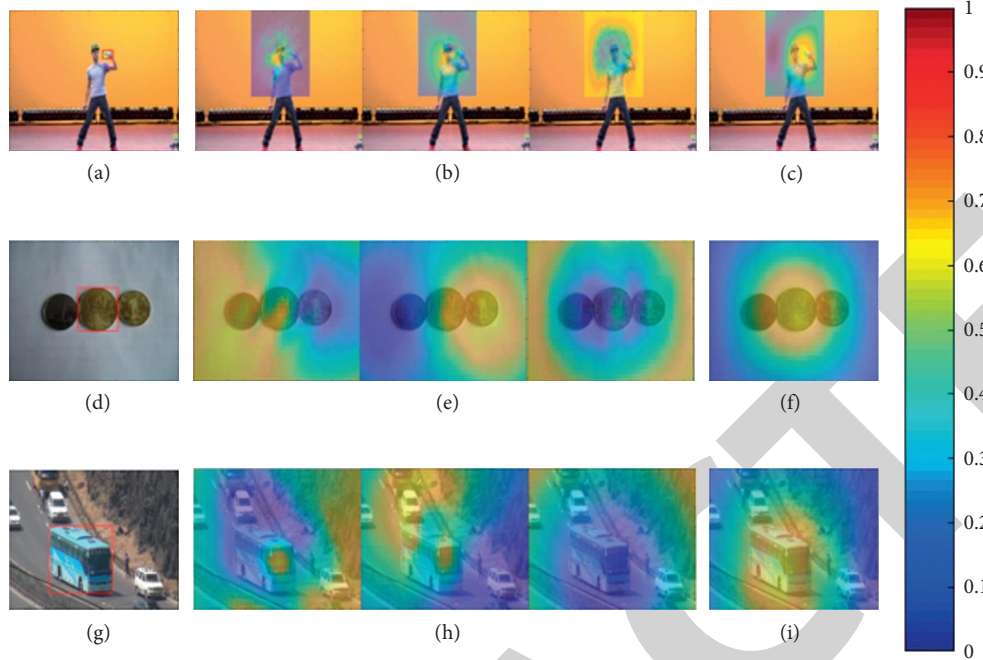


FIGURE 1: Comparison of pretrained (VGG-16) features responses and our target-specific features responses in LaSOT [5] test sets. For example, the left figure is the original frame in the first row, including the target (yoyo) and the human (pretrained distractor). The central figure is the responses generated from VGG-16, which focuses on the human head. The right figure is the response from our methods, which focuses on the target (yoyo). (a) Yoyo-19. (b) Responses from pretrained features. (c) Ours. (d) Coin-6. (e) Responses from pretrained features. (f) Ours. (g) Bus-2. (h) Responses from pretrained features. (i) Ours.

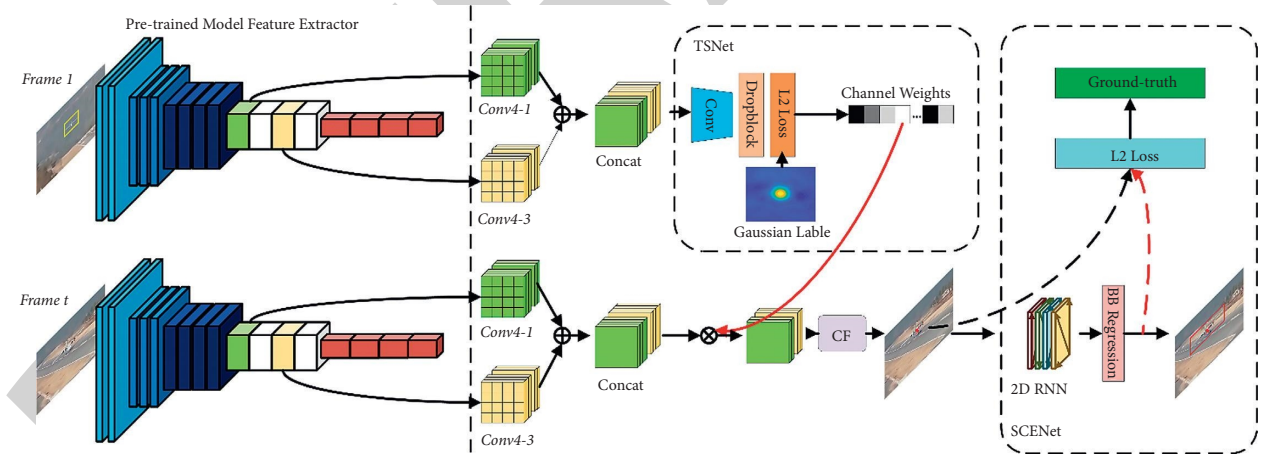


FIGURE 2: The framework of our work.

reduce the redundant channels of the features, and then these features are then sent into the correlation filter for target center localization. We calculate the attention of each feature channel to the target and activate the channels that focus on the given target. Thirdly, we regress the proposal target region on the SCENet to estimate the target's scale with the predicted target center. We construct the spatial structural backbone to establish the spatial relationship of the target image region. The main contributions of our proposed algorithm are summarized as follows:

- (1) We propose the TSNet to generate channel attention and select the effective channels for arbitrary targets, significantly reducing the redundant feature channels and locating the specific target more precisely.
- (2) 2D RNN feature is utilized to represent the spatial structure of the target for scale estimation in the SCENet, which describes the spatial relationship and enhances the response of the target's boundary.
- (3) The proposed algorithm has achieved superior performance on different datasets with many

challenges, including OTB-2013 [6], OTB-2015 [7], TC-128 [8], TrackingNet [9], LaSOT [5], and GOT-10k [10]. The experimental results show that our method effectively improves the precision of both the target's center and scale estimation.

The rest of the paper is organized as follows: we survey related works on visual object tracking in Section 2. Our proposed algorithm is described in Section 3. The following Section 4 demonstrates the experimental evaluations and results. Finally, we conclude the paper in Section 5.

2. Related Work

In this section, we introduce the researches closely related to this work. First, we discuss the features utilized in visual object tracking. Second, we present the scale estimation methods used to estimate the target's shape during tracking.

2.1. Features in Visual Object Tracking. Tracking algorithms utilize features to represent targets' appearance, mainly including traditional manual features, features extracted from end-to-end networks, and features extracted from pretrained models.

Traditional manual features are efficient for visual object tracking. The color attributes method was proposed as a fast color representation, which allowed the tracker to operate at more than 100 frames per second without significant loss in precision [11]. KCF [12] and DSST [13] utilized fHOG to extract features efficiently from different scales and achieved real-time speed. However, traditional manual features perform weakly in complex tracking scenes, such as targets with color change and blurred appearance.

Features extracted from the end-to-end network are adaptive for given tasks and scenes. The end-to-end tracking networks include deep feature methods and matching methods. On the one hand, deep feature methods predict accurate location during tracking. Multi-Domain Convolutional Neural Networks (MDNet) [14] used a multiple-domain network to obtain the convolutional neural network (CNN) features, which achieved 94.8% precision on the OTB50 dataset [7]. To enhance the backbone network without detriment to the tracking speed, DML proposed a mutual-learning-based methodology to use a heavyweight network [15]. On the other hand, matching methods, such as SiamFC [16], Reinforced Attentional Representation (RAR) [17], and SiamRPN [18], utilized Siamese convolution networks to achieve high-speed trackers. End-to-end learning methods use high-dimensional convolutional features to accurately separate foreground and nonsemantic background, but they are computationally expensive to train huge amounts of data on semantic negative pairs [19, 20].

Features that distinguish objects from the pretrained model are efficient without online learning and update. Some correlation filters methods [21–23] combine pretrained features from VGG [24, 25] and ResNet [26, 27]. Joint Group Feature Selection and Discriminative Filter (GFSDCF) [26] combined features of ResNet with correlation filters to achieve group channel responses. ROI Pooled

Correlation Filters (RPCF) [28] introduced deep network and traditional handcraft features for target localization. However, an arbitrary target may not be in the training dataset, making the pretrained models focus on background objects. The pretrained models utilize categories information but lack spatial information, which leads to an inaccurate scale estimation.

2.2. Scale Estimation. The relative motion changes the target's state in the complex tracking scene, including spatial location and scale. The consequent spatial location can be predicted by temporal estimation, such as trajectory [29] and RNN [30], but scale estimation is also important for the target's shape prediction. Tracking methods based on the correlation filter [13, 31–33] applied a multiscale correlation filter with prior parameters to select the scale of the maximum response. Tracking methods based on the regression network used a deep network to predict the bounding box of the target. Some tracking methods introduced the region proposal network (RPN) [18, 19] to estimate the scale variation based on the Siamese network. In addition, ATOM [34] used Intersection over Union (IoU) predictors to estimate IoU for each proposal bounding box. However, deep regression networks, IoU, and RPN networks rely on a huge amount of training data and cost much time. The prior scale coefficient is used in the IoU network and the one-dimensional scale filter, making the scale prediction inaccurate. The object recognition estimated the object's shape accurately [35]. In our work, we learn a spatial response from the first frame and use fast regression to predict the scale variation accurately.

3. Method

This section introduces our novel visual object tracking algorithm for learning enhanced feature response tracking (LEFRT). First, we introduce the framework of our tracking method. Second, we propose the TSNet to select effective features. Finally, we propose a novel SCENet for the scale estimation.

3.1. Tracking Framework. Generally, the proposed LEFRT includes TSNet and SCENet for location prediction and scale estimation separately. As shown in Figure 2, the TSNet reduces redundant feature channels during the training process and retains features focusing on the given target. The SCENet containing 2D RNN layers is trained to enhance the boundary response of the target. During the tracking process, the correlation filters (CF) compute the center coordinates of the target with the input of target-specific features. The scale estimation module uses the spatial structural information in 2D RNN features to estimate the scale variation of the target.

For the TSNet, the pretrained feature extracted model based on VGG-16 is trained with large-scale image database ImageNet. The TSNet fine-tunes the pretrained model using the target appearance from the first frame that focuses on the specific target. The TSNet learns the channel weights of the pretrained

features to pay more attention to the given target in each sequence. We fine-tune the TSNet in the first frame and apply the models in the subsequent frames. We extract the pretrained features and apply the features' channel weights from the trained TSNet to select the most relative channels to the given target when a new frame comes. After feature selection, we use the correlation filter to localize the center location of the target.

For the SCENet, we train the model utilizing a set of annotated video sequences. We combine the CNN layers trained by the Stochastic Gradient Decent (SGD) method and 2D RNN layers to model spatial-relationship between local object areas and obtain a confidence map to estimate object scale. We apply RNNs to modeling the object's structure and use such structure information to enhance the response of objects in the confidence map. The RNNs include four directed acyclic graphs: southeast, southwest, northwest, and northeast. With these directed acyclic graphs, we can perform forward and backward propagation on each directed acyclic graph. The recurrent layer is trained by the method in Section 3.3. We apply the bounding box regression technique [36] to estimate the target's scale.

We also apply a redetection model for our tracking method. We utilize a detection model every 500 frames to obtain ten candidates in the current frame. Then, we compare the similarity between the candidates and the target from the first frame with cosine distance, and the most similar candidate is initialized as the new target template.

3.2. TSNet. Given an arbitrary sequence, we train our TSNet in the first frame to obtain feature channel weights, and then we generate target-specific features with these weights to track the target in the subsequent frames.

3.2.1. Target-Specific Features Extraction. The TSNet is proposed to learn the feature channel weights for the pretrained model. We utilize the VGG-16 network as the pretrained model to extract basic features and discriminate between different videos objects. To reduce the redundant channels that focus on the target, we use the target's appearance in the first frame to train our TSNet. We crop the frame with the given ground truth to achieve the training samples during the training process in the first frame. After extracting the features from the pretrained model, we input the pretrained features into our TSNet.

We regress the pretrained features $\phi(\mathbf{i}, \mathbf{j})$ to the Gaussian labels $\mathbf{y}(\mathbf{i}, \mathbf{j}) = e^{-i^2+j^2/2\sigma^2}$, where (i, j) represents the relative coordinate difference of samples against the target center, and σ is the width of the Gaussian kernel. The loss function for regression can be formulated as

$$L_f = \|\mathbf{y}(\mathbf{i}, \mathbf{j}) - \mathbf{W}_n * \phi(\mathbf{i}, \mathbf{j})\|^2 + \lambda_f \|\mathbf{W}_n\|^2, \quad (1)$$

where $*$ denotes the convolution operation, \mathbf{W}_n is the learnable parameters, and λ_f is a regularization parameter. The center of 2D Gaussian kernel is aligned with the specific target center. For each sequence, we train our TSNet with the feature in the first frame extracted from the pretrained model with given epochs. To alleviate target's deformation, we add a dropblock layer after the convolution layer. The closer the image is to the target center, the higher response $\mathbf{W}_n * \phi(\mathbf{i}, \mathbf{j})$ the features produce.

In order to estimate the importance of each pretrained feature channel in producing the regression response, the learnable parameter \mathbf{W}_n represents the contribution of feature $\phi(\mathbf{i}, \mathbf{j})$ to fit the Gaussian label in equation (1). We utilize the average value of the weight \mathbf{W}_n to compute the importance of the n -th channels Δ_n :

$$\Delta_n = \text{avg}(\mathbf{W}_n). \quad (2)$$

Here, we select the C channels with positive values $\Delta_n > 0$, which have a positive relationship with the target loss. Then, we obtain the target-specific features of the given target in the first frame. In the following video frames, we use the selected channel features to localize the target in the search area.

3.2.2. Target Center Localization. During the tracking process in the subsequent frames, the target-specific features are extracted and input to the correlation filter for target center localization. The appearance model of the correlation filters trackers \mathbf{W}_{CF} is trained on a $M \times N$ pixels sample $\mathbf{x}(\mathbf{i}, \mathbf{j})$. The training samples are generated by circular shifts of $x_{m,n}$, where $(m, n) \in \{1, 2, \dots, M\} \times \{1, 2, \dots, N\}$ with the Gaussian label $y_{m,n}$. In our method, we adopt the target-specific features $\phi(\mathbf{x}(\mathbf{i}, \mathbf{j}))$ instead of the image samples $\mathbf{x}(\mathbf{i}, \mathbf{j})$. The filter \mathbf{W}_{CF} can be solved by minimizing the loss function\cite{kcf}:

$$L_{CF} = \phi(\mathbf{x}(\mathbf{i}, \mathbf{j})) \odot \mathbf{W}_{CF} + \lambda_{CF} \|\mathbf{W}_{CF}\|^2 = \sum_{m,n} (\mathbf{W}_{CF}^T \phi(x_{m,n}) - y_{m,n})^2 + \lambda_{CF} \sum_{m,n} \|\mathbf{W}_{CF}\|^2, \quad (3)$$

where λ_{CF} is a nonnegative regularization parameter. The minimizer of equation (3) has a closed-form, which is acquired by the following formula:

$$\mathbf{W}_{CF} = (\mathbf{X}(\mathbf{m}, \mathbf{n})^T \mathbf{X}(\mathbf{m}, \mathbf{n}) + \lambda_{CF} \mathbf{I})^{-1} \mathbf{X}(\mathbf{m}, \mathbf{n})^T \mathbf{Y}(\mathbf{m}, \mathbf{n}), \quad (4)$$

where the matrix $\mathbf{X}(\mathbf{m}, \mathbf{n})$ has one sample $x_{m,n}$ per row, and each element of $\mathbf{Y}(\mathbf{m}, \mathbf{n})$ is a label $y_{m,n}$. Expressed by fast Fourier transform and inverse fast Fourier transform, the solution \mathbf{W}_{CF} can be acquired by the following formula:

$$\mathbf{W}_{CF} = -F^{-1} \left(\frac{F(x_{m,n}) \odot F(y_{m,n})}{F(x_{m,n}^*) \odot F(x_{m,n}) + \lambda_{CF}} \right), \quad (5)$$

where F , F^{-1} denote the fast Fourier transform and inverse fast Fourier transform, respectively. $x_{m,n}^*$ is a complex-conjugate of $x_{m,n}$ and \odot denotes the element-wise product. The training process is concluded in Algorithm 1.

During tracking, the response map $f(z)$ is calculated by

$$f(z) = -F^{-1} \left(\frac{F(x_{m,n}) \odot F(y_{m,n}) \odot F(z_{m,n})}{F(x_{m,n}^*) \odot F(x_{m,n}) + \lambda_{CF}} \right). \quad (6)$$

Here, $z_{m,n}$ is the patch cropped from the new frame. And the correlation filters place the peak of the response map as the tracking target. In our method, we adopt the target-specific features $\phi(x_{m,n})$ instead of the image samples x in equation (3). The TSNet activates the feature channels that focus on the given target and reduces the feature channels that focus on the background.

3.3. SCENet. After target localization, we propose SCENet to estimate the target's scale and shape. We utilize 2D RNN to establish the spatial relationship between the target and its surrounding regions for scale estimation, which can effectively enhance the boundary responses of the target.

3.3.1. Spatial Structure Feature. RNNs are developed for modeling dependencies in sequential data. Given an input sequence $\{x^{(t)}\}_{t=1,2,\dots,T}$ of length T , the hidden layer $h^{(t)}$ and output layer $f^{(t)}$ at each time step t are calculated with

$$\begin{aligned} h^{(t)} &= \sigma(\mathbf{U}x^{(t)} + \mathbf{W}h^{(t-1)}), \\ f^{(t)} &= \psi(\mathbf{V}h^{(t)}), \end{aligned} \quad (7)$$

where \mathbf{U} , \mathbf{W} and \mathbf{V} represent weight matrices of the hidden layer, the current hidden layer, and the output layer separately; and σ and ψ are nonlinear activation functions. Since the inputs are progressively stored in hidden layers, RNNs can model long-range contextual dependencies among the sequence elements.

Different from one-dimension sequential data, the self-structure of two-dimensional image data is encoded in an undirected cyclic graph. 2D RNN establishes spatial structural relationships from four directions. For each direction, the output of the convolution layer $\phi(\mathbf{x}(\mathbf{i}, \mathbf{j}))$ for the image region $\mathbf{x}(\mathbf{i}, \mathbf{j})$ is input, and the spatial structural relationship can be computed in the 2D RNN unit as follows:

$$\begin{aligned} h_{i,j} &= \sigma \left(\mathbf{U}^R \phi(\mathbf{x}(\mathbf{i}, \mathbf{j})) + \sum_{\alpha} \mathbf{W}^R h_{\alpha} \right), \\ f_{i,j} &= \psi(\mathbf{V}^R h_{i,j}), \end{aligned} \quad (8)$$

where $h_{i,j}$ is the state of the hidden layer, and $f_{i,j}$ is the state of the output layer. \mathbf{U}^R , \mathbf{W}^R , and \mathbf{V}^R are the learnable weight matrices of the hidden layer, the current hidden layer, and the output layer separately. σ and ψ are the nonlinear

activation functions. α is the neighborhood region of $\mathbf{x}(\mathbf{i}, \mathbf{j})$. The parameters of the neighborhood region \mathbf{W}^R establish the relationship between the neighborhood state h_{α} and the current state $h_{i,j}$. The undirected cyclic graph is decomposed into four directions, including southeast, southwest, northwest, and northeast. With equation (8), we can perform forward and backward passes on one directed acyclic graph.

The candidates are sent into the convolution layers during the tracking process to obtain the initial features. With equation (2), our method can generate a new hidden state from the sample to store the representation information, simultaneously containing the location and neighborhood region information. Considering the summation of all hidden layers for the four directions, the forward pass can be calculated as

$$\begin{aligned} h_{i,j}^m &= \sigma \left(\mathbf{U}^{Rm} \phi(\mathbf{x}(\mathbf{i}, \mathbf{j})) + \sum_{\alpha} \mathbf{W}^{Rm} h_{\alpha}^m(\mathbf{x}(\mathbf{i}, \mathbf{j})) \right), \\ f_{i,j}^m &= \psi \left(\sum_m^4 \mathbf{V}^{Rm} h_{i,j}^m \right), \end{aligned} \quad (9)$$

where \mathbf{U}^{Rm} , \mathbf{W}^{Rm} , and \mathbf{V}^{Rm} are matrix parameters for direction m , and the $h_{\alpha}^m(\mathbf{x}(\mathbf{i}, \mathbf{j}))$ is the hidden state of the forward neighborhood of sample $\mathbf{x}(\mathbf{i}, \mathbf{j})$.

3.4. Scale Estimation. Using the target center location predicted from Section 3.3, we can crop k anchor boxes similar with faster-RCNN [37]. For each anchor box p^i , $i \in \{1, 2, \dots, k\}$. We extract the spatial structure feature $\phi(\mathbf{p}^i)$ and then input $\phi(\mathbf{p}^i)$ into the bounding box regression model. The loss function of the bounding box regression model is

$$\mathbf{W}_* = \arg \min_{\mathbf{W}_*} L_2(t_*^i \hat{w}_*^T \phi(\mathbf{p}^i)) + \lambda_{bb} \|\hat{w}_*\|^2, \quad (10)$$

where \hat{w}_* is the learnable parameter. $* \in \{x, y, w, h\}$ denotes the box's center coordinates and its width and height. λ_{bb} is the constant parameter. Label t^i can be calculated as follows:

$$\begin{aligned} t_x &= \frac{\hat{x} - x}{w} \\ t_y &= \frac{\hat{y} - y}{h} \\ t_w &= \log\left(\frac{\hat{w}}{w}\right) \\ t_h &= \log\left(\frac{\hat{h}}{h}\right), \end{aligned} \quad (11)$$

where variables \hat{x} and x are for the predicted box and anchor box, respectively (likewise for y, w, h). The bounding box regression model predicts k bounding boxes, and we average them to obtain the final bounding box.

<p>Input: The first frame of the sequence, I_1; The initial target state, S_1; The pretrained VGGNet (VGG-16), w_1</p> <p>Output: Target-specific features channel weight, Δ_n; Weights of correlation filters W_{CF};</p> <ol style="list-style-type: none"> (1) Crop the search region and extract the pretrained features $\phi(i, j)$ using S_1 for I_1; (2) Generate the Gaussian label map $y(i, j)$; (3) Obtain W_n by equation (1); (4) Compute channel weights Δ_n by equation (2); (5) Obtain target-specific features using channel weights Δ_n and pretrained features; (6) Train correlation filters by equation (4); (7) Return W_{CF}

ALGORITHM 1: Training target center localization model (TSNet).

We propose a fast vision of the SCENet, using the proposal bounding box predicted by correlation filter from Section 3.3 instead of anchor boxes. In the fast vision fSCENet, the feature extraction and bounding box regression model only run once, but the final bounding box's precision might decrease.

4. Experiments

To evaluate the performance of the proposed tracking method, we conduct extensive experiments on three public tracking datasets, including OTB-2013 [6], OTB-2015 [7] dataset, TC-128 [8] dataset, LaSOT [5] dataset, TrackingNet dataset [9], and GOT-10k dataset [10] as shown in Figure 3. Firstly, we detail the implementations and parameters in our experiments and introduce the datasets. Then, in Section 4.1, we evaluate the effectiveness of our tracking method by providing ablation experiments. In Sections 4.2–4.6, we evaluate the performance of the quantitative comparison with several state-of-the-art tracking methods on the OTB-2013/OTB-2015, TC-128, LaSOT, GOT-10k, and TrackingNet datasets, respectively.

Our tracking method is implemented in Matlab using MatconvNet, which runs at an average speed of 30.7 fps with a 2.6 GHz Intel Core i7 CPU with 16 GB RAM and a Titan Xp GPU. In the proposed target-specific features, we utilize the Con43 and Conv41 features from the pretrained model VGG-16. The value of regulation parameter λ_f in equation (1) is set as 0.01. The training epoch for the TSNet is 100, and the learning rate is $5e-7$. The activation functions σ and ψ in the SCENet are tanh and Sigmoid, respectively. The constant parameter λ_{bb} in equation (9) is set as 0.012.

The following datasets are used in our experiments:

OTB-2013/OTB-2015 [6, 7]: OTB-2013/OTB-2015 are popular datasets for visual object tracking. OTB-2013 and OTB-2015 contain 50 and 100 fully annotated videos, respectively, with substantial variations. We adopt the straightforward One-Pass Evaluation (OPE) [6] as the performance evaluation method. For the performance evaluation metrics, we use precision plots and success plots. Following the protocol in the OTB-2015 benchmark, the threshold of 20 pixels and success rate (Succ.) are presented

to compare the representative precision plots and success plots of tracking methods. The main metrics for performance evaluation in the tracking tasks include the precision plot and the success plot. The value in the precision plot can be calculated as $(1/M) \sum_{j=1}^M (\sum_{i=1}^{N_j} 1(|c_{ij} - \hat{c}_{ij}|_2 < \tau_p) / N_j)$, while the value in the success rate can be calculated as $(1/M) \sum_{j=1}^M (\sum_{i=1}^{N_j} \text{IoU}((c_{ij}, \hat{c}_{ij}) < \tau_s) / N_j)$. The ground-truth of the target's state is c_{ij} , while the predicted value is \hat{c}_{ij} . M is the total number of video sequences in the current dataset. N_j represents the total number of images in the j -th video sequence. IoU is the Intersection over Union function to calculate the overlap rate between the ground-truth and the predicted value. When $\tau_p = 20$, the value in the precision plot is chosen as the precision. When $\tau_s = 0.5$, the value in the success plot is chosen as the success rate.

TC-128 [8]: the TC-128 dataset contains 128 fully annotated color video sequences with many challenging factors. The dataset includes sequences from two main sources: 50 previous studies and 78 new collections.

LaSOT [5]: the LaSOT dataset has longer sequences with an average of 2,500 frames per sequence, including a public dataset and test set for tracking. All the videos in the LaSOT test set are annotated with eight challenges, including Aspect Ratio Change (ARC), Low Resolution (LR), Out-of-view (OV), Fast Motion (FM), Full Occlusion (FO), Scale Variation (SV), Rotation (RO), and Deformation (DEF). We evaluate our approach on the test set consisting of 280 videos, and online model adaption is crucial for this dataset.

GOT-10k [10]: GOT-10k is a large benchmarking dataset for visual object tracking. It has 180 test video sequences in 84 categories. The main metrics of GOT-10k include AO, SR0.5, SR0.75, and speed (Hz). SR0.5 and SR0.75 are the SR metric with a threshold of 0.5 and 0.75 separately.

TrackingNet [9]: the TrackingNet Dataset contains 511 sequences for visual object tracking in the wild. The main metrics of TrackingNet Dataset include Precision, NormalizePrecision, and Success. The precision and success are measured separately as the centers' distance and the Intersection over Union (IoU). The NormalizePrecision is measured as the weighted precision, and the weights are calculated by the size of the ground truth bounding box.



FIGURE 3: Some samples of the sequences in our evaluation dataset.

4.1. Ablation Studies. In this section, we conduct the ablation study on OTB-2013 and OTB-2015 to evaluate the effectiveness of each component. The main components contain the TSNet and the SCENet. To show the contribution of the components, four variants of LEFRT are designed:

Baseline: the original correlation filter tracker (see Section 3.2.2) uses the pretrained VGG-16 features, and the scale estimation is based on the prior scale coefficients.

Baseline + TSNet: a variant without the proposed scale estimation strategy, which applies the TSNet to localize the target center.

Baseline + TSNet + SCENet-: a variant replaces the SCENet in scale estimation strategy with CNN layers SCENet-

LEFRT-HS: an accelerated version of LEFRT, which applies fSCENet to estimate the scale.

As shown in Table 1, we evaluate three variants on the OTB-2015 dataset and compare their tracking performance with the proposed LEFRT. All the variants perform worse than LEFRT in terms of tracking accuracy. Pre-Train-T extracts the features from the Conv41 and Conv43 layers of pretrained VGG-16 models, and it estimates the target's scale in the subsequent frames using the prior scale coefficient. The baseline method utilized the pretrained CNN feature from VGG-16. The high dimensional pretrained features are trained for the classification task, and not each channel is useful to distinguish the object and the background. Redundant features increase the parameters of the

model and pay too much attention to the background image, resulting in model drift. Our TSNet selected the feature channels that focused on the target itself and inactivated the redundant channels, improving the tracking model's speed and precision. The baseline method utilized the scale pyramid with the a priori scale coefficients, which limited the accuracy of prediction results. We utilized the bounding box regression to estimate a more accurate target's shape. What is more, the 2D-RNN features contain rich self-structure information of the target, which improves the ability of features to describe the target shape. In the comparison, the accelerated version LEFRT-HS using the standard high-speed features for scale estimation obtains a higher speed than LEFRT. The tracking speed of LEFRT-HS is similar to TS, which indicates that our scale estimation strategy has a fast speed.

4.2. Evaluation on OTB-2013/OTB-2015 Dataset. We conduct the quantitative, qualitative, and challenging attributes on the OTB-2015 dataset. What is more, we also conduct the ablation study experiment on the OTB-2013 dataset.

4.2.1. Quantitative Comparison. We compare the precision score and success rate obtained by our LEFRT and several state-of-the-art tracking methods including SiamFC++ [38], ASRCF [24], ARCF [39], UDT [20], ECO [40], CNN-SVM [41], fDSST [13], ATTF [21], TADT [42], MLT [43], CFNet [44], SiamFC [16], Siam-tri [45], EACOFT [22], and Dinesh et al. [23]. Since our tracking method adopts target-specific

TABLE 1: The precision scores (Prec₂₀), the success rate (Succ.) scores and speed (fps) on the OTB-2015 dataset. The best results are displayed in bold fonts.

Methods	TSNet	SCENet	Prec ₂₀	Succ.	Speed (FPS)
Baseline			81.1	65.6	46.1
Baseline + TSNet	✓		86.5	69.7	49.0
Baseline + TSNet + SCENet	✓	SCENet	86.5	70.1	35.0
LEFRT-HS	✓	fSCENet	90.4	70.6	42.3
LEFRT	✓	fSCENet	90.4	71.2	30.7

TABLE 2: The precision score and the success rate (Succ.) on the OTB-2013 and OTB-2015 datasets. The best results are displayed in bold fonts.

Methods	OTB-2013		OTB-2015	
	Precision	Succ.	Precision	Succ.
CNN-SVM	76.0	58.2	81.4	55.4
fDSST	74.0	55.4	69.3	52.0
STDCE	63.9	—	—	—
MLT	—	62.1	74.5	59.3
SiamFC	70.8	52.2	68.4	49.8
TADT	80.2	66.5	78.4	65.0
ASRCF	86.0	71.4	84.1	68.8
ARCF	77.1	63.9	74.8	61.2
UDT	73.1	60.8	68.7	57.9
SiamFC++	—	—	—	68.3
EACOF	—	—	—	67.6
AFFT-50	—	—	89.9	69.3
Dinesh et al.	84.6	71.2	83.7	53.5
LEFRT	91.3	73.5	90.4	71.8

features to select the channels of pretrained features, we choose ASRCF, ARCF, UDT, CNN-SVM, and ECO, because they use pretrained models to extract features. The pretrained features contain redundant channels, which lead to the model drifting for the tracking task. AdaCFNet, TADT, and MLT applied attention mechanism to optimize features. CFNet, SiamFC, Siamtri, fDSST, MLT, and RAR use prior scale coefficient to estimate the scale, limiting the accuracy for target’s shape estimation.

Table 2 demonstrates the comparisons of the precision scores, success rate, and speeds obtained by our LEFRT and other state-of-the-art tracking methods. Our LEFRT performs favorably against other state-of-the-art tracking methods in precision scores and success rate. Compared with high-speed trackers based on correlation filters such as fDSST, our LEFRT improves accuracy in precision scores and success rate noticeably. Compared with the trackers based on the correlation filters with deep learning features, such as ASRCF and ARCF, our LEFRT shows performance superiority. The proposed scale estimation strategy provides high accuracy to track the target with scale variation compared with the other methods.

4.2.2. Challenges Attributes. In this subsection, we compare the success plots between our LEFRT and other tracking methods with eleven challenging attributes, including background clutter (BC), deformation (DEF), fast motion

(FM), in-plane rotation (IR), low resolution (LR), illumination variation (IV), motion blur (MB), occlusion (OCC), out-of-plane rotation (OR), out of view (OV), and scale variation (SV).

In Figure 4, compared with the state-of-the-art tracking methods, our LEFRT obtains much better performance under the most challenging attributes, beneficial for target-specific features. Targets that did not get trained in the pretrain process can still be tracked robustly, because our TSNet dropped the redundant feature channels for the background objects. Our LEFRT also achieves significant performance improvements under SV. This proves that the proposed scale estimation effectively predicts the target’s size. For the deformation attribute, LEFRT performs worse than ECO and ASRCF since the model is sensitive to the deformation of the targets. Even so, LEFRT obtains a higher tracking accuracy than ECO and ASRCF on the whole dataset.

4.2.3. Qualitative Comparison. In Figure 5, we qualitatively compare the performance obtained by our LEFRT, ASRCF [24], ECO [40], ARCF [39], TADT [42], CFNet [44], and SiamFC [16] on five challenging sequences.

Our LEFRT accurately tracks the object in terms of both positions and scale for the most challenging sequences, while most tracking methods fail to locate the target positions or incorrectly estimate the target scale. For the sequences of CarScale (row 1), the compared tracking methods locate the

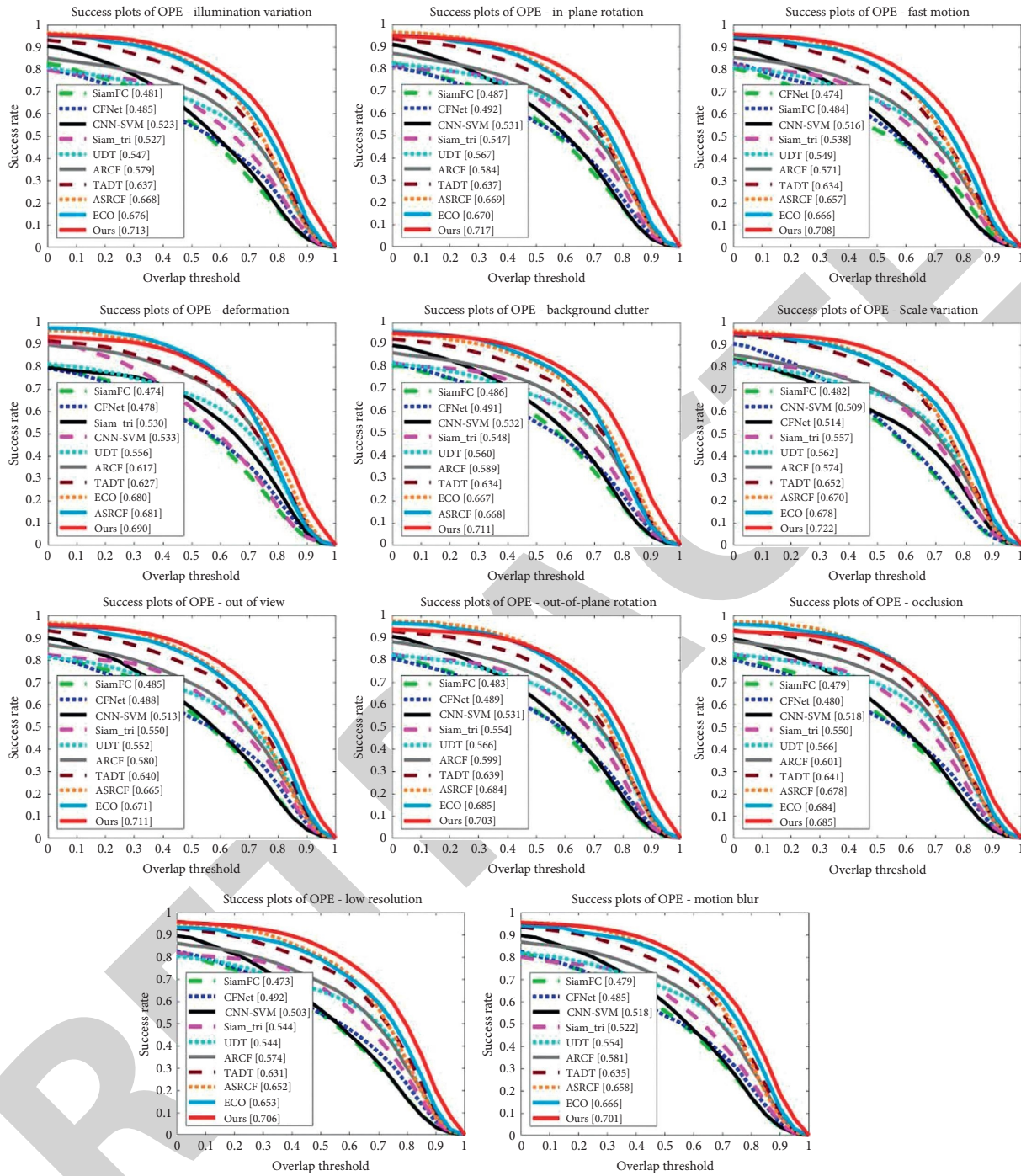


FIGURE 4: The success plots on the OTB-2015 dataset for eleven challenging attributes, including background clutter, deformation, fast motion, in-plane rotation, low resolution, illumination variation, motion blur, occlusion, out-of-plane rotation, out of view, and scale variation. Only shown are the top ten performance tracking methods.

target position correctly, but they only discriminate a part of the object instead of the whole object when they undergo the large-scale variation. At the same time, our LEFRT correctly estimates both the position and scales of the object. For the sequences of Ironman and Matrix (row 2 and row 3), the most compared tracking methods drift away because of the significant illumination variation and occlusion. In contrast,

our LEFRT successfully handles these challenges and accurately tracks the object despite the complex backgrounds. In the sequences of MotorRolling and Skiing (row 4 and row 5), the compared tracking methods are hard to realize the stable tracking when encountering fast motion and significant rotation, while our LEFRT keeps robust tracking of the object throughout the sequences.

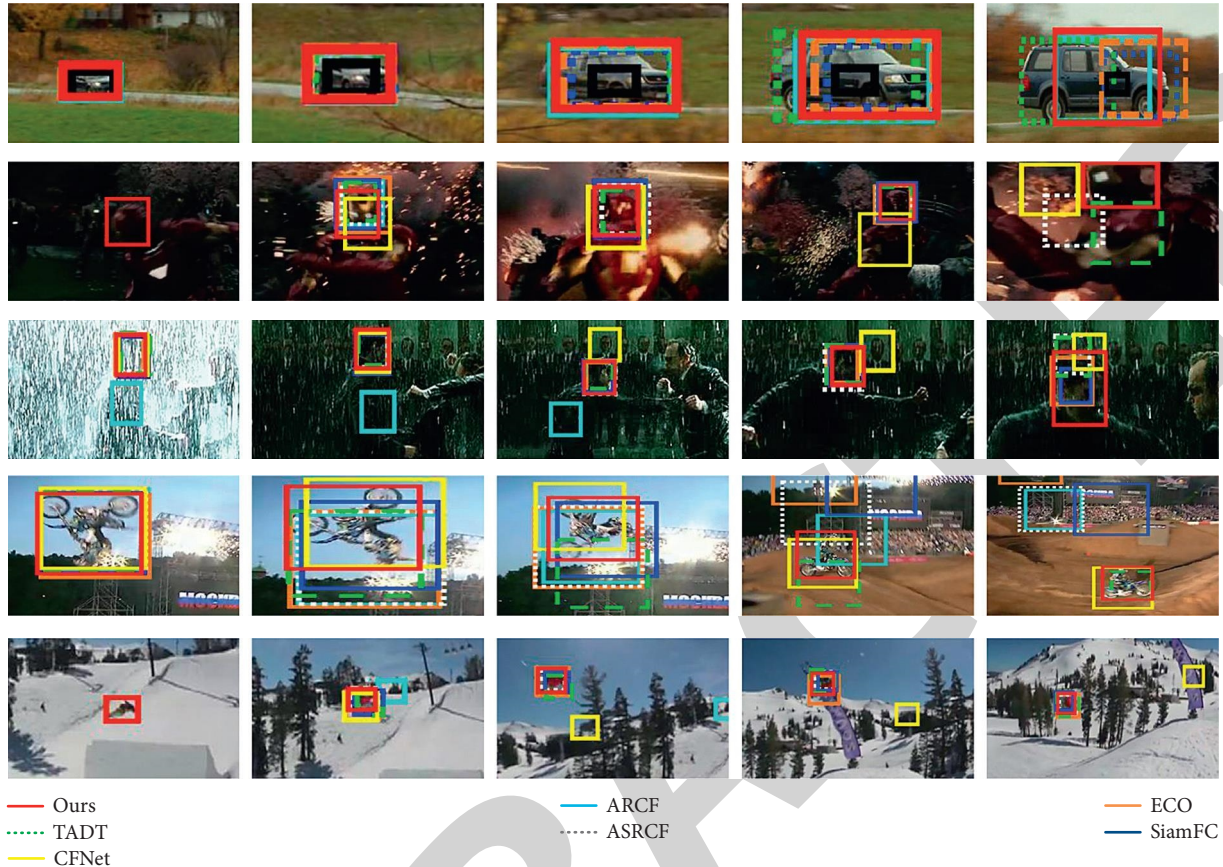


FIGURE 5: Qualitative results of our LEFRT, ASRCF [24], ECO [40], ARCF [39], TADT [42], CFNet [44], and SiamFC [16] on OTB-2015 dataset (from top to down: CarScale, Ironman, Matrix, MotorRolling, and Skiing, respectively).

4.3. Evaluation on TC-128 Dataset. We compare our LEFRT with several state-of-the-art tracking methods that have publicly available results on the TC-128, including ASRCF [24], ARCF [39], ECO [40], DSST [13], TADT [42], UDT [20], GFSDCF [26], and IGSSRTCF [46]. As shown in Table 3, we can observe that our LEFRT achieves the best performance in both precision plots and success plots among all the compared tracking methods. LEFRT outperforms the other tracking methods that use deep features,

That is, there are ECO, ASRCF, and TADT, with a relative improvement of 0.6% and 0.8% compared with ECO, respectively. Compared with the tracking methods based on end-to-end networks, such as SiamFC and CFNet, our LEFRT achieves higher tracking accuracy, and the visual results are shown in Figure 6.

4.4. Evaluation on LaSOT Dataset. We conduct the quantitative comparison and the challenge attributes on the LaSOT dataset.

4.4.1. Quantitative Comparison. We compare our LEFRT with several state-of-the-art tracking methods that have publicly available results on the LaSOT, including ATOM

[34], SiamFC++ [38], Ocean [47], ASRCF [24], ARCF [39], ECO [40], SiamFC [16], CFNet [44], MCCT [33], TADT [42], UDT [20], SiamDW [48], and SiamRPN++ [19]. Table 4 shows the precision plots and the success plots of the comparisons between our LEFRT and the state-of-the-art tracking methods. Our LEFRT ranks first with a precision of 0.522 and a success rate of 0.541 in this dataset, verifying the effectiveness of the proposed target-specific features and scale estimation strategy. Ocean trains a special model for the LaSOT dataset and achieves better performance than ours. However, the model of the proposed method is universal and does not need to be fine-tuned for each dataset.

4.4.2. Challenge Attributes. We further analyze the performance of LEFRT under different challenges on the LaSOT test dataset.

For the most attributes in LaSOT test dataset, our LEFRT performs the first rank against other compared tracking methods. Benefiting from localization based on the target-specific features, our LEFRT achieves high precision in most cases in Figure 7. At the same time, our LEFRT can obtain the best performance in scale variation. Compared with other tracking methods that utilize the prior scale

TABLE 3: Precision plots and success plots on TC-128 dataset.

Methods	Precision	Success rate
UDT [20]	0.310	0.103
DSST [13]	0.543	0.477
ARCF [39]	0.702	0.646
TADT [42]	0.765	0.703
ASRCF [24]	0.784	0.723
ECO [40]	0.789	0.729
GFSDCF [26]	0.831	0.605
IGSSRTCF [46]	0.777	0.573
Ours	0.841	0.755



FIGURE 6: Qualitative results of our LEFRT, ASRCF [24], ECO [40], ARCF [39], TADT [42], DSST [13], and UDT [20] on TC-128 dataset (from top to down: Basketball, Iceskater, Jogging, Singer1, and Skyjumping, respectively).

TABLE 4: The precision scores, the precision (threshold = 25), and success rate (threshold = 0.5) on the LaSOT test dataset.

Methods	Precision	Success rate
SiamFC++	0.512	0.537
Ocean	0.566	0.560
ATOM	0.479	0.515
SiamRPN++	0.467	0.496
SiamDW	0.378	0.385
TADT	0.370	0.397
ASRCF	0.328	0.344
ARCF	0.302	0.324
MCCT	0.236	0.259
UDT	0.245	0.250
CFNet	0.263	0.275
ECO	0.302	0.324
SiamFC	0.331	0.336
Ours	0.522	0.541

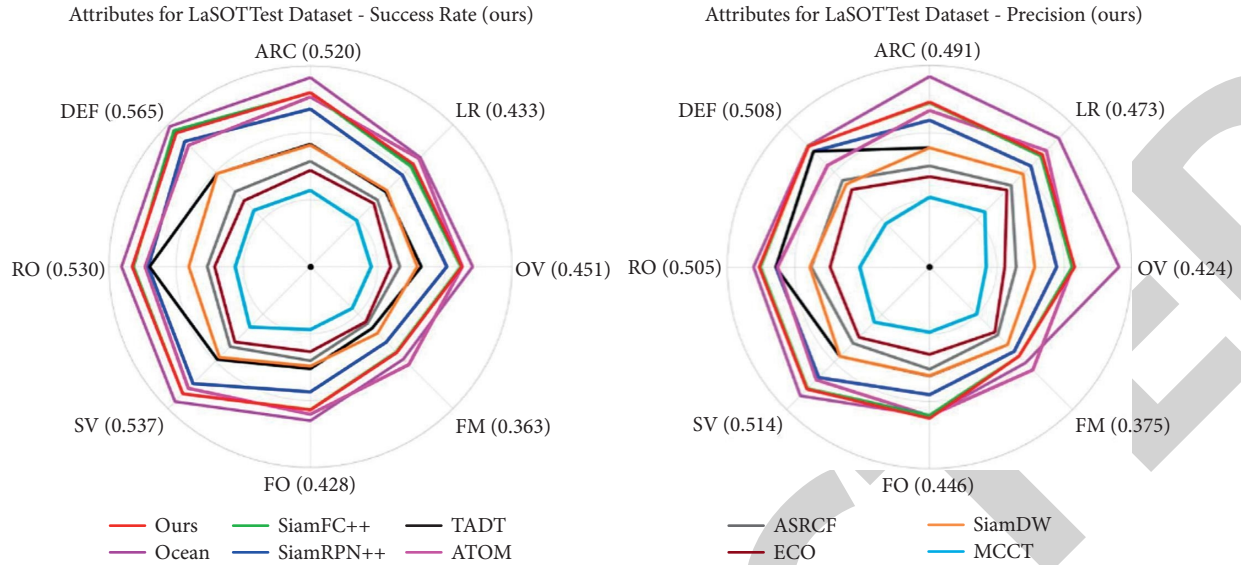


FIGURE 7: The precision and success rate plots (top ten) on the LaSOT test dataset for challenging attributes.

parameters, our SCENet predicts a more accurate target's shape. Our LEFRT performs worse than SiamFC++ for deformation, which is likely that the feature focuses on the target's appearance in the first frame, and the model is sensitive to the target deformation. Even so, LEFRT obtains a higher tracking accuracy than others on the whole dataset. The visual results are shown in Figure 8.

4.5. Evaluation on GOT10K Dataset. In Table 5, we compare our LEFRT with several state-of-the-art tracking methods that have publicly available results on the GOT-10K, including SiamFC++ [38], Ocean [47], TADT [42], SiamDW [48], SiamRPN++ [19], ASRCF [24], and Autotrack [49]. Our LEFRT achieves 0.619 AO, 0.721 $SR_{0.5}$, and 0.477 $SR_{0.75}$ separately, which is better than most of the methods based on the correlation filter and Siamese network. The pretrained features utilized in the correlation filter lead to the model drifting, which decrease the precision and success rate. Compared with the end-to-end Siamese network, the proposed LEFRT utilized the incremental update to avoid model drifting by the deformation targets. Our method is not as good as Ocean, because Ocean utilizes anchor-free image segmentation for target shape estimation, and the metrics of GOT-10k pay more attention on the overlap rate. Same as the LaSOT dataset, Ocean also trains a special model for the GOT-10k dataset. Compared with Ocean, the model of the proposed method is more comprehensive.

4.6. Evaluation on TrackingNet Dataset. In Table 6, we compare our LEFRT with several state-of-the-art tracking methods on the TrackingNet, including SiamFC++ [38], Ocean [47], TADT [42], SiamDW [48], SiamRPN++ [19], ASRCF [24], and Autotrack [49]. The proposed method has the advantage in precision and achieves 71.23% Precision and 81.92% NormalizePrecision separately. Compared with the state-of-the-art tracking methods based on the correlation filters, our method improves the effectiveness of the pretrained features. However, the speed of LEFRT may decrease because the time complexity of 2D-RNN is higher than the convolutional layer. Different from the results on other datasets, Ocean achieves a lower overlap rate and obtains a lower Success score than ours. We check the results and find that Ocean trained independent models for each of the other datasets except TrackingNet dataset. Because the TrackingNet dataset does not provide the ground truth file, the Ocean cannot fine-tune the model specifically. And in this experiment, we have to test the Ocean tracker using its OTB2015 model. The result proves that the comprehensiveness of our method is higher. SiamDW utilized the ResNet as the feature extraction backbone, which has deeper network structure. SiamDW enlarged the receptive field for the convolutional layer, which enlarged the feature size and stride. Therefore, the estimation for the target's shape could be carried out on a larger feature map, which achieves a higher success rate than ours.



FIGURE 8: Qualitative results of our LEFRT, ASRCF [24], ECO [40], ARCF [39], TADT [42], SiamRPN++ [19], and SiamDW [16] on LaSOT test dataset (from top to down: Basketball-6, Bear-4, Bicycle-9, Bicycle-18, and Bird-15, respectively).

TABLE 5: The precision scores, the average overlap (AO), success rate (SR) $SR_{0.5}$, $SR_{0.75}$, and speed (Hz) on the GOT-10k dataset.

Methods	AO	$SR_{0.5}$	$SR_{0.75}$	Speed (FPS)
SiamFC++	0.610	0.716	0.463	40.65
Ocean	0.649	0.728	0.597	22.79
SiamRPN++	0.611	0.717	0.492	34.05
SiamDW	0.607	0.731	0.476	27.37
TADT	0.367	0.389	0.130	11.39
ASRCF	0.313	0.317	0.113	11.11
AutoTrack	0.286	0.277	0.109	15.37
Ours	0.522	0.721	0.477	21.19

TABLE 6: The precision, normalize precision, and success on the TrackingNet dataset.

Methods	Precision	Normalize precision	Success rate
SiamFC++	70.88	81.06	74.33
Ocean	68.28	79.01	69.22
SiamRPN++	69.38	79.98	73.30
SiamDW	63.75	76.06	79.92
TADT	53.99	67.11	59.28
AutoTrack	44.08	55.56	57.63
Ours	71.23	81.92	75.71

5. Conclusions

This paper proposes a novel visual object tracking method, LEFRT, which can effectively track an arbitrary object and estimate the scale variation. Specifically, the TSNet can generate effective responses for target localization even if the target is not in the pretrained dataset. Moreover, the 2D RNN structure in the SCENet enhances the boundary responses of the target, which produces a precise shape for the target's scale estimation. Experiments on six challenging datasets demonstrate that our method effectively improves the target's center precision and the average overlap rate. Although the proposed tracking method has achieved competitive tracking results, its performance can be improved for long-term tracking in complex scenes. In future work, our research will further focus on long-term tracking in complex scenes. To improve the tracking performance in the deformation and occlusion scenes, we will utilize the Transformer to extract the temporal information for the target localization.

Data Availability

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

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

- [1] Q. Wang, L. Zhang, L. Bertinetto, W. Hu, and P. H. Torr, "Fast online object tracking and segmentation: a unifying approach," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 1328–1338, Singapore, June 2019.
- [2] L. Bazzani, M. Cristani, D. Tosato et al., "Social interactions by visual focus of attention in a three-dimensional environment," *Expert Systems*, vol. 30, no. 2, pp. 115–127, 2013.
- [3] Z. Qiu, F. Wang, and Z. Pan, "Adaptive adjustment object detection algorithm under multiple mechanisms based on gan," *Scientific Programming*, vol. 2021, Article ID 5875320, 7 pages, 2021.
- [4] H. Wang, C. C. Cheah, W. Ren, and Y. Xie, "Passive separation approach to adaptive visual tracking for robotic systems," *IEEE Transactions on Control Systems Technology*, vol. 26, no. 6, pp. 2232–2241, 2018.
- [5] H. Fan, L. Lin, F. Yang et al., "Lasot: A high-quality benchmark for large-scale single object tracking," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 5374–5383, June 2019.
- [6] Y. Wu, J. Lim, and M.-H. Yang, "Online object tracking: a benchmark," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 2411–2418, Portland, Oregon, June 2013.
- [7] Y. Wu, J. Lim, and M. H. Yang, "Object tracking benchmark," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, p. 1, 2014.
- [8] P. Liang, E. Blasch, and H. Ling, "Encoding color information for visual tracking: algorithms and benchmark," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5630–5644, 2015.
- [9] M. Muller, A. Bibi, S. Giancola, S. Alsubaihi, and B. Ghanem, "Trackingnet: a large-scale dataset and benchmark for object tracking in the wild," in *European Conference on Computer Vision*, pp. 300–317, Springer Science, New York, NY, USA, 2018.
- [10] L. Huang, X. Zhao, and K. Huang, "Got-10k: a large high-diversity benchmark for generic object tracking in the wild," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 1, pp. 1–10, 2019.
- [11] M. Danelljan, F. Shahbaz Khan, M. Felsberg, and J. Van de Weijer, "Adaptive color attributes for real-time visual tracking," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 1090–1097, Columbus, OH, USA, June 2014.
- [12] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 583–596, 2014.
- [13] M. Danelljan, G. Häger, F. S. Khan, and M. Felsberg, "Discriminative scale space tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 8, pp. 1561–1575, 2016.
- [14] H. Nam and B. Han, "Learning multi-domain convolutional neural networks for visual tracking," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pp. 4293–4302, Las Vegas, NV, USA, June 2016.
- [15] H. Zhao, G. Yang, D. Wang, and H. Lu, "Deep mutual learning for visual object tracking," *Pattern Recognition*, vol. 112, Article ID 107796, 2021.
- [16] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, and P. H. Torr, "Fully-convolutional siamese networks for object tracking," in *European Conference on Computer Vision*, pp. 850–865, Springer Science, New York, NY, USA, 2016.
- [17] P. Gao, Q. Zhang, F. Wang, L. Xiao, H. Fujita, and Y. Zhang, "Learning reinforced attentional representation for end-to-end visual tracking," *Information Sciences*, vol. 517, pp. 52–67, 2020.
- [18] B. Li, J. Yan, W. Wu, Z. Zhu, and X. Hu, "High performance visual tracking with siamese region proposal network," in *Proceedings of the International Conference on Computer*

Retraction

Retracted: Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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] N. Krishnaraj, B. Sivakumar, R. Kuppasamy, Y. Teekaraman, and A. R. Thelkar, "Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8501738, 13 pages, 2022.

Research Article

Design of Automated Deep Learning-Based Fusion Model for Copy-Move Image Forgery Detection

N. Krishnaraj ¹, B. Sivakumar,² Ramya Kuppusamy ³, Yuvaraja Teekaraman ⁴,
and Amruth Ramesh Thelkar ⁵

¹Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, India

²Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kattankulathur 603203, Chennai, India

³Department of Electrical and Electronics Engineering, Sri Sairam College of Engineering, Bangalore 562106, Karnataka, India

⁴Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 3JD, UK

⁵Faculty of Electrical & Computer Engineering, Jimma Institute of Technology, Jimma University, Jimma, Ethiopia

Correspondence should be addressed to Yuvaraja Teekaraman; yuvarajastr@ieee.org and Amruth Ramesh Thelkar; amruth.rt@gmail.com

Received 28 October 2021; Revised 3 January 2022; Accepted 12 January 2022; Published 31 January 2022

Academic Editor: Deepika Koundal

Copyright © 2022 N. Krishnaraj 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.

Due to the exponential growth of high-quality fake photos on social media and the Internet, it is critical to develop robust forgery detection tools. Traditional picture- and video-editing techniques include copying areas of the image, referred to as the copy-move approach. The standard image processing methods physically search for patterns relevant to the duplicated material, restricting the usage in enormous data categorization. On the contrary, while deep learning (DL) models have exhibited improved performance, they have significant generalization concerns because of their high reliance on training datasets and the requirement for good hyperparameter selection. With this in mind, this article provides an automated deep learning-based fusion model for detecting and localizing copy-move forgeries (DLFM-CMDFC). The proposed DLFM-CMDFC technique combines models of generative adversarial networks (GANs) and densely connected networks (DenseNets). The two outputs are combined in the DLFM-CMDFC technique to create a layer for encoding the input vectors with the initial layer of an extreme learning machine (ELM) classifier. Additionally, the ELM model's weight and bias values are optimally adjusted using the artificial fish swarm algorithm (AFSA). The networks' outputs are supplied into the merger unit as input. Finally, a faked image is used to identify the difference between the input and target areas. Two benchmark datasets are used to validate the proposed model's performance. The experimental results established the proposed model's superiority over recently developed approaches.

1. Introduction

Recently, the extension of Internet services and the strengthening and proliferation of social networks such as Reddit, Facebook, and Instagram had had an important effect on the number of content prevailing in digital media. As per the International Telecommunication Union (ITU), by the end of 2019, 53.6% of the world's population utilizes the Internet, which implies around 4.1 billion peoples have access to these technologies, as well as with distinct

mechanisms accessed on the Internet [1]. Even though in many situations has only been manipulated or content shared is original for entertainment purposes only, in another case the manipulation might be intended for falsehood purposes, using forensic and political consequences, for example, utilizing the false contents as digital proof in criminal investigations. Video/Image manipulation represents few actions that are accomplished on the digital content via software editing tools (e.g., GIMP, PIXLR, Adobe Photoshop) or artificial intelligence. Especially, the

copy-move techniques copy a portion of the image and paste it onto similar images [2]. Since editing tools advance, the quality of false images rises and it seems to be original images. Furthermore, postprocessing manipulations, such as brightness equalization/changes and JPEG compression, might decrease the traces left by manipulation and make it very complex to identify [3]. The copy-move forgery detection (CMFD) consists of deep learning- and hand-crafted-based approaches. The previous one is largely separated into hybrid, block, and key point-based methods and next employs convention framework from fine-tuned/scratch algorithms.

Block-based methods utilized distinct kinds of feature extraction, for example, Tetrolet transforms/Fourier transform, and DCT (discrete cosine transform). The major concern is the performance reduction while the copied objects are resized/rotated since the recognition of forging can be performed by a matching procedure [4]. Conversely, key point-based methods such as SURF (Speed-Up Robust Features) and SIFT (scale-invariant feature transform) are very stronger to lighting and rotation differences; however, they have many problems to conquer, for example, natural duplicate objects spotted as false duplicate objects and reliance on original key point in an image, and detect forgeries in the area of uniform intensity [5]. A hybrid method provides constant results by means of F1-score, precision (P), and recall (R) for an individual dataset.

1.1. Motivation. There is a current development of deviating traditionally handcrafted feature extraction for employing convolutional neural network (CNN)-based extractor. But, in few conventional CNN-based forensic detectors is usually not real world for several details, for example, by means of strength in feature extraction and solution of tampering position. Thus, there are various attempts to develop a preprocessing layer for enhancing the strength of feature extraction [6] and combine several detector-based likelihood maps and individual CNN-based consistency maps for improving the solution of tampering location. But still, they endure numerous limits in the abovementioned methods. Initially, current pixel-wise tampering detector adapts an autonomous patch-based approach instead of utilizing the related data amongst patches [7]. Moreover, the lack of statistical features on flat regions (blue ocean, clear sky, and so on) leads to uncertainty approximation and degradation of recognition accuracy. In this situation, the texture of an image content has become a decisive factor to enhance recognition performance. In addition, with the quick growth of image-editing software, the remainder left by the manipulation process has behavior like its pristine versions (viz., tampering trace is difficult to identify) [8]. Then, decreasing the possibility of recognition mismatch and enhancing the solution of localization (managed by the small units of finding) still remain an open challenge.

1.2. Scope of the Research Work. This article presents an automated DL-based fusion model for copy-move forgery detection and localization (DLFM-CMDFC). The proposed

DLFM-CMDFC technique comprises the fusion of generative adversarial network (GAN) and densely connected network (DenseNet) models.

2. Related Works

Yao et al. [9] develop efficient detectors, which can complete image fake localization and detection. Particularly, based on the developed continuous high-pass filter, they initially determine an effective CNN framework automatically for and adaptively extracting features and propose an RFM model for improving tamper recognition performance and localization solution. Abdalla et al. [10] examine copy-move counterfeit findings with a fusion processing method including an adversarial method and deep convolution method. Four databases were employed. The result indicates a considerably higher recognition accuracy (~95%) shown by the discriminator counterfeit detector and DL-CNN models. Accordingly, an end-to-end trained DNN method for counterfeit finding seems to be an optimum approach.

Diallo et al. [11] introduce an architecture enhancing strength for image counterfeit recognition. The vital stage of this architecture is to consider the image quality matching to the selected application. Consequently, it is based on a camera recognition method-based CNN model. Lossy compressions like JPEG are taken into account as general kind of inadvertent/intentional concealment of image counterfeit, which results in manipulation. Consequently, the trainable CNN is fed into a combination of distinct amounts of uncompressed and compressed images. Rodriguez-Ortega et al. [12] present 2 methods, which utilize the DL method, an approach with a convention framework, and a method with the TL model. In all the cases, the effect of depth of the network can be examined by means of F1-score, precision (P), and recall (R). In addition, the challenge of generalization can be resolved from 8 distinct open-access databases.

In the study by Doegar et al. [13], CNN-based pretrained AlexNet method deep feature was employed, which is effective and efficient than that of current advanced methods on open-source standard database MICC-F220. Marra et al. [14] introduce a CNN-based image counterfeit recognition architecture that makes decisions according to the full resolution data collected from the entire image. Because of gradient checkpointing, the architecture can be trained end to end using constrained memory resources and weak (image-level) supervision, which enables the joint optimization of each parameter.

Dixit and Bag [15] presented a technique where SWT and spatial-limited edge-preserving watershed segmentation are employed on input images in the preprocessing phase. Descriptor computation and key point extraction were implemented. Outlier removal can be executed by the RANSAC approach. Furthermore, counterfeit areas are positioned by relation map generation. In Bi et al. [16], a counterfeit localization generator GM has been presented on the basis of a multidecoder single task method. Through adversarial training 2 generators, the presented alpha-learnable WCT blocks in GT suppress manually the

tampering artifact in the counterfeit images. In the meantime, the localization and detection capacities of GM would be enhanced by learning the phony images restored by GT.

Ghai et al. [17] aim at designing a DL-based image counterfeit recognition architecture. The presented model focuses on detecting images counterfeit with splicing and copy-move methods. The image conversion method supports the detection of related features to the network for training efficiently. Next, the pretrained personalized CNN is utilized for training the public standard databases. In Rao et al. [18], a new image counterfeit localization and detection system has been presented on the basis of the DCNN model that integrates a multisemantic CRF-based attention method. The presented model depends on the main findings that the boundary transition artifact arising from the blending operation is global in several image counterfeit manipulations, that is, established in this model using a method with CRF-based attention method through making attention mapping to characterize the possibility of being counterfeit for all the pixels in an image.

3. The Proposed Model

In this study, an efficient DLFM-CMDFC technique is presented for automated copy-move forgery detection and localization model. The proposed DLFM-CMDFC technique encompasses the fusion of GAN and DenseNet models. In DLFM-CMDFC technique, the two outcomes are combined into a layer to define the input vectors with the initial layer of the ELM classifier. Moreover, the optimal parameter tuning of the ELM model takes place by the use of AFSA. The outcomes of the networks are fed as input to the merger unit. Lastly, the difference between the input and targets areas is identified in a forged image.

3.1. GAN-Based Forgery Image Generation. Advancements of technology are assisting GAN to generate forged images, which fool even the more advanced detector [58]. It must be noted that the main objective of generative adversarial network is to create images that could not be differentiated from the primary source image. As demonstrated, generator G_A was applied for transforming input images A from domain D_A to output domain D_B . Then, generator G_B can be utilized for mapping image B back to domain D_A (the original domain). Thereby, another set of cycle consistency losses are included in the standard adversarial losses borne by the discriminator, therefore, attaining $A = G(G(A))$ and assisting the 2 images to be coupled. Highly advanced editing tools are needed for changing an image context. This tool should be capable of altering images when preserving the original source perspectives, shadowing, etc. Those without forgery detection training will not able to differentiate the actual image from an image forged utilizing this methodology that implies that it is the best candidate to develop support material for false news reports.

GAN task is given in the following: (1) build a discriminator network; (2) load a dataset; (3) generate a sample image; (4) build a generator network; (5) closing thoughts; (7) training difficulties. The GAN network branch is shown in Figure 1 [19].

In the presented GAN network, it is considered 2 major phases: (1) in the initial phase, the generator fashions an image from haphazard noise input, and (2) then, the image, as well as various images based on a similar database, is proposed for the discriminator. (3) After the discriminator is proposed by the real and forged images, it provides likelihoods through numbers in the range of zero and one, extensive. Now, zero denotes a forged image and one represents a higher probability for validity. It should be noted that the discriminator must be pretrained previous to the generator since it generates clear gradients. Retaining the constant values enables the network to possess a good understanding of the gradients, that is, the foundation of its learning. But GAN has been proposed as a kind of game performed among opposite networks, and retaining their balance could be problematic. Inopportunately, learning is hard for GAN when the generator/discriminator is highly proficient since GAN usually needs extensive training time. Thus, for example, a GAN can take a long time for an individual GPU, whereas for an individual CPU, a GAN might need few more days.

3.2. DenseNet Model. In this study, the DenseNet-121 framework is utilized as the foundation. In addition, the transfer learning method has been employed in the DenseNet architecture for enhancing the system performance [20]. DenseNets in contrast to common belief require fewer parameters when compared to traditional CNN models since they do not want to learn unnecessary feature maps. The basic idea of the DenseNet architecture is the feature reuse that leads to tremendously compact version. Consequently, it requires fewer parameters when compared to another CNN model because no feature map is repeated. Once CNN goes further, it faces challenges. DenseNet makes this connectivity much easier by simply interconnecting all the layers straightforwardly with every layer. DenseNets utilize the network's capability by reutilizing features. All the layers in DenseNet obtain further input over every prior layer and transmit its feature map to the succeeding layers.

All the layers receive good understanding from the above layers, namely, the idea of concatenation that is utilized. For maximizing computational recycling among the classifiers, incorporating several classifiers to a model and DCNN and interconnect with dense connectivity for effective image classification [21]. A study has proved that a convolution network with smaller connections among layers and those nearer to the output could be very much deeper, and it would be more precise for training. DenseNet attains important developments over the advanced technology when consuming minimum memory and processing to improve its efficiency. The DL library PyTorch and torchvision are utilized, that is, a pretrained data learning method that contains a maximal control across overfitting and also improves the optimization of results from the very first. It consists of 1 classification layer (16), 2 DenseBlocks (1 × 1 and 3 × 3 convs), 3 transition layers (6, 12, and 24), and 5 convolution and pooling layers.

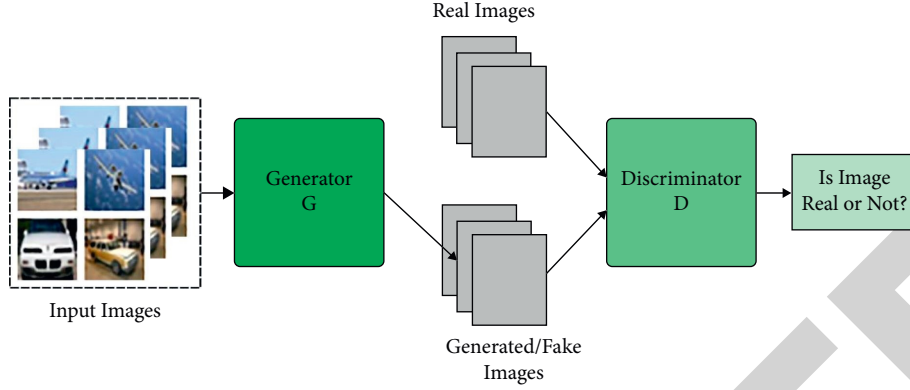


FIGURE 1: Framework of GAN.

3.3. *Optimal ELM Model Using AFSA.* ELM is essentially an SLFN algorithm. The variance among ELM and SLFN exists within the weight of the output layer, and hidden layer neurons are upgraded. In SLFN, the weight of input and output layers is initiated arbitrarily, and the weight of the layers is upgraded using the BP model. In ELM, the weight of the hidden layer is allocated arbitrarily but not upgraded, and the weight of the output layer is upgraded at the time of training. Since in ELM, the weight of single layer is upgraded against both layers of SLFN, it would make ELM quicker when compared to SLFN.

Assume the trained database as (x_j, t_j) in which $x_j = [x_{j1}, x_{j2}, \dots, x_{jN}]^T$ represents the input vector and t_j denotes the output vector. The output of j^{th} hidden layer neuron is represented as $g(w_i, b_i, x_j)$, in which w_i indicates the weight vector connected the input neuron to i^{th} hidden layer neuron, b_i signifies the bias of i^{th} hidden neurons, and g denotes the activation function. All the hidden layer neurons of ELM are interconnected to all the output layer neurons with related weight, and they represent the weight interconnecting the i^{th} hidden layer neuron with output neuron as β_i . This framework is denoted arithmetically by

$$\sum_{i=1}^L \beta_i g(w_i, b_i, x_j) = t_j, \quad (1)$$

where L represents the number of hidden neurons, and j indicates the output or input sample of overall N trained samples. The aforementioned formula is expressed by

$$H\beta = T. \quad (2)$$

In the above formula, consider m output node as

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m} \quad \text{and} \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}, \quad (3)$$

where H denotes the output matrix of the hidden layer, which is given as

$$H = \begin{bmatrix} g(w_1, b_1, x_1) & \cdots & g(w_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ g(w_1, b_1, x_N) & \cdots & g(w_L, b_L, x_N) \end{bmatrix}. \quad (4)$$

The minimum norm least square of (2) is

$$\hat{\beta} = H^+ T, \quad (5)$$

where H^+ is the Moore–Penrose generalized inverse of matrix H . H^+ is evaluated by singular value decomposition (SVD), QR approach, orthogonal projection model [22], and orthogonalization method.

It must standardize the scheme (to avoid overfitting), and the optimization issues turn into

$$\min \left(\|\beta\|^2 + C \sum_{i=1}^N \|\xi_i^T\|^2 \right), \quad (6)$$

where $\xi_i = t_i^T - h(x_i)\beta$ denotes the trained error of i^{th} instance and C denotes the appropriate penalty factor. It might convert these problems to its dual form and create the Lagrangian function as

$$F = \|\beta\|^2 + C \sum_{i=1}^N \|\xi_i\|^2 - \sum_{i=1}^N \sum_{j=1}^L \alpha_{ij} (h(x_i)\beta_j - t_{ij} + \xi_{ij}). \quad (7)$$

Take the partial derivative of the aforementioned formula and apply KKT condition. When $L < N$, the size of matrix $H^T H$ is lesser when compared to matrix HH^T ,

$$\beta = H^+ T = \left(\frac{I}{C} + H^T H \right)^{-1} H^T T. \quad (8)$$

Hence, the last output of ELM is

$$f(x) = h(x)\beta = h(x) \left(\frac{I}{C} + H^T H \right)^{-1} H^T T. \quad (9)$$

Once $L > N$, the size of matrix HH^T is lesser when compared to the matrix $H^T H$, the solution of the equation becomes

$$\beta = H^+T = H^T \left(\frac{I}{C} + HH^T \right)^{-1} T. \quad (10)$$

Thus, the last output of ELM is

$$f(x) = h(x)\beta = h(x)H^T \left(\frac{I}{C} + HH^T \right)^{-1} T. \quad (11)$$

For the binary classification problems, the decision function of ELM can be expressed by

$$f(x) = \text{sign}(h(x)\beta). \quad (12)$$

For multiclass instance, the class label of instance is expressed by

$$\text{label}(x) = \arg \max_{1 \leq i \leq m} \{f_i(x)\}. \quad (13)$$

Then

$$f(x) = [r_1(x), f_2(x), f_3(x), \dots, f_n(x)]^T. \quad (14)$$

ELM was employed for the classification and prediction tasks in various fields. To optimally adjust the learning rate of the ELM model, the AFSA is used, which is a kind of swarm intelligence method depending on the behavior of the animal. It was developed by Li et al. in 2002 [23]. Its fundamental is the inspiration of collision, foraging, and clustering behavior of fish and the collective support in a fish swarm for realizing a global optimum points. The highest distance pass through in the artificial fish method can be determined by *Step*, the apparent distance pass through by the artificial fish can be determined by *Visual*, the retry amount represent the *Try_Number* also the factors of crowd amount represent η . The location of a single artificial fish is defined by the resulting vectors $X = (X_1, X_2, \dots, X_n)$, and the distance among artificial fish i and j denotes $d_{ij} = |X_i - X_j|$. The behavior function for the artificial fish can be determined by random, prey, swarm, and follow.

Assume that the fish observe their food using their eyes and the present location is X_i , as well as an arbitrarily elected location is X_j within their perceptive range:

$$X_j = X_i + \text{Visual} \times \text{rand}(0 \sim 1), \quad (15)$$

where *rand* (0-1) represents an arbitrary value between zero and one. When $Y_i > Y_j$, the fish move in this direction. Or else, the method arbitrarily selects a novel location X_j for judging whether it fulfills the moving criteria. When it performs,

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\|X_j - X_i^t\|} \times \text{Step} \times \text{rand}(0 \sim 1). \quad (16)$$

When it does not *Try_Number* times, an arbitrary movement can be generated by

$$X_i^{t+1} = X_i^t + \text{Visual} \times \text{rand}(0 \sim 1). \quad (17)$$

In order to prevent overcrowding, an artificial present location X_i is fixed. Next, the amount of fish in its n_f company and X_c center in the region (i.e., $d_{ij} < \text{Visual}$) are

defined. When $Y_c/n_f < \eta \times Y_i$, the position of companion represents the optimal number of food and lower crowding. Subsequently, the fish moves to its companion region center position:

$$X_i^{t+1} = X_i^t + \frac{X_c - X_i^t}{\|X_c - X_i^t\|} \times \text{Step} \times \text{rand}(0 \sim 1). \quad (18)$$

Or else it starts to perform the behavior of prey.

The present location of artificial fish swarm can be determined by X_i . The swarm defines its main company Y_j as X_j in the region (i.e., $d_{ij} < \text{Visual}$). When $Y_j/n_f < \eta \times Y_i$, the position of companies represents the optimal number of food and lesser crowd [24]. Next, the swarm moves to X_j :

$$X_i^{t+1} = X_i^t + \frac{X_j - X_i^t}{\|X_j - X_i^t\|} \times \text{Step} \times \text{rand}(0 \sim 1). \quad (19)$$

It enables artificial fish to attain company and food through a large regional area. A location is arbitrarily chosen, as well as artificial fish moves to it. Figure 2 illustrates the flowchart of AFSA.

With the searching space of D dimensional, highly probable distance amid 2 artificial fishes is utilized for vigorously limiting the *Visual* & *Step* of an artificial fish. It is determined by *MaxD*:

$$\text{MaxD} = \sqrt{(x_{\max} - x_{\min})^2 \times D}, \quad (20)$$

where x_{\min} and x_{\max} represent the lower and upper bounds of the optimization range, respectively, and D indicates the dimension of the search space.

4. Experimental Validation

This section investigates the result analysis of the proposed model on MNIST and COCO datasets. Figure 3 shows a few sample image, tampered image, and localization image.

Table 1 and Figure 4 provide the performance analysis of the proposed model on the applied MNIST dataset under varying runs. The results demonstrated that the proposed model has gained effective outcomes under distinct runs. For instance, under run-1, the proposed model has attained effective outcome with the prec_n of 96.38%, rec_l of 93.71%, acc_y of 94.29%, and F_{score} of 95.98%. Also, under run-3, the presented manner has reached effective outcome with the prec_n of 93.54%, rec_l of 97.30%, acc_y of 94.88%, and F_{score} of 97.19%. Besides, under run-5, the presented technique has obtained effective outcome with the prec_n of 96.80%, acc_y of 97.43%, acc_y of 96.87%, and F_{score} of 94.69%.

Figure 5 demonstrates the ROC analysis of the DLFM-CMDFC technique on the test MNIST dataset. The figure has shown that the DLFM-CMDFC technique has resulted in an effective outcome with a maximum ROC of 98.5180.

Figure 6 portrays the accuracy analysis of the DLFM-CMDFC technique on the test MNIST dataset. The results demonstrated that the DLFM-CMDFC technique has accomplished improved performance with increased training and validation accuracy. It is noticed that the DLFM-

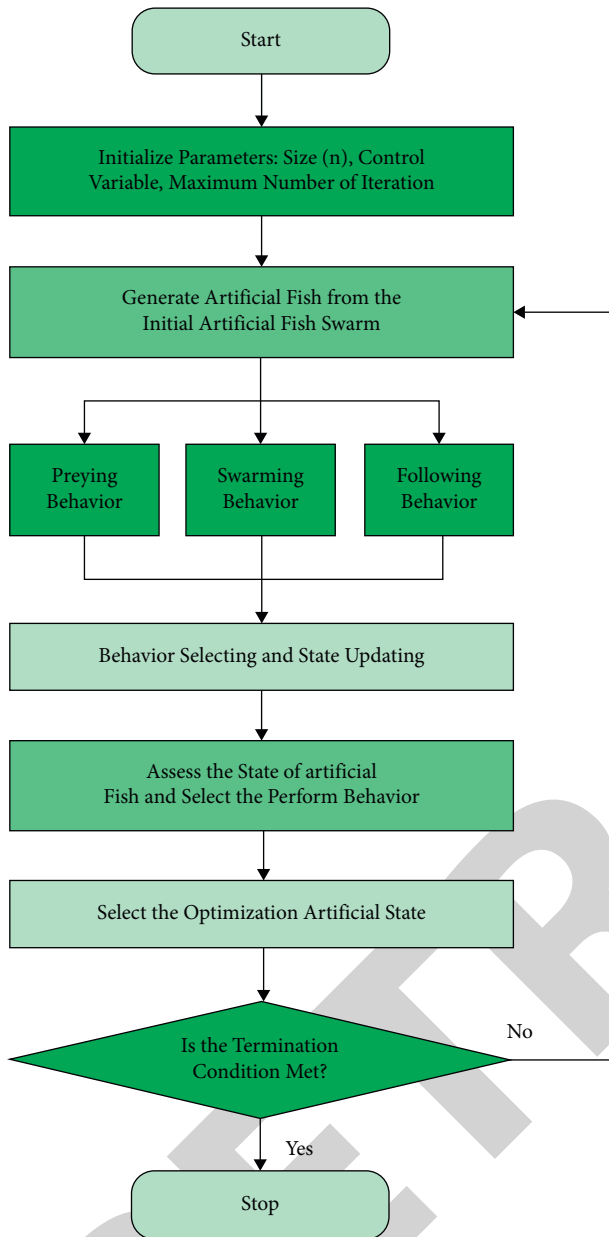


FIGURE 2: Flowchart of AFSA.

CMDFC technique has gained improved validation accuracy over the training accuracy. Similarly, Figure 7 depicts the loss analysis of the DLFM-CMDFC technique on the test MNIST dataset. The results established that the DLFM-CMDFC technique has resulted in a proficient outcome with reduced training and validation loss. It is observed that the DLFM-CMDFC technique has offered reduced validation loss over the training loss.

Table 2 and Figure 8 offer the performance analysis of the presented technique on the applied CIFAR-10 dataset under varying runs. The outcomes exhibited that the presented approach has reached effectual outcomes under different runs. For instance, under run-1, the presented manner has attained

effective outcome with the $prec_n$ of 96.52%, rec_l of 96.15%, acc_y of 96.36%, and F_{score} of 96.66%. Followed by, under run-3, the proposed model has attained effective outcome with the $prec_n$ of 97.95%, rec_l of 96.68%, acc_y of 97%, and F_{score} of 96.57%. In addition, under run-5, the projected system has achieved effective outcome with the $prec_n$ of 97.46%, rec_l of 96.50%, acc_y of 97.35%, and F_{score} of 94.52%.

Figure 9 depicts the ROC analysis of the DLFM-CMDFC technique on the test CIFAR-10 dataset. The figure outperformed that the DLFM-CMDFC scheme has resulted in an effective outcome with the maximal ROC of 98.7262.

Figure 10 demonstrates the accuracy analysis of the DLFM-CMDFC technique on the test CIFAR-10 dataset. The outcomes showcased that the DLFM-CMDFC technique has accomplished improved efficiency with increased training and validation accuracy. It can be noticed that the DLFM-CMDFC manner has gained increased validation accuracy over the training accuracy.

Figure 11 represents the loss analysis of the DLFM-CMDFC manner on the test CIFAR-10 dataset. The outcomes recognized that the DLFM-CMDFC approach has resulted in a proficient outcome with the decreased training and validation loss. It can be stated that the DLFM-CMDFC technique has obtainable minimum validation loss over the training loss.

The $prec_n$ analysis of the DLFM-CMDFC technique with existing ones on the test dataset is given in Table 3.

Figure 12 illustrates the $prec_n$ analysis of the DLFM-CMDFC technique with existing ones. The figure has shown that the IFD-AOS-FPM and CMFD-BMIF techniques have obtained reduced $prec_n$ of 53.90% and 54.40%. At the same time, the CMFD and BB-KB-ICMFD techniques have resulted in moderate $prec_n$ of 57.34% and 56.62%, respectively. Moreover, the CMFD-GAN-CNN technique has accomplished near optimal $prec_n$ of 69.64%. However, the DLFM-CMDFC technique has resulted in superior performance with the $prec_n$ of 97.27%.

Figure 13 illustrates the rec_{ll} analysis of the DLFM-CMDFC approach with current ones. The figure exhibited that the CMFD and CMFD-BMIF algorithms have obtained reduced rec_{ll} of 49.39% and 80.20%, respectively. Concurrently, the CMFD-GAN-CNN and BB-KB-ICMFD techniques have resulted in a moderate rec_{ll} of 80.42% and 80.40%, respectively. In addition, the IFD-AOS-FPM system has accomplished near optimal rec_{ll} of 83.27%. But, the DLFM-CMDFC technique has resulted in a maximal performance with the rec_{ll} of 96.46%.

Figure 14 depicts the F_{score} analysis of the DLFM-CMDFC system with present ones. The figure portrayed that the IFD-AOS-FPM and CMFD techniques have obtained reduced F_{score} of 54.39% and 49.26, respectively. Simultaneously, the CMFD-BMIF and BB-KB-ICMFD techniques have resulted in a moderate F_{score} of 59.43% and 60.55%, respectively. Also, the CMFD-GAN-CNN algorithm has accomplished near optimal F_{score} of 88.35%. Eventually, the DLFM-CMDFC manner has resulted in increased efficiency with the F_{score} of 96.06%.

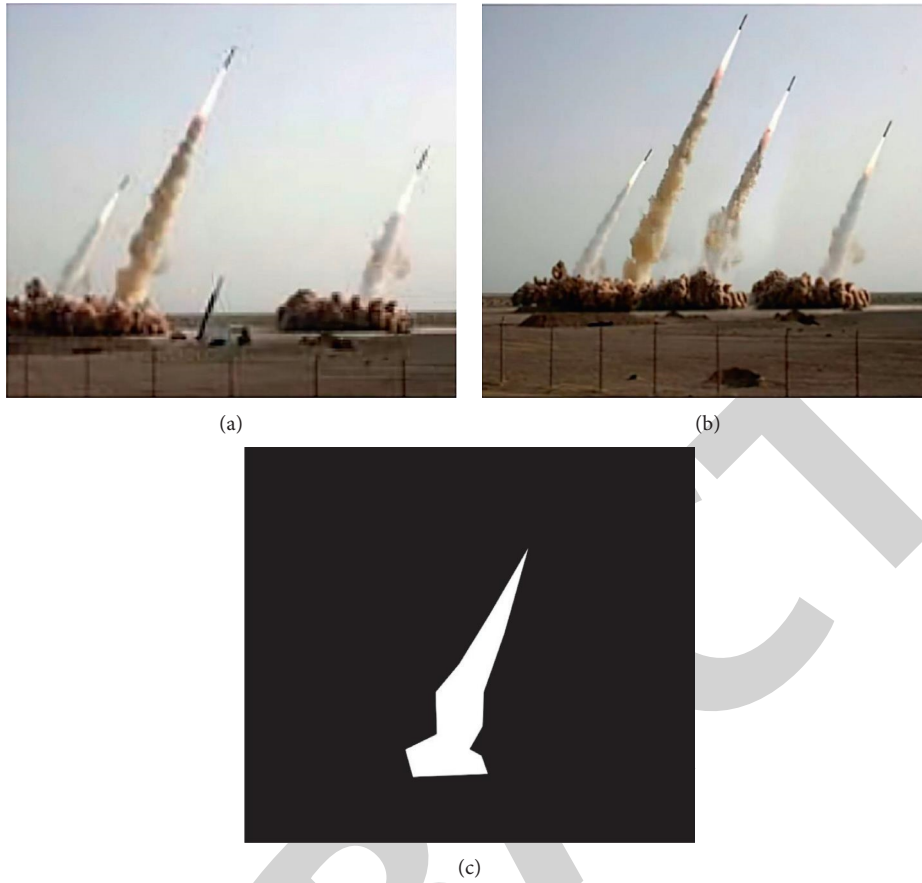


FIGURE 3: Sample image forgery detection results. (a) Original image. (b) Tampered image. (c) Localization image.

TABLE 1: Result analysis of DLFM-CMDFC model on MNIST dataset.

No. of runs	Precision	Recall	Accuracy	F-score
Run-1	96.38	93.71	94.29	95.98
Run-2	93.83	95.51	94.61	93.93
Run-3	93.54	97.30	94.88	97.19
Run-4	96.54	95.52	96.45	97.32
Run-5	96.80	97.43	96.87	94.69
Average	95.42	95.89	95.42	95.82

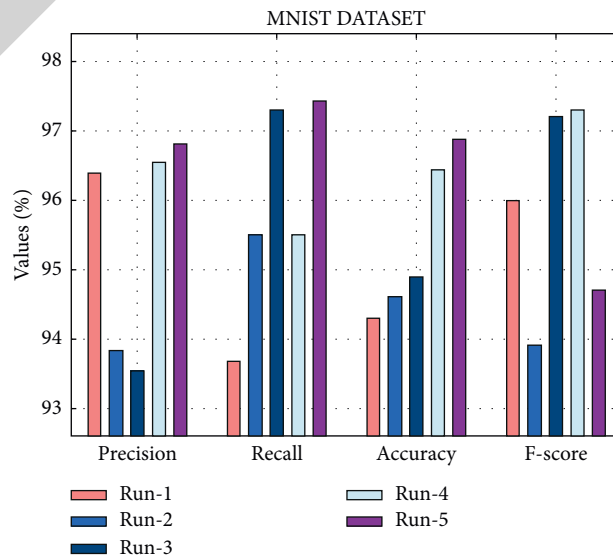


FIGURE 4: Result analysis of DLFM-CMDFC model on MNIST dataset.

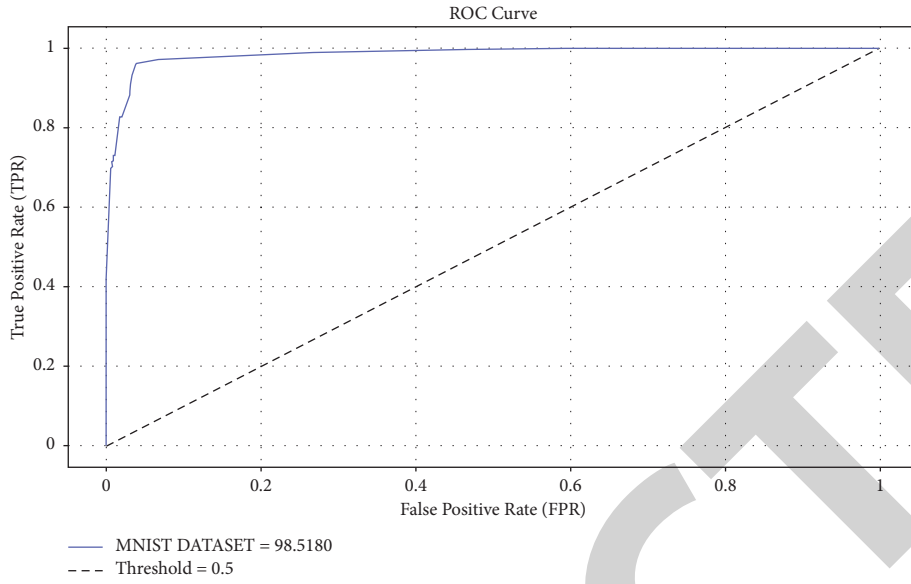


FIGURE 5: ROC analysis of the DLFM-CMDFC model on the MNIST dataset.

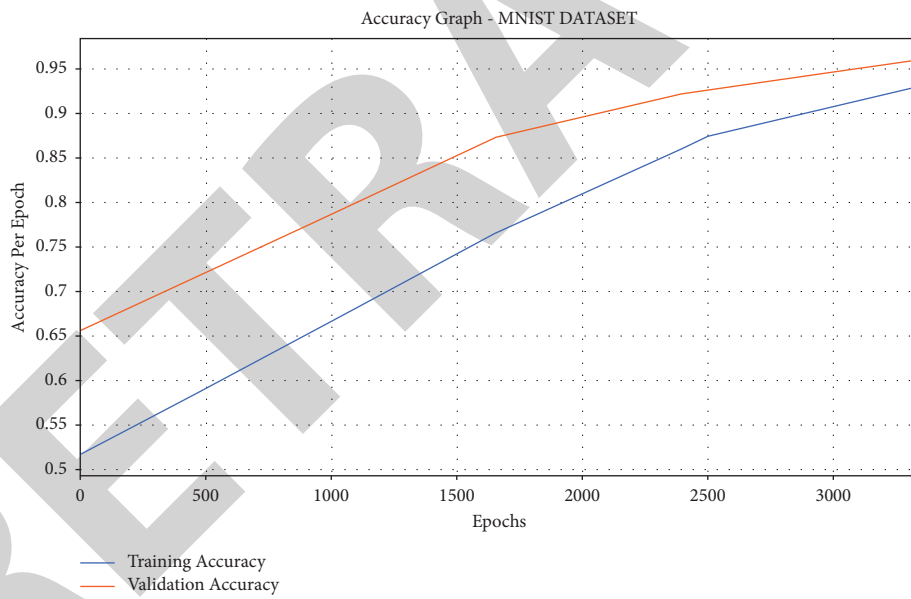


FIGURE 6: Accuracy analysis of the DLFM-CMDFC model on the MNIST dataset.

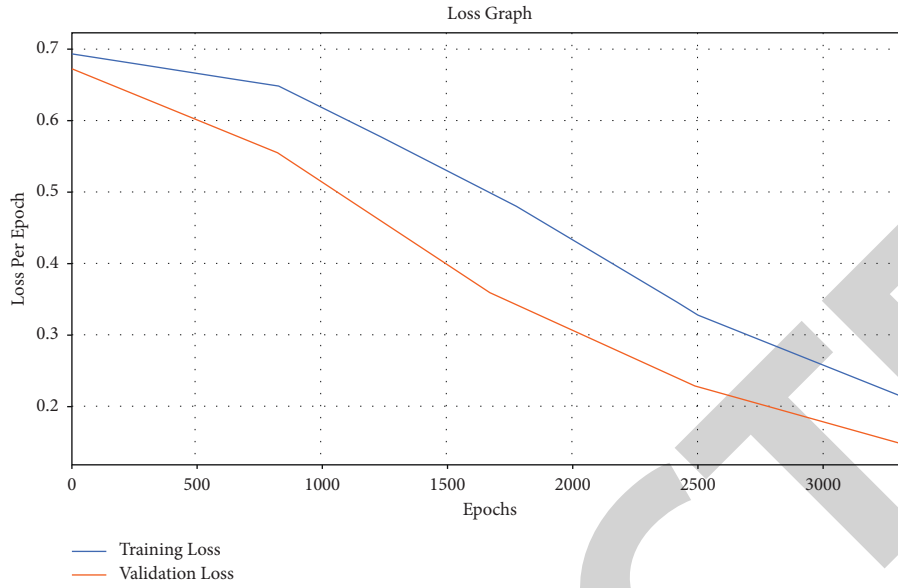


FIGURE 7: Loss analysis of the DLFM-CMDFC model on the MNIST dataset.

TABLE 2: Result analysis of the DLFM-CMDFC model on the CIFAR-10.

No. of runs	Precision	Recall	Accuracy	F-score
Run-1	96.52	96.15	96.36	96.66
Run-2	95.75	97.45	96.90	94.77
Run-3	97.98	96.68	97.00	96.57
Run-4	97.51	95.93	97.02	93.50
Run-5	97.46	96.50	97.35	94.52
Run-6	97.78	96.70	97.20	97.23
Run-7	97.71	96.03	97.22	96.86
Run-8	96.98	96.68	96.00	96.82
Run-9	97.31	95.73	96.82	96.51
Run-10	97.66	96.70	97.55	97.17
Average	97.27	96.46	96.94	96.06

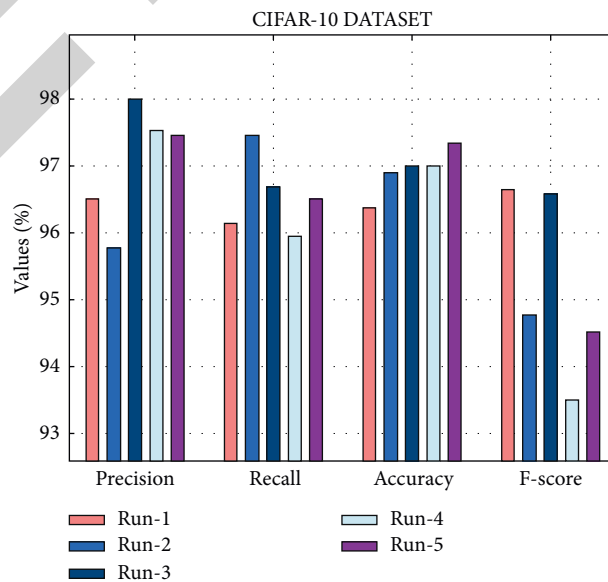


FIGURE 8: Result analysis of the CIFAR-10 model on the DLFM-CMDFC dataset.

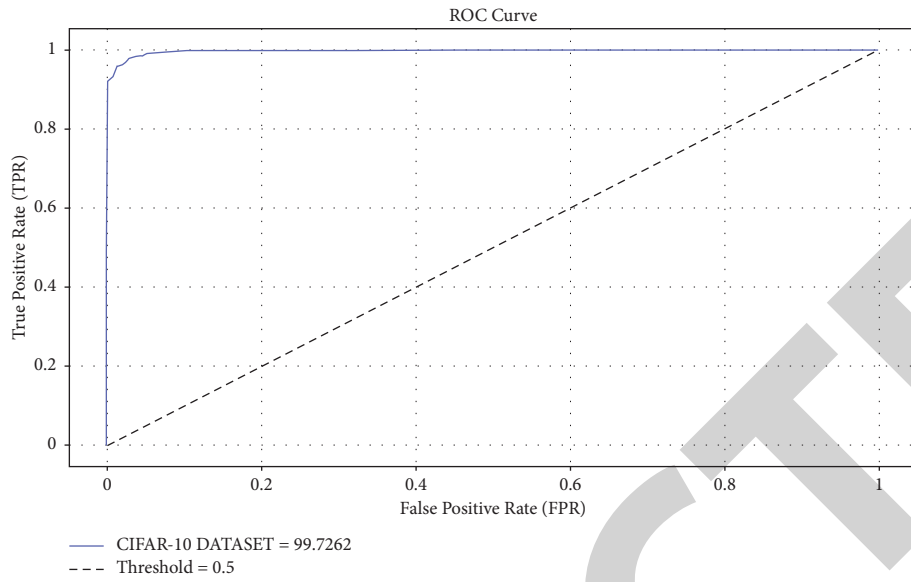


FIGURE 9: ROC analysis of the DLFM-CMDFC model on the CIFAR-10 dataset.

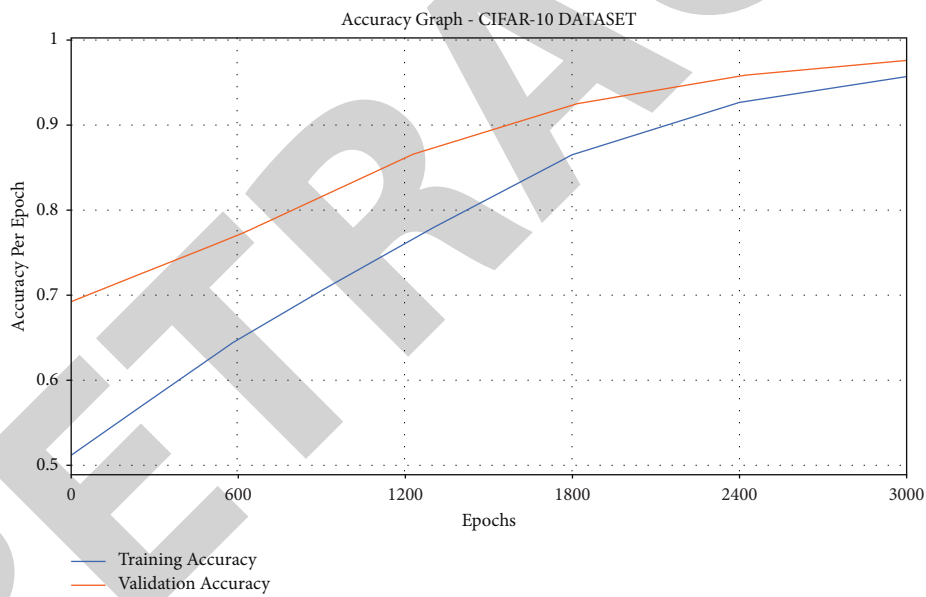


FIGURE 10: Accuracy analysis of the DLFM-CMDFC model on the CIFAR-10 dataset.

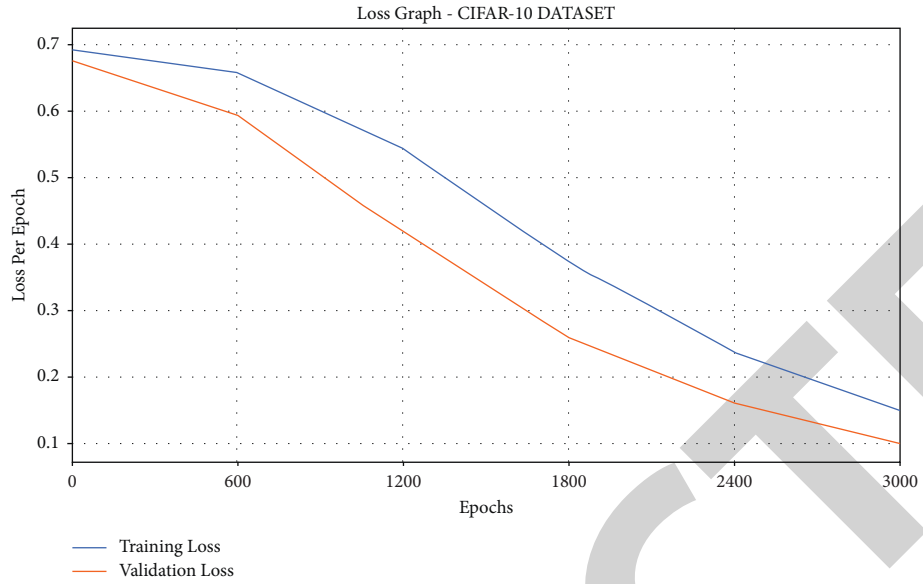


FIGURE 11: Accuracy analysis of THE DLFM-CMDFC model on THE CIFAR-10 dataset.

TABLE 3: Comparative analysis of DLFM-CMDFC mode with existing techniques.

Methods	Precision	Recall	F-score
CMFD	57.34	49.39	49.26
IFD-AOS-FPM	53.90	83.27	54.39
CMFD-BMIF	54.40	80.20	59.43
BB-KB-ICMFD	56.62	80.40	60.55
CMFD-GAN-CNN	69.63	80.42	88.35
DLFM-CMDFC	97.27	96.46	96.06

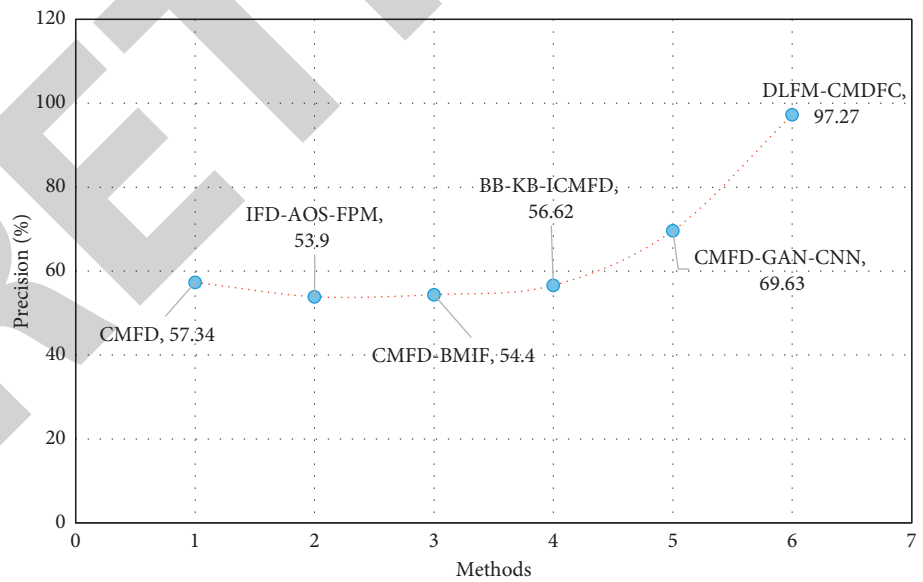


FIGURE 12: Precision analysis of DLFM-CMDFC technique with existing manners.

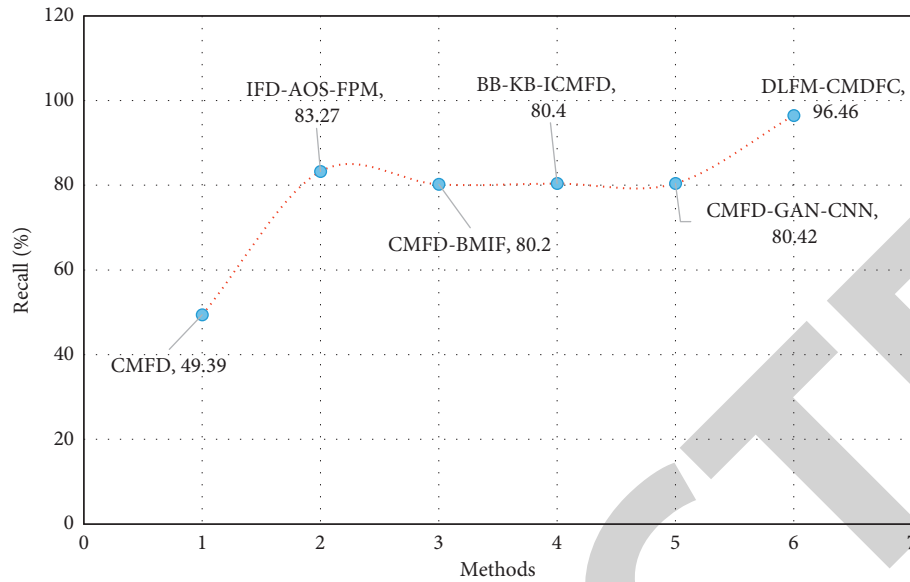


FIGURE 13: Recall analysis of DLFM-CMDFC technique with existing manners.

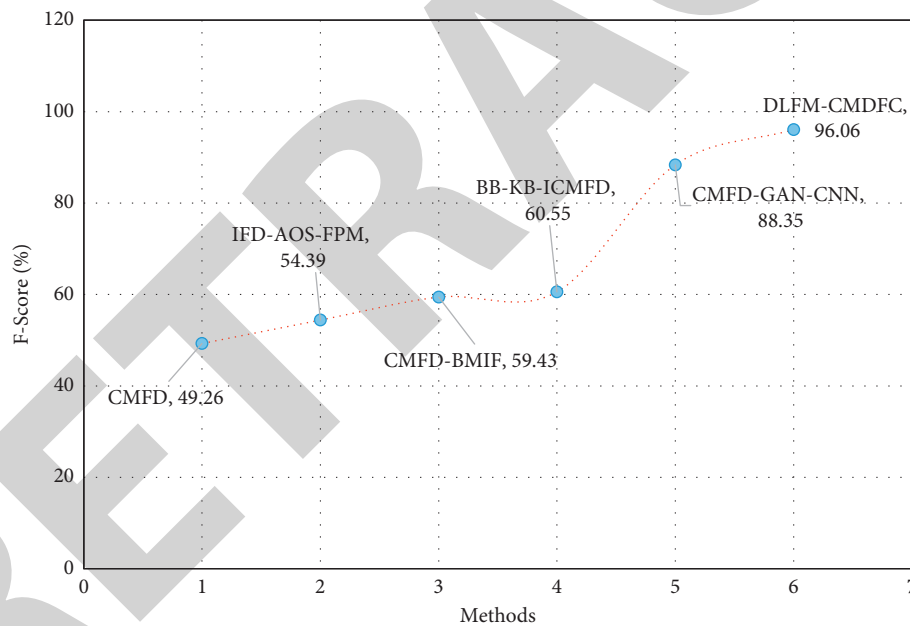


FIGURE 14: F-Score analysis of DLFM-CMDFC model with existing manners.

5. Conclusion

This article has presented an automated copy-move forgery detection and localization model, named DLFM-CMDFC. The proposed DLFM-CMDFC technique encompasses the fusion of GAN and DenseNet models. In DLFM-CMDFC technique, the two outcomes are combined into a layer to define the input vectors with the initial layer of the ELM classifier. Moreover, the optimal parameter tuning of the ELM technique takes place by the use of AFSA. The outcomes of the networks are fed as

input to the merger unit. Lastly, the difference between the input and targets areas is identified in a forged image. The performance validation of the proposed manner takes place using two benchmark datasets. The proposed research work outperforms with 97.27% of precision, 96.46% of recall, and 96.06% of F-score. The experimental outcomes pointed out the supremacy of the proposed technique on the recently developed approaches. As a part of future scope, the detection performance can be improved by the use of generative adversarial network (GAN) model.

Retraction

Retracted: Applying Dynamic Systems to Social Media by Using Controlling Stability

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Abdullah Hamad, M. L. Thivagar, M. Bader Alazzam, F. Alassery, F. Hajje, and A. A. Shihab, "Applying Dynamic Systems to Social Media by Using Controlling Stability," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4569879, 7 pages, 2022.

Research Article

Applying Dynamic Systems to Social Media by Using Controlling Stability

Abdulsattar Abdullah Hamad,¹ M. Lellis Thivagar ,¹ Malik Bader Alazzam ,² Fawaz Alassery ,³ Fahima Hajje ,⁴ and Ali A. Shihab⁵

¹School of Mathematics Madurai Kamaraj University, Madurai, India

²Faculty of Computer Science and Informatics, Amman Arab University, Amman, Jordan

³Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁴Department of Information Systems, College of Computer and Information Sciences,

Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

⁵Mathematics Department, Education for Pure Sciences College, Tikrit University, Tikrit, Iraq

Correspondence should be addressed to Malik Bader Alazzam; m.alazzam@aau.edu.jo

Received 30 October 2021; Revised 9 December 2021; Accepted 15 December 2021; Published 31 January 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Abdulsattar Abdullah Hamad 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 focuses on hybrid synchronization, a new synchronization phenomenon in which one element of the system is synced with another part of the system that is not allowing full synchronization and nonsynchronization to coexist in the system. When $\lim_{t \rightarrow \infty} Y - \alpha X = 0$, where Y and X are the state vectors of the drive and response systems, respectively, and α ($\alpha = \mp 1$), the two systems' hybrid synchronization phenomena are realized mathematically. Nonlinear control is used to create four alternative error stabilization controllers that are based on two basic tools: Lyapunov stability theory and the linearization approach.

1. Introduction

Alazzam et al.' study [1] is an example. Control and hybrid three-dimensional synchronization (HPS) procedures are for a unique hyperchaotic system. To begin, the revolutionary hyperchaotic system is regulated to an unstable equilibrium position or limit cycle using only one scalar controller with two state variables. Using Lyapunov's direct approach, the HPS between two new hyperchaotic systems is studied. A nonlinear feedback vector controller is presented to establish perfect synchronization between two new hyperchaotic systems, which can then be condensed further into a single scalar controller. Finally, simulation data are supplied to ensure the effectiveness of these strategies.

The proposed approaches have some implications for lowering controller installation costs and complexity. Dynamical systems have received a lot of attention. It is one of

the first attempts in a Lu model, and a new hyperchaotic model with three unstable equilibrium points is disclosed. Despite the fact that the newly built system is basic, it is six-dimensional (6D) and has eighteen terms in 2021. [2]. It presents the new structure of high dimension (6D), novel king of quaternion complete, and has some unusual properties [3–6]. There is another study which introduces another chaotic and hyperchaotic complex nonlinear, and this type has a significant stake in its phase-space behavior [7–9]. It has been organized in the previous time, for example, a 3D auto system, which is not differ-isomorphic with the Lorenz attractor. In the arrangement of values for a parameter k , [10] has proposed another 3D attractor that shows chaotic behavior in distinct respects and not diffeomorphic with Lorenz [11–18]. The first chaotic nonlinear system has been suggested by Lorenz [19–22] in which is a generalization of the Lorenz system. The Lorenz system's messy structure is utilized.

2. Hybrid Synchronization between Two Similar Systems

We have already learned about the dynamic system [18], which is in the following formula:

$$\begin{cases} \dot{x}_1 = a(x_2 - x_1) + x_4, \\ \dot{x}_2 = cx_1 - x_2 - x_1x_3 + x_5, \\ \dot{x}_3 = -bx_3 + x_1x_2, \\ \dot{x}_4 = dx_4 - x_1x_3, \\ \dot{x}_5 = -kx_2, \\ \dot{x}_6 = hx_6 + rx_2, \end{cases} \quad (1)$$

which represents the driving system as $x_1, x_2, x_3, x_4, x_5, x_6$ are the variables representing the system states and that $a, b, c, d, k, h,$ and r are the real positive parameters and their values are 11, $7/3, 27, 2, 7.4, 1,$ and $1,$ respectively.

While, the response system can be written as follows:

$$\begin{cases} \dot{y}_1 = a(y_2 - y_1) + y_4 + u_1, \\ \dot{y}_2 = cy_1 - y_2 - y_1y_3 + y_5 + u_2, \\ \dot{y}_3 = -by_3 + y_1y_2 + u_3, \\ \dot{y}_4 = dy_4 - y_1y_3 + u_4, \\ \dot{y}_5 = -ky_2 + u_5, \\ \dot{y}_6 = hy_6 + ry_2 + u_6. \end{cases} \quad (2)$$

The dynamic error of the hybrid synchronization between the 6D chaos system (1) and system (2) is defined by the following relationship.

$$\begin{aligned} e_i &= y_i - x_i, \quad i = 1, 3, 5 \longrightarrow \lim_{t \rightarrow \infty} e_i = 0, \\ e_j &= y_j - x_j, \quad j = 2, 4, 6 \longrightarrow \lim_{t \rightarrow \infty} e_j = 0. \end{aligned} \quad (3)$$

Thus, the error is calculated for the dynamic system as follows:

$$\begin{cases} \dot{e}_1 = ae_2 - ae_1 + e_4 - 2ax_2 - 2x_4 + u_1, \\ \dot{e}_2 = ce_1 - e_2 + e_5 - y_1e_3 + x_3e_1 - 2y_1x_3 + 2cx_1 + 2x_5 + u_2, \\ \dot{e}_3 = -be_3 + e_1e_2 - x_2e_1 + x_1e_2 - 2x_1x_2 + u_3, \\ \dot{e}_4 = de_4 - y_1e_3 + x_3e_1 - 2y_1x_3 + u_4, \\ \dot{e}_5 = -ke_2 + 2kx_2 + u_5, \\ \dot{e}_6 = he_6 + re_2 + u_6. \end{cases} \quad (4)$$

Using the method of linear approximation to find the characteristic equation and the intrinsic values of the system before the control, which is witnessing a state of instability for the system before the control, this method confirms this thing.

$$\lambda^6 + \frac{31}{3}\lambda^5 - \frac{3069}{15}\lambda^4 + \frac{1558}{15}\lambda^3 + \frac{64004}{15}\lambda^2 - \frac{8496}{5}\lambda - 348 = 0,$$

$$\begin{cases} \lambda_1 = 2, \\ \lambda_2 = 1, \\ \lambda_3 = -\frac{7}{3}, \\ \lambda_4 = 10.9659 - 8.10^{-9}i, \\ \lambda_5 = -12.6916 - 3.92820323010^{-9}i, \\ \lambda_6 = 0.3257 + 8.92820323010^{-9}i. \end{cases} \quad (5)$$

The results of the distinctive equation confirm that the error of the dynamic system is in the position of instability.

Theorem 1. Let U be the controller of the system:

$$\begin{cases} u_1 = 2ax_2 + 2x_4 - ce_2 - x_3e_2 + x_2e_3 - x_3e_4, \\ u_2 = -ae_1 + 2y_1x_3 - 2cx_1 - 2x_5 - x_1e_3 + ke_5 - re_6, \\ u_3 = y_1e_2 - e_1e_2 + 2x_1x_2 + y_1e_4, \\ u_4 = -2de_4 - e_1 + 2y_1x_3, \\ u_5 = -e_2 - 2kx_2 - e_5, \\ u_6 = -2he_6. \end{cases} \quad (6)$$

Thus, the hybrid synchronization between system (1) and system (2) can be observed in two ways:

Proof. After compensating the control (6) in the dynamic system error (4), we get

$$\begin{cases} \dot{e}_1 = ae_2 - ae_1 + e_4 - ce_2 - x_3e_2 + x_2e_3 - x_3e_4, \\ \dot{e}_2 = ce_1 - e_2 + e_5 - y_1e_3 + x_3e_1 - ae_1 - x_1e_3 + ke_5 - re_6, \\ \dot{e}_3 = -be_3 - x_2e_1 + x_1e_2 + y_1e_2 + y_1e_4, \\ \dot{e}_4 = -de_4 - y_1e_3 + x_3e_1 - e_1, \\ \dot{e}_5 = -ke_2 - e_2 - e_5, \\ \dot{e}_6 = re_2 - he_6. \end{cases} \quad (7)$$

The first method is the method of linear approximation:

$$\lambda^6 + \frac{53}{3}\lambda^5 + \frac{12834}{25}\lambda^4 + \frac{286498}{75}\lambda^3 + \frac{286609}{25}\lambda^2 + \frac{373231}{25}\lambda + \frac{169704}{25} = 0,$$

$$\begin{cases} \lambda_1 = -1, \\ \lambda_2 = \frac{7}{3}, \\ \lambda_3 = -2.3511, \\ \lambda_4 = -2.6530, \\ \lambda_5 = -4.4979 + 19.6945i, \\ \lambda_6 = -4.4979 - 19.6945i. \end{cases} \quad (8)$$

The linear approximation method succeeded in systems 1 and 2 showing the hybrid synchronization between the two systems and the Lyapunov Method is failed. We get $\dot{V}(e_i)$ as follows:

And they use the Lebanov method [18–20], as follows. After differentiating the function $V(e_i)$, we get

$$\begin{aligned} \dot{V}(e) &= \mathbf{e}_1 (ae_2 - ae_1 + e_4 - ce_2 - x_3e_2 + x_2e_3 - x_3e_4) \\ &\quad + \mathbf{e}_2 (ce_1 - e_2 + e_5 - y_1e_3 + x_3e_1 - ae_1 - x_1e_3 + ke_5 - re_6) \\ &\quad + \mathbf{e}_3 (-be_3 - x_2e_1 + x_1e_2 + y_1e_2 + y_1e_4) + \mathbf{e}_4 (-de_4 - y_1e_3 + x_3e_1 - e_1) \\ &\quad + \mathbf{e}_5 (-ke_2 - e_2 - e_5) + \mathbf{e}_6 (re_2 - he_6), \\ \dot{V}(e_i) &= -ae_1^2 - e_2^2 - be_3^2 - de_4^2 - e_5^2 - he_6^2 = -e_i^T Q_0 e_i, \\ Q_0 &= \text{diag}(a, 1, b, d, 1, h). \end{aligned} \quad (9)$$

Thus, $Q_0 > 0$, and this leads to $\dot{V}(e_i)$ being negatively defined in \mathbf{R}^6 . Thus, the nonlinear control unit is suitable, in which there is no synchronization.

The initial values (2, 1, 8, 6, 12, 4), (-18, -9, -1, -5, -20, 15) are used to illustrate how the hybrid synchronization occurs between the two systems (1) and (2). Figures 1 and 2 show the verification of these results numerically. \square

Theorem 2. The nonlinear U control of system (4) and Figure 3 is designed as follows:

$$\begin{cases} u_1 = 2ax_2 + 2x_4 - ce_2 - x_3e_2 + x_2e_3 - x_3e_4, \\ u_2 = -ae_1 + 2y_1x_3 - 2cx_1 - 2x_5 - x_1e_3 - e_6, \\ u_3 = y_1e_2 - e_1e_2 + 2x_1x_2 + y_1e_4, \\ u_4 = -2de_4 - e_1 + 2y_1x_3, \\ u_5 = -2kx_2 - e_5, \\ u_6 = -2he_6. \end{cases} \quad (10)$$

The hybrid synchronization between the two systems (1) and (2) can be explained in two ways.

Proof. By relying on control (10) with system (4),

$$\begin{cases} \dot{e}_1 = ae_2 - ae_1 + e_4 - ce_2 - x_3e_2 + x_2e_3 - x_3e_4, \\ \dot{e}_2 = ce_1 - e_2 + e_5 - y_1e_3 + x_3e_1 - ae_1 - x_1e_3 - e_6, \\ \dot{e}_3 = -be_3 - x_2e_1 + x_1e_2 + y_1e_2 + y_1e_4, \\ \dot{e}_4 = -de_4 - y_1e_3 + x_3e_1 - e_1, \\ \dot{e}_5 = -ke_2 - e_5, \\ \dot{e}_6 = -he_6 + re_2. \end{cases} \quad (11)$$

The first method is the method of linear approximation:

$$\lambda^6 + \frac{53}{3}\lambda^5 + \frac{2167}{5}\lambda^4 + \frac{38509}{15}\lambda^3 + \frac{90806}{15}\lambda^2 + \frac{31068}{5}\lambda + \frac{11552}{5} = 0,$$

$$\begin{cases} \lambda_1 = -1, \\ \lambda_2 = \frac{7}{3}, \\ \lambda_3 = -1.2965, \\ \lambda_4 = -1.9544, \\ \lambda_5 = -5.3745 + 17.6928i, \\ \lambda_6 = -5.3745 - 17.6928i. \end{cases} \quad (12)$$

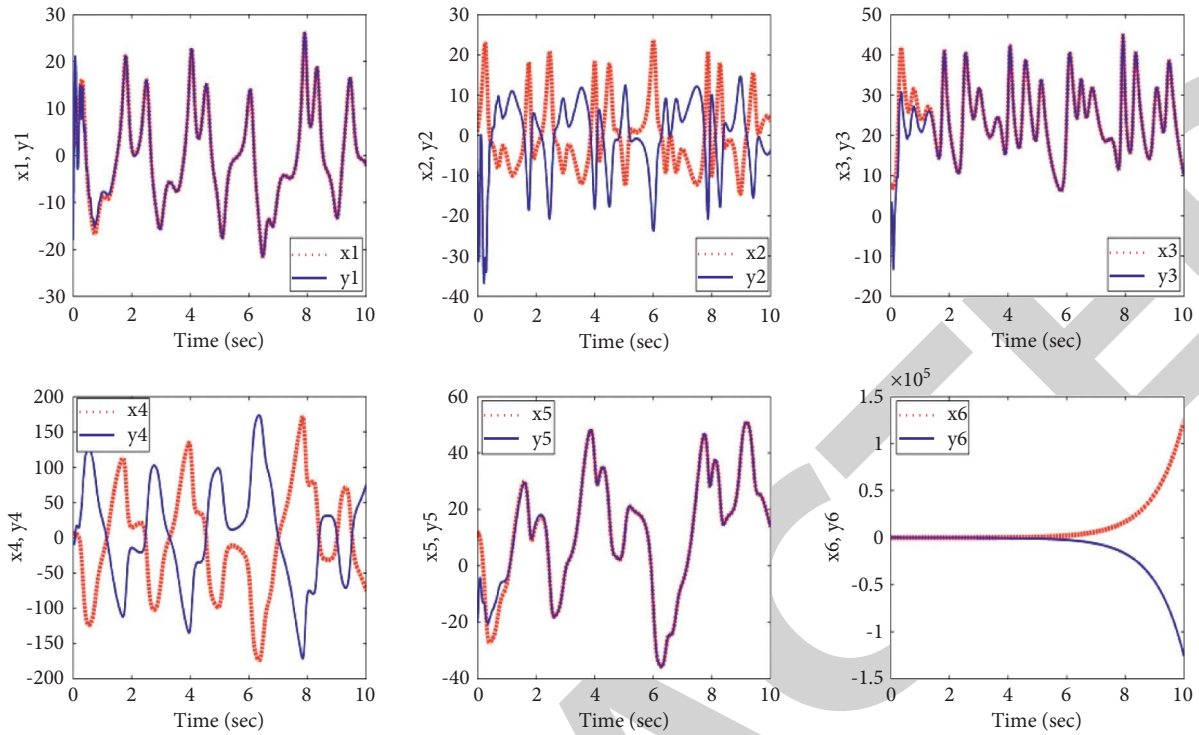


FIGURE 1: Hybrid synchronization between two systems (1).

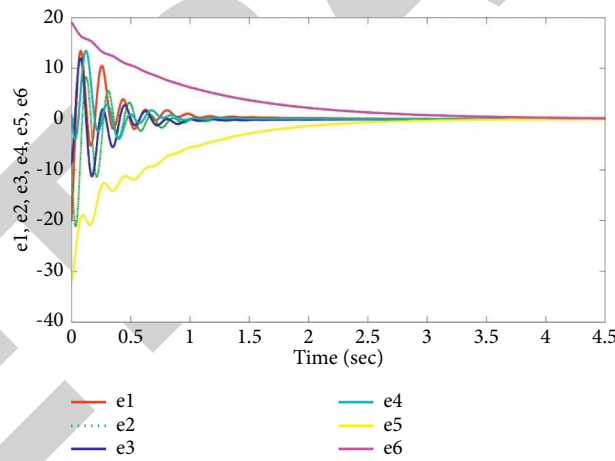


FIGURE 2: The convergence of system (4) with control (10).

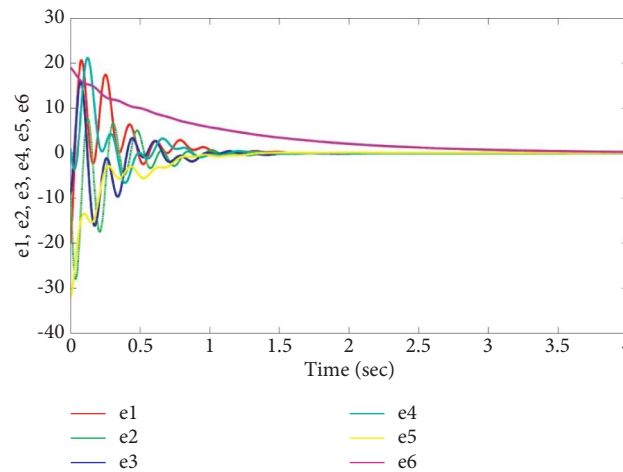


FIGURE 3: The convergence of system (4).

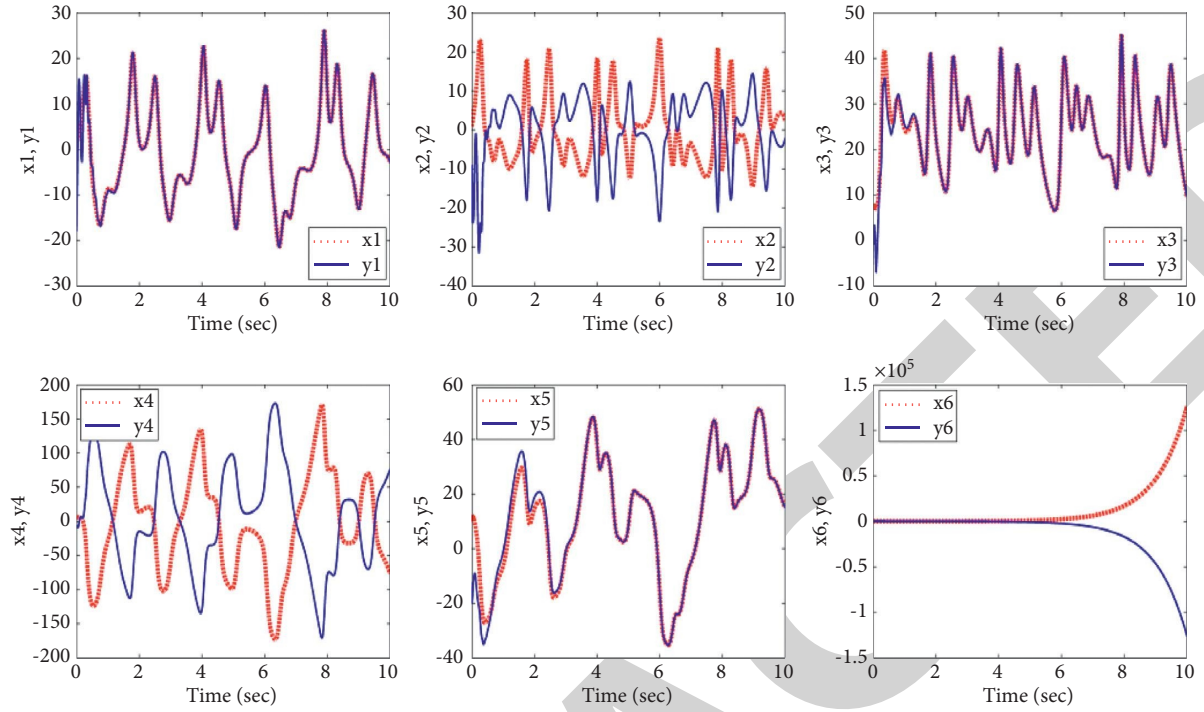


FIGURE 4: Hybrid synchronization between the two systems (1).

The linear approximation method succeeded in showing the hybrid synchronization between the two systems.

The second method is the method of Lebanov. The Lebanov derivative with control (10) is as follows:

$$\dot{V}(e) = -ae_1^2 - e_2^2 - be_3^2 - de_4^2 - e_5^2 - he_6^2 + e_2e_5(1-k) + e_2e_6(r-1) = -e^T Q_1 e. \quad (13)$$

So, we get the matrix

$$Q_{1_4} = \begin{bmatrix} a & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{(k-1)}{2} & \frac{(1-r)}{2} \\ 0 & 0 & b & 0 & 0 & 0 \\ 0 & 0 & 0 & d & 0 & 0 \\ 0 & \frac{(k-1)}{2} & 0 & 0 & 1 & 0 \\ 0 & \frac{(1-r)}{2} & 0 & 0 & 0 & h \end{bmatrix}. \quad (14)$$

The matrix Q_{1_4} is nondiagonal.

Now, find the parameters to confirm that the array is negatively defined.

$$\left\{ \begin{array}{l} 1, \quad a > 0, \\ 2, \quad b > 0, \\ 3, \quad d > 0, \\ 4, \quad 1 > \frac{(k-1)^2}{4}, \\ 5, \quad \left(h - \frac{h(k-1)^2}{4} - \frac{(1-r)^2}{4} \right) > 0. \end{array} \right. \quad (15)$$

There are some inequalities that are not correct, and therefore, Q_{1_4} is negatively defined, so the control failed to achieve hybrid synchronization between the two systems, and to overcome this problem, we update the P-matrix with the same control as follows:

$$P_{1_4} = \text{diag}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{5}{84}, \frac{1}{2}\right). \quad (16)$$

The simulation was implemented via the wolf algorithm and MATLAB software 2020, with parameters $a = 11, b = -7/3, c = 25, d = 2, h = 9.1, r = 1, p = 1, q = 2$ and control parameter $k = 13.5$, and the new model has five +ve Lyapunov spectra.

The derivative of Lebanov is as follows:

$$\dot{V}(e_i) = -10e_1^2 - e_2^2 - \frac{8}{3}e_3^2 - 2e_4^2 - \frac{5}{42}e_5^2 - e_6^2 = -e^T Q_{24} e, \quad (17)$$

such that $Q_{24} = \text{diag}(10, 1, 8/3, 2, 5/42, 1)$; it is a positive definition matrix, thus achieving hybrid synchronization between the two systems (1) and (2) [23] Figure 4 shows the numerical validity of what we have arrived at results. \square

3. Conclusion

In Figure 2, the convergence system of the complete synchronization scheme, we focused on the nonlinear control strategy, and another method was suggested, namely, linearization; in addition, we used the Lyapunov method which is adopted in all previous works in order to compare and verify between the two methods.

The results show that the linearization method is the best for achieving the synchronization since the stability Lyapunov method needs to construct an auxiliary function (Lyapunov function) and may need to update this function sometimes. At other times, it is difficult for us to create a suitable auxiliary function which leads to the fall of this method; thus, the failure and success of the method depend on the additional auxiliary factor, in addition to the control factor. While, the linearization method dispenses for this auxiliary factor, which gives it extra strength compression of the stability of the Lyapunov method. It also addressed the issue of known parameters and unknown.

As for the phenomenon of projective synchronization, which is the most comprehensive among the phenomena, we were only using the method of Lyapunov in achieving the phenomenon, and the results were good.

Data Availability

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

Disclosure

It was performed as a part of the employment of institutions.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors deeply acknowledge Taif University for supporting this study through Taif University Researchers Supporting Project Number (TURSP-2020/150), Taif University, Taif, Saudi Arabia and Princess Nourah bint Abdulrahman University Researchers Supporting Project

number (PNURSP2022R236), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

- [1] M. B. Alazzam, F. Alassery, and A. Almulihi, "Development of a mobile application for interaction between patients and doctors in rural populations," *Mobile Information Systems*, vol. 2021, Article ID 5006151, 8 pages, 2021.
- [2] A. Barzin, A. Sadeghieh, H. KhademiZare, and M. Honarvar, "Hybrid bio-inspired clustering Algorithm for energy efficient wireless sensor networks," *Journal of Information Technology Management*, vol. 11, no. 1, pp. 76–101, 2019.
- [3] Al. Abaji and M. Abdulkareem, "Cuckoo search Algorithm: review and its application," *Tikrit Journal of Pure Science*, vol. 26, no. 2, pp. 137–144, 2021.
- [4] R. N. Salih and M. A. Al-jawaherry, "Finding minimum and maximum values of variables in mathematical equations by applying firefly and PSO Algorithm," *Tikrit Journal of Pure Science*, vol. 25, no. 5, pp. 99–109, 2020.
- [5] B. Bao, J. Xu, Z. Liu, and Z. Ma, "Hyperchaos from an augmented lü system," *International Journal of Bifurcation and Chaos*, vol. 20, no. 11, pp. 3689–3698, 2010.
- [6] C. Zhu, "Control and synchronize a novel hyperchaotic system," *Applied Mathematics and Computation*, vol. 216, no. 1, pp. 276–284, 2010.
- [7] F. M. Abdoon, A. I. Khaleel, and M. F. El-Tohamy, "Utility of electrochemical sensors for direct determination of nicotinamide (B3): comparative studies using modified carbon nanotubes and modified β -cyclodextrin sensors," *Sensor Letters*, vol. 13, no. 6, pp. 462–470, 2015.
- [8] S. A. Salih and Z. G. Atiya, "Applying a mathematical model to simulate the ground water reservoir in Al-Alam area/Northeast Tikrit city/Iraq," *Tikrit Journal of Pure Science*, vol. 26, no. 3, pp. 60–66, 2021.
- [9] M. B. Alazzam, F. Alassery, and A. Almulihi, "Diagnosis of melanoma using deep learning," *Mathematical Problems in Engineering*, vol. 2021, Article ID 1423605, 9 pages, 2021.
- [10] A. Khadidos, A. Khadidos, O. M. Mirza, T. Hasanin, W. Enbeyle, and A. A. Hamad, "Evaluation of the risk of recurrence in patients with local advanced rectal tumours by different radiomic analysis approaches," *Applied Bionics and Biomechanics*, vol. 2021, Article ID 4520450, 9 pages, 2021.
- [11] F. M. Abdoon and S. Y. Yahyaa, "Validated spectrophotometric approach for determination of salbutamol sulfate in pure and pharmaceutical dosage forms using oxidative coupling reaction," *Journal of King Saud University Science*, vol. 32, no. 1, pp. 709–715, 2020.
- [12] G. Alshammari, A. A. Hamad, Z. M. Abdullah et al., "Applications of deep learning on topographic images to improve the diagnosis for dynamic systems and unconstrained optimization," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 4672688, 7 pages, 2021.
- [13] F. Q. Dou, J. A. Sun, and W. S. Duan, "Anti-synchronization in different hyperchaotic systems," *Communications in Theoretical Physics*, vol. 50, pp. 907–912, 2008.
- [14] W. A. Saeed and J. S. Abdulghafoor, "Convergence solution for some harmonic stochastic differential equations with application," *Tikrit Journal of Pure Science*, vol. 25, no. 5, pp. 119–123, 2020.
- [15] A. C. Fowler, J. D. Gibbon, and M. J. McGuinness, "The complex Lorenz equations," *Physica D: Nonlinear Phenomena*, vol. 4, no. 2, pp. 139–163, 1982.

Retraction

Retracted: TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] L. Lin, S. Yang, Q. Bai, and Y. An, "TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 4268681, 7 pages, 2022.

Research Article

TCM Treatment and Drug Co-Occurrence Analysis of Psoriasis

Li Lin ¹, Suqing Yang,² Qingsong Bai,¹ and Yuepeng An ²

¹School of Graduate, Heilongjiang University of Chinese Medicine, Harbin 150040, Heilongjiang, China

²Department of Dermatology, First Affiliated Hospital, Heilongjiang University of Chinese Medicine, Harbin 150040, Heilongjiang, China

Correspondence should be addressed to Yuepeng An; yanyan7878@21cn.com

Received 7 December 2021; Revised 25 December 2021; Accepted 28 December 2021; Published 29 January 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Li Lin 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.

Psoriasis is a long-term immune-mediated disease. Patients with a long history and slow progress are more common, and its treatment is difficult. This study proposes to use traditional Chinese medicine to treat psoriasis. Through the follow-up of all participants for 12 and 24 weeks, a large number of comparative experiments effectively verify the effectiveness of the method proposed in this study. The research results of this study can provide some reference ideas for follow-up research.

1. Introduction

1.1. Psoriasis (Chart 1). Approximately 1–3% of the world's population has some degree of psoriasis, which can be remitted and recur. Although a skin biopsy specimen's histological study may assist confirming the diagnosis, it is usually determined based on clinical symptoms alone [1]. Research shows that even though it is not life-threatening, the condition of having psoriasis may be severely debilitating and has substantial social and economic ramifications [2]. The biology of psoriasis is now well understood, and therapeutic advancements are helping even the most severely affected people. A considerable number of patients have a family history of disease, and some families have obvious genetic tendency. It is generally believed that about 30% of them have a family history. The incidence rate varies widely among different races. Psoriasis is a polygenic genetic disease with the interaction of genetic factors and environmental factors. The incidence of some HLA antigens in patients with this disease was significantly increased. Psoriasis may overlap with other diseases (such as rheumatoid arthritis and atopic dermatitis).

Even though it affects people of all ages, psoriasis manifests itself in two stages: the first happens in teens and young adults (between the ages of 16 and 22) and the second occurs in older people (at 57–60 years of age). Typically, the

visible symptoms consist of circumscribed, thickened, scaly plaques that are pruritic and can be found on the elbows, knees, thighs, and buttocks as well as the scalp and areas of local trauma (Koebner's phenomenon). It is customary to measure the severity of involvement by using the psoriasis region and severity index, which consider factors such as the extent of the area involved, the redness, the thickness, and the scaling [3]. The plaques are characterized by a significantly thicker, horny layer (hyperkeratosis), and the dryness and cracking of this layer are responsible for much of the pain, itching, and inflammatory changes that result. On the psoriasis index, the maximum possible score is 72, and mild, moderate, and severe psoriasis are indicated by values of less than 10, between 10 and 50, and greater than 50, respectively. Plaques of psoriasis may naturally regress without leaving scars after a few weeks, months, or even years. Relapses are also prevalent, and by avoiding exacerbating variables, it is possible to lessen the frequency and severity of relapses in affected people [4]. These relapses are likely to occur in certain cases as a result of the activation of local wound-healing mechanisms in the body. Many scholars have confirmed that streptococcal infection is related to the onset and course of psoriasis from the aspects of humoral immunity (antistreptococcal group), cellular immunity (peripheral blood and skin lesion T cells), bacterial culture, and treatment. In patients with psoriasis, *Staphylococcus aureus*

infection can aggravate skin lesions, which is related to the superantigen of *Staphylococcus aureus* exotoxin. The occurrence of this disease is related to virus. Although there is a certain relationship between viral (such as HIV virus) and fungal (such as *Malassezia*) infection, its exact mechanism has not been finally confirmed.

1.2. Symptoms of Psoriasis. The signs and symptoms of psoriasis may vary from person to person. Here are some examples of frequent indications and symptoms:

- (i) The crimson flesh on the arms and legs is covered in large, silvery scales
- (ii) Scaling spots of varied sizes (often observed in youngsters)
- (iii) Dry, cracked, and itchy skin with the potential to bleed or itch. Itching, burning, or pain are all possible symptoms.
- (iv) Nails that are enlarged, pitted, or ridged swollen and stiff joints. The following formula was used to determine the sample size.

$$n = \frac{z_2 \cdot p \cdot q}{d_2}, \quad (1)$$

where n represents the sample size requested. The standard normal deviation is denoted by z (1.96). p represents the the estimated percentage of the target population who have the attributes being targeted. $q = 1 - p$ represents the statistical significance level (5%).

Psoriasis vulgaris is the most common type, with frequent acute onset. The typical manifestation is erythema with clear boundary and different shapes and sizes, surrounded by inflammatory erythema, slightly infiltrated and thickened. The surface is covered with multiple layers of silvery-white scales. The scales are easy to scrape off. After scraping, it is a light red and shiny translucent film, and small bleeding points can be seen when scraping the film. Skin lesions often occur in the head, sacrum, and lateral extension of the limbs. Some patients felt different degrees of pruritus.

According to the data in Table 1, the vast majority of 96% of patients had psoriasis vulgaris, 16% had pustular psoriasis, 0.7% had erythrodermic psoriasis, and 1.3% had psoriatic arthritis. Of the 11751 patients who had nonmissing disease activity data, 43.3% were in the active stage, 34.1% were stable, 15.3% were resolving, and 7.3% were remitting their symptoms.

2. Types of Psoriasis

In China and other countries, there exist different types of psoriasis, which have been occasionally treated by TCM.

2.1. Plaque Psoriasis. Plaque psoriasis is a persistent autoimmune disorder that affects the skin. It manifests itself on the skin as thick, red, and scaly areas of the skin (Figure 1).

Plaque psoriasis can be an extremely itchy and unpleasant condition, especially in the early stages [5]. It can also be embarrassing, and it does not always respond to medical therapy as expected. It is occasionally mistaken as another skin ailment, such as dermatitis or eczema, and treated as such. Plaque psoriasis is characterized by regions of rough, red skin that are covered with silvery-white scales. This occurs as a result of the skin cells receiving a signal to make new skin cells at an excessively rapid rate [6]. They accumulate and shed in the form of scales and patches. Arthritis psoriasis is also known as psoriatic arthritis. Psoriatic patients have rheumatoid arthritis like joint damage at the same time, which can affect the joints of the whole body, but the lesions of the distal finger (toe) internode joints are the most characteristic. The affected joints are red, swollen, and painful, and the skin around the joints is often red and swollen. The joint symptoms are often aggravated or alleviated at the same time as the skin symptoms. The blood rheumatoid factor is negative.

In addition to causing red and silvery spots on the skin, this buildup of the skin also causes pain and inflammation. Scratching can result in abrasions, bleeding, and infection of the skin. Affected body parts include the elbows, knees, and scalp, which are among the most regularly observed. In these regions, the majority of people who suffer from plaque psoriasis will grow patches of the skin. However, some people will develop psoriasis patches on other parts of their bodies as well. The appearance of psoriasis patches on the body might be unpredictable in its spread. Some patches may be large enough to cover large areas of the body, while others may be a little larger than a dime in diameter.

Once a person has developed psoriasis, it may manifest itself in a variety of different ways and in different locations. The pubic area and armpits are frequently spared when it comes to plaque psoriasis, in contrast to inverse psoriasis.

As areas of plaque psoriasis heal, the position of the lesions may shift. During subsequent attacks, new patches may appear in other locations. Plaque psoriasis manifests itself differently in each individual. No two persons will have the same symptoms as one another.

This kind of medicine includes carboplatin and tacca-thitol. It has good curative effect on plaque psoriasis. Car-potriol cream, ointment, and lotion (for the head) are applied externally twice a day. It usually takes the effect within 8 weeks and will not be dependent after long-term use. The combination of this drug with glucocorticoid or UVB can improve the curative effect. It should be used with caution in patients with bone diseases, calcium metabolism disorders, and renal insufficiency, so as not to cause hypercalcemia.

Psoriasis can manifest itself as a range of nail lesions that are both distinct and ambiguous (Figure 2). Fingernails are more frequently affected than toenails (Figure 3), which is likely because fingernails develop more quickly.

As observed in this graph, there was no statistically significant difference in overall NAPS scores between the left and right hands or between fingers with symmetrical lengths. NAPS was greater for the thumb than for the other fingers, on the other hand. The forefinger, ring finger, and

TABLE 1: Comparison of lesion distribution and triggering factors between genders with psoriasis.

	Total $N = 12031$, n (% = n/N)	Male ($N_1 = 7206$), n_1 (% = n_1/N_1)	Female ($N_2 = 4825$), n_2 (% = n_2/N_2)	P
Triggering factors				
Season	7243 (60.2)	4388 (60.9)	2861 (59.3)	0.203
Stress	4151 (34.5)	2414 (33.5)	1790 (37.1)	<0.001
Smoking	626 (5.2)	533 (7.4)	92 (1.9)	<0.001
Alcohol	2214 (18.4)	1794 (24.9)	405 (8.4)	<0.001
Sphagitis	3296 (27.4)	1816 (25.2)	1486 (31.0)	<0.001
Lesions distribution				
Leg	9348 (77.7)	5355 (74)	3993 (83)	<0.001
Scalp	9024 (75.0)	5797 (80)	3227 (67)	<0.001
Elbow	8069 (67.1)	4757 (66)	3312 (69)	0.183
Thigh	8042 (66.8)	4558 (63)	3484 (72)	<0.001
Forearm	7620 (63.3)	4402 (61)	3218 (67)	<0.05
Back	7602 (63.2)	4173 (58)	3429 (71)	<0.001
Upper arm	7298 (60.7)	4104 (57)	3194 (66)	<0.001
Waist	6639 (55.2)	3719 (52)	2920 (61)	<0.001
Abdomen	6405 (53.2)	3442 (48)	2963 (61)	<0.001

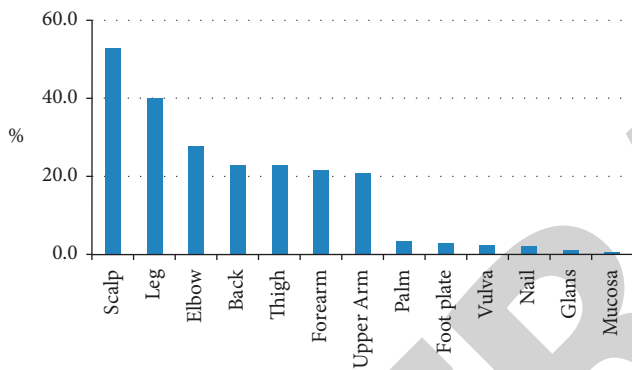


FIGURE 1: The first onset of psoriasis in patients.



FIGURE 2: Plaque psoriasis infected skin. Nail psoriasis and its co-occurrences.

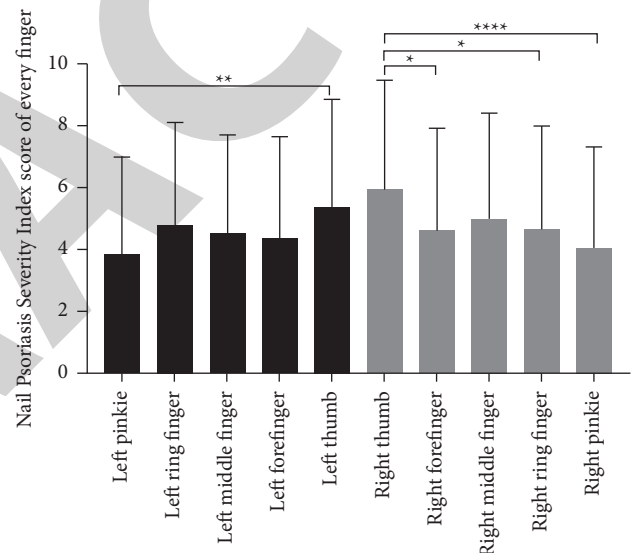


FIGURE 3: Comparison of nail psoriasis (NP) severity of different fingers of each hand ($n = 126$).



FIGURE 4: Nail psoriasis.

little finger all had higher NAPSI scores than a thumb ($p 0.01$) and forefinger ($p 0.01$), although the thumb ($p 0.01$) had lower NAPSI scores than all three ($p 0.05$, 0.05 , and 0.0001) (Figures 4–6).

A variety of tests have determined that the response is either all-encompassing, restricted to the nails exclusively, or

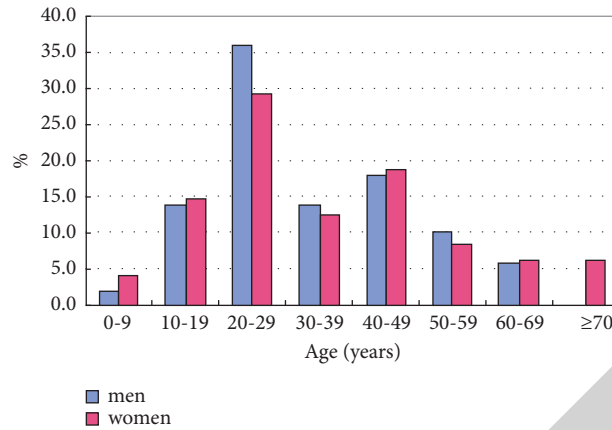


FIGURE 5: Gender psoriasis patients seeking TCM medication in China.

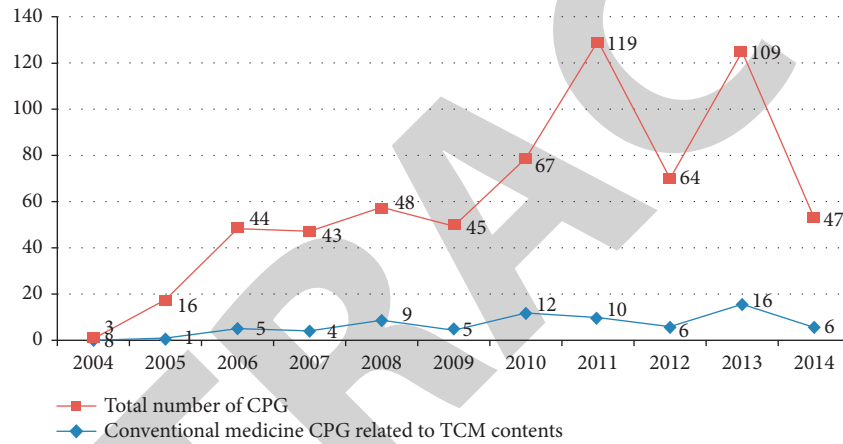


FIGURE 6: The approval of TCM as a treatment for drug co-occurrence of psoriasis in Europe.

a hybrid of the two. Pitches are the small, sharp depressions on the nail's surface are the most common and immediately obvious symptom of a problem [7]. There are no differences in terms of size or depth between everything (Table 2). It is possible to distribute things at random or in transverse or short longitudinal lines that are parallel to one another. Small psoriatic foci develop in the apical matrix of these locations, producing parakeratosis and breaking of the proximal nail fold as it emerges. There is parakeratosis if the nail plate has an ivory-colored region in the proximal half [8]. As shown in Figure 4, It is possible to have one or many pits, and they may be associated with psoriasis or exist independently of the skin disorder. To have psoriasis, you need to have at least 10 craters on one of your nails.

The etiology of psoriasis includes genetic factors, immune factors, infection factors, endocrine factors, living habits, drug factors, environmental factors, and mental factors. Among them, endocrine factors and pregnancy can make the skin lesions disappear or reduce and can also aggravate the skin lesions. Thyroid diseases, diabetes, and other endocrine diseases have no obvious effect on the disease.

TABLE 2: Disorder of nail with psoriasis.

Characteristics	$N = 12031$ (% = n/N)
Nail damage, n (%)	3592 (29.9)
Nail pitting	2167 (18)
Oil drops	1765 (14.7)
Subungual hyperkeratosis	1395 (11.6)
Longitudinal ridging of nails	1326 (11.0)
Onycholysis	856 (7.1)
Onychomadesis	300 (2.5)
Joint damages ^a , n (%)	740 (6.2)
Digital joints	406 (3.4)
Knee joint	356 (3.0)
Toe joint	190 (1.6)
Ankle joint	131 (1.1)
Elbow joint	114 (0.9)
Wrist joint	114 (0.9)
Vertebral column	94 (0.8)

2.2. Treatment of Nail Psoriasis and Its Co-Occurrences. All therapies for nail psoriasis take a long time, since the nail is a slow epidermal attachment. The effectiveness of any treatment is seldom proved before 3–6 months, and it may

take a year or more to get the greatest improvement feasible with a specific therapy [9]. As shown in Figure 5, to help patients visualize the therapeutic effects, photos taken before treatment is highly recommended. These photos should be presented to patients at all follow-up sessions [10]. As a result, before beginning treatment for nail psoriasis, patients should be screened for concomitant onychomycosis. A changed nail that remains despite appropriate treatment might be an indication of concomitant fungal infection. When onychomycosis is present, it might make it more difficult to treat the patient.

There are a variety of therapy options. As shown in Figure 6, some of the considerations that go into their decision-making include the degree of nail involvement and the impact on quality of life, as well as other conditions such as rheumatoid arthritis and other comorbidities [11]. Even though topical treatments for nail psoriasis are difficult to find, skin-cleansing systemic medicines work effectively for this condition [12]. There is a limited amount of topical medication that can reach the affected tissues because of the small nail fold and subungual hyperkeratosis in matrix lesions and nail bed psoriasis. Pits are the toughest wounds to cure, despite their subtle nature. An antipsoriatic topical therapeutic trial of three months is recommended (Figure 1). A shorter nail length increases the drug's ability to reach the psoriatic lesion. To treat nail bed psoriasis, the onycholytic nail must be cut over the infected nail bed. Drilling holes in the nail plate using mechanical burrs or ablative lasers improves nail penetration after filing or grinding. Laser ablation is a method used to analyze solid samples in inductively coupled plasma mass spectrometry. The laser used for laser shirt burning is a high-power, high repetition laser with large area sampling. In this way, the sampling uniformity is good and is not limited by the type of sample. Laser ablation mass spectrometry can be used to analyze small areas of samples.

2.3. Guttate Psoriasis. This subtype tends to affect youthful adults and children. It is frequently caused by a bacterial infection, such as strep throat. It is distinguished by tiny, drop-shaped, scaling sores on the torso, limbs, or legs.

2.4. Inverse Psoriasis. This primarily affects the groin, buttocks, and breast skin folds. Smooth areas of the red skin are caused by inverse psoriasis, which worsens with friction and sweating. This kind of psoriasis can be caused by fungal infections.

2.5. Psoriatic Arthritis. The disease is characterized by inflamed and aching joints, which are archetypal of arthritis symptoms. Sometimes, the symptoms of psoriasis in the joints are the first or only symptom or sign of the disease. There are times when only nail changes are visible. Psoriatic arthritis manifests itself in a variety of ways, ranging from moderate to severe, and can affect any joint. Stagnant joints can become stiffer as a result of this, and in the most severe cases, joint damage can become irreversible.

2.5.1. Causes of Psoriasis. An issue with the immune system produces psoriasis, which causes the skin to regenerate at a quicker pace than normal. Plaque psoriasis is the most prevalent variety of psoriasis, and it is characterized by a rapid turnover of cells that results in scales and red patches of the skin. We still do not know exactly what it is that causes the immune system to malfunction [6]. Both genetics and environmental variables, according to the researchers, play a role. It is not contagious to have this illness.

The immune system is an important system for the body to perform immune response and immune function. It is composed of immune organs, immune cells, and immune molecules. The immune system has the function of identifying and eliminating antigenic foreign bodies, coordinating with other systems of the body, and jointly maintaining the stability of the internal environment and physiological balance.

2.5.2. TCM Treatment and Drug Co-Occurrence of Plaque Psoriasis. In the Western world, the use of Chinese herbs to treat psoriasis is considered an alternative therapy because there is insufficient clinical evidence to support its effectiveness [13]. The traditional Chinese medicines Sheng Di Huang (also known as Qing Dia), Qian Cao Gin (also known as Qian Hua She Cao), and Bai Hua She Cao are considered mainstream options for the treatment of this skin condition among the billion or so people who live in China. When it comes to psoriasis treatment, traditional Chinese medicine (TCM) takes a completely different approach than when it comes to diagnosing and treating the condition in the West [14]. When it comes to traditional Chinese culture, the disease known as Bai Bi is supposed to be caused by "heat pathogens" which are responsible for the characteristic redness of the skin (Table 3).

According to some theories, blood "dryness" and "stagnation" are responsible for the itching, pain, and scaly plaque formation that characterize the autoimmune condition as the disease progresses.

Because traditional Chinese medicine is based primarily on impression and personal experience, the therapies prescribed by one practitioner may differ from another [15]. In contrast to Western treatment, traditional Chinese medicine therapies do not so much treat the sickness as they do correct the imbalances that cause situations such as heat, dryness, and stagnation.

TCM remedies are frequently combined and matched to target certain imbalances. As shown in Table 4, it would be more appropriate to employ herbs that "clear heat" in higher quantities than herbs that ease stagnation or dryness if blood heat is the primary concern in a given situation.

Qian Cao gin is the Chinese name for the herb common madder, which is also known as yarrow (*Rubix rubiae*). This herb is said to contain blood-cooling and antiproliferative effects, which may help to temper, if not prevent the growth of psoriasis plaques. It is also used to treat acne.

A Chinese herb known as Qing Dai (Indigo Naturalis) is considered to have great heat-clearing qualities, as well

TABLE 3: Clinical features and treatment of psoriasis by TCM.

Characteristics	No. of cases (%)
Total	N = 12031
Treatment history	
Systemic treatments	9600 (79.8)
TCM ^a	8661 (72.0)
Vitamin A derivatives	2604 (21.6)
Immune depressants	1297 (10.8)
Glucocorticoids	1577 (13.1)
Topical treatments	10787 (89.7)
Glucocorticoids	9113 (75.7)
Vitamin D3 analog	3634 (30.2)
Tazarotene	2976 (24.7)
TCM ^a	2802 (23.3)
Salicylic Acid	2295 (19.1)
Other therapy	2634 (21.9)
Photo therapy	2220 (18.5)

TABLE 4: A summary of the patient population.

Characteristics	N = 12031
Sex, <i>n</i> (%)	
Male	7206 (59.9)
Female	4825 (40.1)
Age, <i>y</i>	
Minimum	1
Maximum	86
Mean (SD)	39.42 (18.4)
Family history of psoriasis, <i>n</i> (%)	
Family history of psoriasis	2781 (23.1)
Relatives	N = 4995
First degree relatives	2323 (46.5)
Second degree relatives	1538 (30.8)
Third degree relatives	1134 (22.7)
History of smoking and alcohol, <i>n</i> (%)	
Smoking history	4023 (33.4)

as antiproliferative and anti-inflammatory properties. Qing Dai is also known as Indigo Naturalis in the west.

2.5.3. TCM Treatment. TCM, which is a system of ancient medical practice that differs from contemporary medicine in its substance, methodology, and philosophy, plays an essential part in the health maintenance of Asian peoples, and it is becoming more frequently adopted in countries around the Western world [16]. Despite their growing popularity, there are still some misconceptions about traditional medicines and the standards they must satisfy to be effective. In this area, we try to dispel some of the most common myths.

For the longest time, Asia's medical system has relied on traditional Chinese medicine (TCM) [1]. Traditional Chinese Medicine was developed after tens of thousands of years of empirical research and refining (TCM). Until English missionaries arrived in China during the early 1800s, it was the only medical specialty available in the country [17]. They brought modern medicine's medications, equipment, and practices with them.

2.6. Psoriasis Dosage

$$RBW = \frac{ABT (g)}{IB (g)}. \quad (2)$$

RBW represents the relative body weight, ABT represents the absolute body weight, and IB represents the psoriasis patient body weight.

These doses were calculated for every psoriasis patient concerning their weight.

TCM comprises a wide range of techniques (chart 4) including herbal medicine and acupuncture that are familiar to Westerners as well as practices that are new to most Westerners such as cupping (hot cup therapy), Tuina (massage), and Qigong (burnt mugwort therapy). Traditional Chinese medicine (TCM) research is currently focused on understanding the mechanisms of action of these treatments [18, 19]. For the most part, this Nature Outlook will focus on herbal remedies as their efficacy is almost equivalent to that of current pharmaceuticals.

2.7. Acupuncture TCM as an Analysis of Psoriasis. Using small needles inserted through the skin, practitioners apply pressure to certain points on the body to stimulate healing. Studies have shown that acupuncture, whether real or fake, is equally effective, demonstrating that the placebo effect is at play. By altering brain regions involved in processing pain, acupuncture is thought to boost the body's generation of natural analgesics [20]. True acupuncture, according to the results of several studies, may be useful in treating chronic pain problems like osteoarthritis/knee pain and carpal tunnel syndrome as well as low back pain and neck discomfort. The frequency of tension headaches may be reduced, and migraine headaches may be prevented using this supplement. This is provided in Center for Chronic Disease and Hypertension Fact Sheet on Acupuncture (NCCIH) [21, 23].

2.8. Tai Chi TCM as an Analysis of Psoriasis. Postures, gentle motions, mental attention, breathing, and relaxation are part of a workout program [17]. Patients with Parkinson's disease, as well as the elderly and those who practice Tai Chi, may find it helpful with balance and stability. It has also been shown to help people with fibromyalgia and back pain, as well as enhance the quality of life and mood in those with heart failure. More information can be found about Tai Chi by going to the National Center for Chronic Disease and Hypertension's Tai Chi Fact Sheet.

2.9. Aspects of TCM an Analysis of Psoriasis. A wide range of medical illnesses has been researched using TCM medications, including stroke, heart disease, mental illness, and respiratory ailments (such as bronchitis and the common cold). About one out of every five Americans, according to the results of a national poll, uses Chinese herbal medications to treat a variety of medical ailments (Parker et al., 2016). It is impossible to draw firm conclusions about the value of many studies because they are of low quality.

Retraction

Retracted: DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. Zaguia, D. Pandey, S. Painuly, S. K. Pal, V. K. Garg, and N. Goel, "DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 8393498, 11 pages, 2022.

Research Article

DNA Methylation Biomarkers-Based Human Age Prediction Using Machine Learning

Atef Zaguia ¹, Deepak Pandey ², Sandeep Painuly ², Saurabh Kumar Pal ²,
Vivek Kumar Garg ³, and Neelam Goel ²

¹Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

²Department of Information Technology, University Institute of Engineering and Technology, Panjab University, Chandigarh 160014, India

³Department of Medical Lab Technology, University Institute of Applied Health Sciences, Chandigarh University, Gharuan, Mohali 140413, Punjab, India

Correspondence should be addressed to Neelam Goel; erneelam@pu.ac.in

Received 30 October 2021; Revised 20 November 2021; Accepted 22 December 2021; Published 24 January 2022

Academic Editor: Alexander Hošovský

Copyright © 2022 Atef Zaguia 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.

Purpose. Age can be an important clue in uncovering the identity of persons that left biological evidence at crime scenes. With the availability of DNA methylation data, several age prediction models are developed by using statistical and machine learning methods. From epigenetic studies, it has been demonstrated that there is a close association between aging and DNA methylation. Most of the existing studies focused on healthy samples, whereas diseases may have a significant impact on human age. Therefore, in this article, an age prediction model is proposed using DNA methylation biomarkers for healthy and diseased samples. **Methods.** The dataset contains 454 healthy samples and 400 diseased samples from publicly available sources with age (1–89 years old). Six CpG sites are identified from this data having a high correlation with age using Pearson's correlation coefficient. In this work, the age prediction model is developed using four different machine learning techniques, namely, Multiple Linear Regression, Support Vector Regression, Gradient Boosting Regression, and Random Forest Regression. Separate models are designed for healthy and diseased data. The data are split randomly into 80:20 ratios for training and testing, respectively. **Results.** Among all the techniques, the model designed using Random Forest Regression shows the best performance, and Gradient Boosting Regression is the second best model. In the case of healthy samples, the model achieved a MAD of 2.51 years for training data and 4.85 for testing data. Also, for diseased samples, a MAD of 3.83 years is obtained for training and 9.53 years for testing. **Conclusion.** These results showed that the proposed model can predict age for healthy and diseased samples.

1. Introduction

Aging is the process of getting older. It has been an irrevocable biological practice in an individual lifespan inspired by many aspects. It is related to the transformations in vibrant physiological, biological, and environmental methods [1–3]. The modification in the aging process can be done by chemical or physical changes in a DNA structure at the genetic level that can affect the aging process [4]. Aging can be predicted using many different methods; however, the problem of low prediction accuracy hinders the

possibility of any breakthrough in such research. A number of strategies have been utilized for predicting an aging process but often face the issue of low accuracy in prediction, which obstructs the potentials of several developments in such a domain. Nowadays, DNA methylation (DNAm) data have been emerging as a popular research area employed to predict the epigenetic age of organisms [5, 6]. Recently, it has been shown that the process of aging is extremely associated with the alterations of DNAm in genome-particular situations. DNAm is a genetic means in which methyl clusters are supplemented to the molecule of DNA. Beneath this

practice, an active methyl has been conveyed to a certain base on the chain of DNA in the DNA methyltransferase (DNMT) catalysis [7]. A number of approaches have been given for age prediction, but one of the approaches for predicting human age is measuring and analyzing the skeletal markers is becoming very popular. However, these presented approaches are not trustworthy due to low prediction accuracy and difficulty in performing [8]. It is an extremely well-known fact that aging affects organisms on a macro, molecular, and microscopic level. The practice of DNAm to attain supplementary data in the investigations of forensic sciences has shown to be a favorable field of interest [9]. It has been shown that the interest has been grown in this field of study to find the association between individual age and age-dependent variations in DNAm of particular CpG sites within the genome. With the developments in DNAm research, it is feasible for predicting individuals' age by a quantifiable statistical correlation between DNAm and diverse ages has been ascertained on the basis of the modification rule of methylation with age [7]. Machine learning is one of the most advanced fields that have been employed to make predictions based on the available data by developing models [10–12]. In one of these studies, a supervised machine learning technique has been presented for fitting the protein features model to the set of known nonaging and aging-associated proteins for the prediction of aging-related proteins to determine aging-related properties of the proteins simultaneously. Very little consideration has been performed utilizing supervised learning models to predict aging-related genes of human DNA repair genes.

This article has presented the comparative analysis of three machine learning techniques such as support vector machine (SVM) as a binary classifier for training the data that are linearly nonseparable, logistic regression analysis of the binary sequences, and XGBoost as a scalable tree boosting system for the classification of human proteins as nonaging or aging. These models have been implemented on 21,000 protein features that have been extracted from various databases (Gene Ontology, GeneFriends, and UniProt) and are appropriate to well-known aging-associated human proteins (extracted from GenAge). However, various works have been presented in the literature to predict age employing DNAm. Still, much more work is required to be done in this field.

Therefore, the aim of the presented research work is the utilization of the potential of machine learning and statistical analysis techniques to identify the effect of aging on DNA methylation data of specific CpG sites. Different machine learning techniques like artificial neural network, Random Forest Regression, Support Vector Regression, multiple linear regression, and Gradient Boosting Regression have been applied in the past to design age prediction models. However, due to the limited data used in this study, the artificial neural network is excluded from the present work. This work tries to create robust machine learning models to predict human age using the methylation data from CpG sites in human cells. The goal is to show the correlation between these epigenetic modifications in DNA and human age. The effects of diseases on such correlation between the

molecular-level changes in the human body and human age have also been observed.

The key contributions of this research work are as follows: (i) in this study, six CpG sites having high correlation with age are identified from both healthy and diseased data; (ii) age prediction models are designed using four machine learning techniques, namely, Random Forest Regression, Support Vector Regression, multiple linear regression, and Gradient Boosting Regression; (iii) the impact of human biological age on disease is analyzed by comparing it with predictions from healthy data.

The rest of the article is structured as follows: Section 2 presents the related work. Section 3 discusses the material, methodology, techniques, and performance evaluation parameters used in the present study. The experimental results are summarized in Section 4. Finally, Section 5 concludes the given work.

2. Related Work

The existing work related to human age prediction using machine learning techniques is given below.

Lau et al. have integrated the four variable selection methods with the statistical and machine learning model. A total of 991 whole blood samples of age between 19 years to 101 years have been used. From experimental results, it has been observed that the 16 markers have been chosen with multiple linear regression from the forward selection approach for predicting age. Instead, the machine learning model with a very superior high dimensional variable selection method has appeared superfluous for DNAm-based age predictions [13]. Further, Liu et al. have developed a prevailing Web server named BioSeq-Analysis to construct the predictor. It produced the improved forecaster and utilized the three sequence evaluation chores. From investigational outcomes, it has been revealed that the forecasters produced by BioSeq-Analysis even surpassed state-of-the-art approaches [14]. Additionally, Thong et al. have determined the adequate age predictors by making a comparative analysis of prediction accuracy between the regression model and artificial neural networks (ANN). It also investigated the impact of covariables like sex and ethnicity on predicting age and revealed the less amount of input DNA entailed for bisulfite medication and pyrosequencing for age prediction [15].

Further, Aliferi et al. have performed analysis on 110 blood samples collected from individuals aged between 11 and 93 years using massively parallel sequencing (Illumina MiSeq) based DNAm quantification assay of 12 CpG sites and bisulfite conversion. Employing these data, 17 diverse statistical modeling methods have been contrasted with SVM with a polynomial function (SVMp) model using root mean square error (RMSE) for further testing. To select the models (RMSE = 4.9 years), the mean average error (MAE) of the blind test ($n = 33$) had been computed at 4.1 years, whereas 86% with less than 7 years and 52% with less than 4 years of error had been predicted [16]. Vidaki et al. have introduced the prospective age-related markers. Additionally, a methodology has been given for machine

learning-based prediction analysis using ANN. The given model not only exhibited a good accuracy of prediction but also has the potential to be applied to individuals of various nonblood tissues, ethnic backgrounds, and underage children. However, it has been noted that the predictions can be enhanced in the future by the normalization of various technologies of DNAm analysis. Moreover, in unhealthy individuals, the model worked less accurately; therefore, testing of marker's resistance to DNA methylation alterations in a diseased state needs to be tested further [17]. Similarly, a method for age prediction has been introduced for solving the multivariate regression problem from DNAm data with the optimization ANN model utilizing the Cell Separation Algorithm (CSA) [18].

The CSA impersonates the cell separation action by employing a differential centrifugation method, including manifold centrifugation stages. The saliva samples and diseased/healthy blood samples are utilized for testing the performance of the method. The comparative analysis of CSA has been performed with genetic algorithms, ADAM, and stochastic gradient descent. The results have shown that CSA outperformed other methods [18]. Further, bisulfite sequencing data have been generated for 95 saliva samples utilizing massively parallel sequencing (MPS). It is then contrasted with methylation SNaPshot data from the 95 samples. The age predicted by utilizing MPS data to a model developed for methylation SNaPshot data has diverged significantly from the sequential age because of platform variances. Thus, variables were presented for indicating the type of platform and constructing the platform-independent age predictive models utilizing multivariate linear regression and neural networks. The platform-independent age prediction method has been built on a growing number of platforms introducing platform variables, and this idea can be employed to model age prediction for other body fluids [19].

Smeers et al. assessed the alternate methods for giving more accuracy for age-dependent prediction intermissions. A quantile regression model with weighted least squares (WLS) has been presented. The given model has been contrasted against other regression models on similar data. Both of the models offered the age-dependent prediction intervals, considered for the growing variance with age, but WLS regression outperformed with success rate. Though, quantile regression might be a chosen way to deal with a variance as it is nonconstant and not normally distributed. Besides, deep learning models have shown good findings in disease heterogeneity [20]. Additionally, MethylNet, a DNAm deep learning model, has been built for the construction of embeddings, making predictions, generating fresh data, and uncovering the unspecified heterogeneity with negligible human intervention [21]. Further, an epigenetic timer has been introduced utilizing a suite of methylation markers of five distinct genes of the Italian population samples of various ages enfolding the entire duration of individual life [22]. Moreover, a survey has been done in which a relationship among certain forms of DNA repair and aging has been discovered with numerous aging biomarkers using machine learning. Besides, novel

candidate proteins with robust computational signs of their significant function in the aging have been attained [23, 24]. State-of-the-art machine learning models have been employed for classifying 36 human protein features as nonaging-related or aging-related [25].

Though much work has been done in this area, very few studies focused on comparing the performance of age prediction models based on healthy and diseased samples. In this article, different models using healthy and diseased samples are developed and their effect on age is analyzed.

3. Materials and Methods

3.1. Data Collection. In this work, the DNA methylation data from human blood samples are required. All the data used in the present study are collected from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) [24]. Many GEO datasets were explored to gather data. The data were collected under two categories, namely, healthy individuals and unhealthy individuals. A total of 11 blood datasets are considered for this work. Only those datasets are selected, which provide the individuals' age. All the DNA methylation data taken for the study are acquired from the HumanMethylation27 BeadChip platform [26, 27].

The first dataset is of healthy individuals with little to no gene mutation, between the age range of 4 to 89, and the second dataset was comprised of DNA information of individuals who had severe genetic mutations and were patients of diseases like cancer, Alzheimer's, and Asthma, between the age ranges of 1 to 86. The details of the dataset for healthy individuals are provided in Table 1.

A total of 454 samples were collected across all ages between 4 and 89 using six different datasets, with a mean age of 33.46 and a median age value of 30 years. On further cleaning of data, many samples were rejected as these were outliers and they adversely affected the prediction accuracy in genomic data [7]. For unhealthy individuals, a total of 400 samples are collected across five datasets. The age range was 1 to 86 for this dataset, with a mean age of 41.50 and a median of 41 years. It further approves our understanding of how advancing age is a major risk factor for diseases in humans. Table 2 shows the distribution of unhealthy data.

3.2. Selection of CpG Sites. Initially, 8 CpG sites were selected for this study. These sites are cg22736354, cg19283806, cg18473521, cg02228185, cg06493994, cg19761273, cg01820374, and cg09809672. The inspiration for choosing these CpG sites came from Li et al. [7]. Among these 8 CpG sites, cg22736354, cg06493994, cg19283806, and cg18473521 were positively correlated and cg09809672, cg02228185, cg01820374, and cg19761273 were negatively correlated.

At a later stage, cg19283806 is dropped as there were many missing instances of it in different datasets. cg18473521 was showing high levels of collinearity (a threshold of 0.75 was set for datasets) with cg09809672 and cg19761273 so it also dropped afterward. Finally, the data had 'G' sites and all the values were then recorded manually

TABLE 1: Data collection for healthy individuals.

DNA origin	Platform (K)	No.	Age range	Availability
Blood PBMC 1	27	80	3.6–18	GSE27097
Whole blood	27	93	49–74	GSE20236
Blood CD4+ CD14	27	50	16–69	GSE20242
White blood	27	60	18–89	GSE32396
Blood PBMC	27	80	24–45	GSE37008
Whole blood	450	91	26–101	GSE40279

CD: cluster differentiation; PBMC: peripheral blood mononuclear cell.

TABLE 2: Data collection for unhealthy individuals.

DNA origin	Platform (K)	No.	Age range	Availability
Blood	27	80	23–85	GSE49904
Whole blood	27	100	50–85	GSE19711
Whole blood	27	120	1–32	GSE20067
White blood	27	62	16–86	GSE41037
Blood	450	38	34–72	GSE51032

across each marker for two different datasets. Among these 6 CpG sites, a positive correlation of cg06493994 and cg227363 with age has been found. However, age was found to be negatively correlated with cg01820374, cg19761273, cg02228185, and cg09809672. These results comply with Horvath’s research data results [27]. Figure 1 illustrates the relationship between CpG site Beta-value and healthy dataset Age, whereas the relation between CpG site Beta-value and unhealthy dataset Age is shown in Figure 2.

The combination of these six CpG sites was performing best; thus, data were collected for these six sites. However, each dataset had its own local characteristics infused in the data collected, providing us with noise and outliers, which are to be handled in later stages.

3.3. Methodology. The methodology used in this work is presented in Figure 2. After collecting the dataset, the next task is to make these data useful for prediction. The dataset initially obtained could not be used directly with machine learning models, as this dataset was uncleaned and it had many outliers and all the features were not scaled. Moreover, due to the local noise per array dataset, an uneven distribution was there hindering the performance of the designed models. The data was thoroughly cleaned and processed before using it, four machine learning algorithms were then selected and four different evaluation metrics were used to evaluate the performance of the machine learning models. A detailed description of these steps is given as follows.

3.3.1. Preprocessing. The raw datasets were highly unevenly distributed and had a large number of outliers. As the data were generated at different times, we can easily observe batch effects in the data between different data platforms. These batch effects were removed by normalizing the methylation levels between different datasets. The data were log-transformed to create a normal distribution before sending them to regression models. The purpose of this was to improve the generalization of the model and allow the use of standard

scaling on our dataset, as the Beta-value of each marker had a difference in 10 to 100 s of magnitudes, causing an unequal contribution in final mapping. The data were then confined between the quantile range of 0.20 to 0.80, which resulted in the huge loss of data points, so this strategy was later replaced with manual inspection and removal of extreme values for each feature. After cleaning, a total of 15 healthy and 13 diseased samples were removed.

After cleaning, the dataset had an almost normal distribution, as the features were scaled to create a more robust model. The below-given histogram shows the age distribution of datasets after cleaning. The age distribution histograms for healthy individuals and unhealthy individuals are illustrated in Figures 3(a) and 3(b), respectively.

After manually cleaning both the datasets, the healthy dataset had a total of 439 samples, each with six features and one continuous label as Age; also, the count of the unhealthy dataset dropped to 377 after cleaning. The next step was to pass these data to machine learning pipelines. One column was dropped from both the datasets because of the high level of multicollinearity, which refers to a condition where more than two explanatory variables in a multiple regression model are highly linearly associated. The threshold was kept at 0.70; if the value is more than this, then one of the features can be dropped as it saves against the curse of dimensionality. After cleaning up the dataset, a total of 439 samples were selected for the healthy dataset and 377 samples were selected for the unhealthy dataset. The first attempt is a 60 : 40 split, which reduces accuracy. Finally, after testing multiple splits, an 80 : 20 random split is selected.

3.3.2. Algorithms. Four different machine learning models are selected for age prediction. The models were chosen which were robust to nonlinear mapping. It could be easily seen that the relationships between the features and the age variable are rather complex to map; thus, ensemble methods are preferred over linear regression models. These methods are metaalgorithms that combine several predictive models to create more robust models. These are generally used to manage bias-variance tradeoffs and improve prediction accuracy. One bagging method, namely, Random Forest Regression, and one boosting method, namely, Gradient Boosting regression, are chosen in this work. We also used Support Vector Regression and multiple linear regression to compare and benchmark the results with these popular methods.

- (i) Multiple Linear Regression: Multiple Linear Regression is a statistical technique used to model the relationship between multiple explanatory variables and a scalar response. Multiple regression uses a linear function to predict values based on ground truth.
- (ii) Support Vector Regression: Support Vector Regression is used to predict discrete values. Support Vector Regression is based on a support vector machine and the goal is to find the optimal

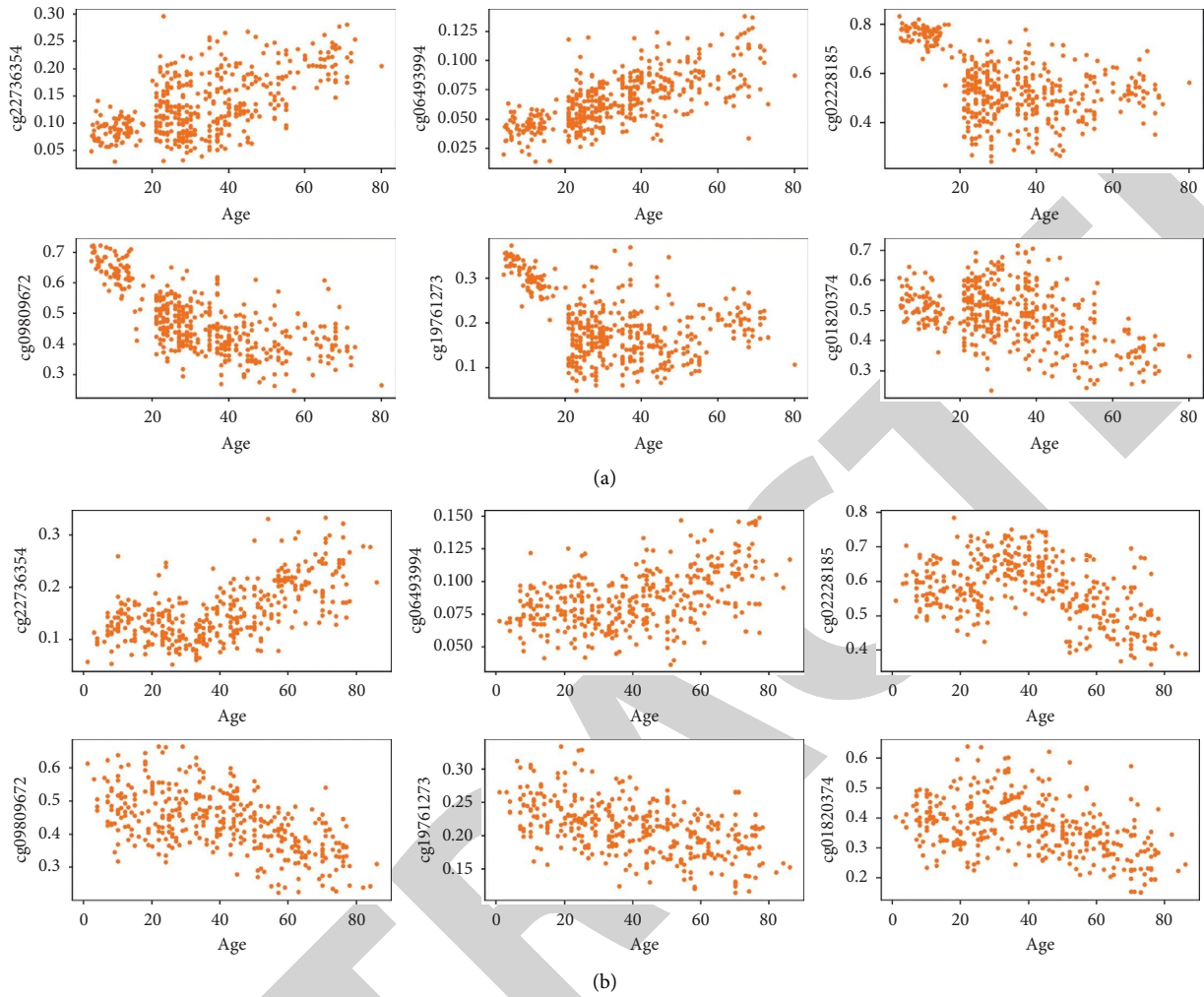


FIGURE 1: (a) Relation between Beta-value of CpG site and Age. (b) Relation between Beta-value of CpG site and Age.

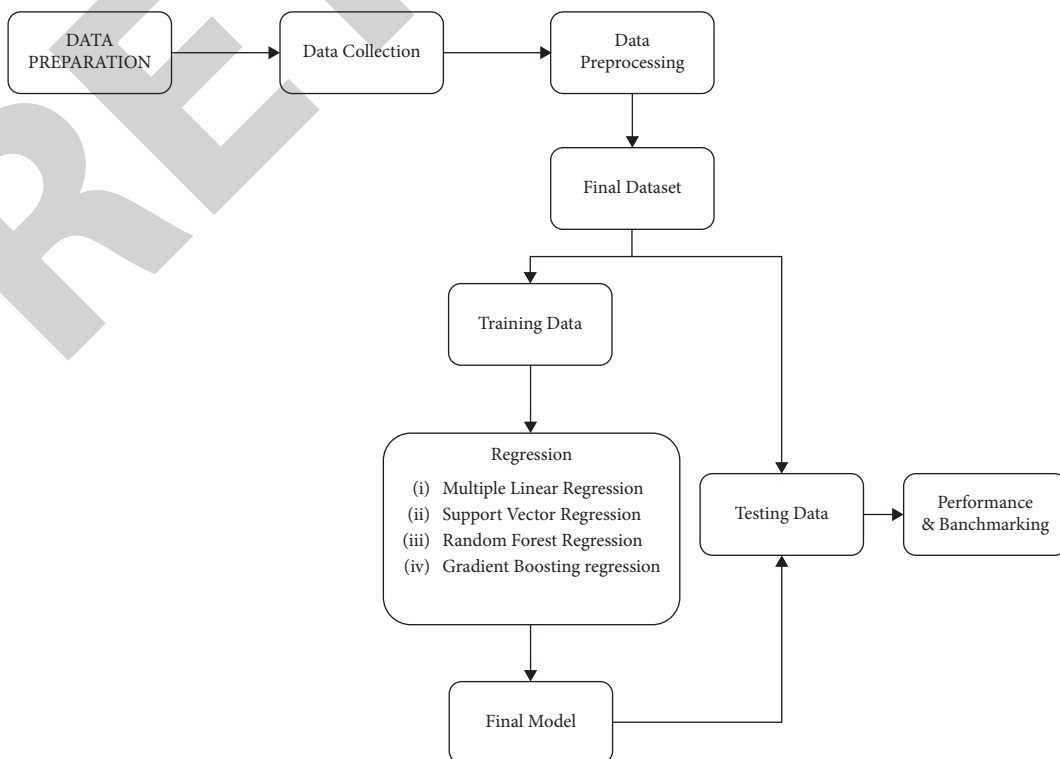


FIGURE 2: Proposed methodology.

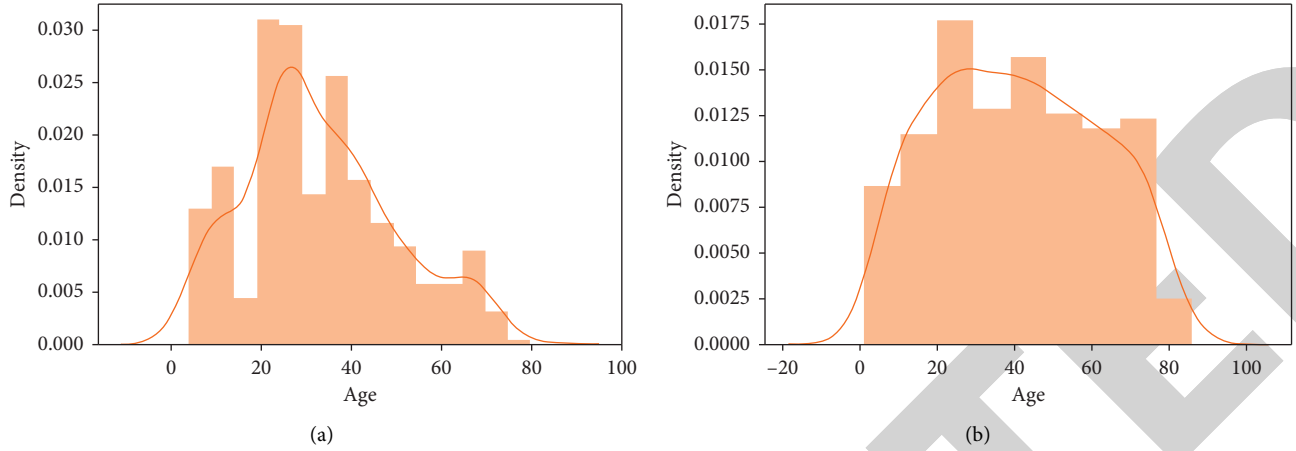


FIGURE 3: (a) A histogram of the age distribution for healthy individuals; (b) disease individuals.

hyperplane that covers the maximum number of data points.

- (iii) Random Forest Regression: Random Forest Regression is an ensemble-based model of machine learning. Now we combine several decision trees to create a robust model for learning the complex associations between features and output variables.
- (iv) Gradient Boosting Regression: Gradient Boosting Regression is another ensemble-based method developed on the principle of boosting along with a suboptimal model, which provides a powerful predictive model.

3.3.3. Evaluation Metrics. As the problem at hand is a regression problem, four statistical metrics were chosen to evaluate the performance of age prediction models. The degree of correlation between true and predicted age was calculated using the R^2 -score:

$$R^2 = 1 - \frac{RSS}{TSS},$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}}, \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2.$$

The age prediction model is evaluated using mean absolute deviation (MAD). The MAD determines the mean absolute deviation between the real age and predicted age using DNAm data [28, 29].

$$\frac{\sum_{i=1}^m |y^i - \bar{y}|}{m}, \quad (2)$$

where m represents the target values $y = (y_1, y_2, \dots, y_m)^T$, y^i signifies the value of prediction, and $f(x^i)$ denotes the feature vector x^i regression function. The MAD represents the

absolute deviation, RMSE (root mean square error), and MSE (mean square error) [28].

- (i) R-Squared (Coefficient of Determinations): it is also called the coefficient of determination. This metric shows how well the model fits into a particular dataset. This shows how close the regression line (that is, the plotted predicted values) is to the actual data values. The coefficient of determination has a value between 0 and 1, where 0 indicates that this model does not match the provided data and 1 means that the model exactly matches the provided dataset.
- (ii) Root-Mean-Squared Error or RMSE: RMSE is the root-mean-squared error that occurs when making data set predictions. This is the same as MSE (mean squared error), but the square root of the value is taken into account when determining the accuracy of the model.
- (iii) Mean Absolute Error or MAE: the MAD, or mean absolute deviation, for a particular dataset is the average of the absolute deviations from the central distribution point. The absolute difference means that the result with a negative sign is ignored. Hence, $MAE = |\text{true values} - \text{predicted values}|$.

3.3.4. Model Performance: Dealing with Overfitting and Underfitting

- (i) Outlier handling: outliers can have a significant impact on the model, as there are relatively small datasets initially. Therefore, it was necessary to identify and delete the outliers. To get a valid model on a small dataset, it is essential to remove the effects of outliers. Also, depending on the use case, we manually excluded outliers to avoid being affected by highly distributed issues. Due to the small size of the dataset and the few extreme values, we have carefully selected and deleted them.
- (ii) Feature selection: explicit feature selection is usually not the best approach, but it can be an important

step if you have limited data. Due to the small number of observations and a large number of predictors, it is difficult to avoid overfitting. There are several approaches to feature selection, including correlation analysis with target variables, importance analysis, and recursive elimination. Also, note that feature selection always benefits from domain expertise. In this use case, we will perform a univariate analysis of the features to see which features contribute significantly to the output variables and use this only as an input. Select a model. This also helped avoid the problem of overfitting.

- (iii) Ensemble-based model: the combination of results from multiple models allows for much more accurate predictions. For example, the final forecast, which is calculated as a weighted average of the forecasts from different individual models, has significantly lower variance and greater generalizability than the forecasts from the individual models. And according to our use case, we have an ensemble technique, NS-Random forest, and this increases generalizability compared to individual models.

3.4. Experimental Setup. The implementation is done in the python programming language (python 3.7.5). Various python packages like pandas, numpy, scipy, seaborn, and sklearn are used. To tune the model to the best parameters, techniques like RandomizedSearchCV and GridSearchCV are used. There are many parameters (hyperparameters) in each machine learning algorithm. Without experimentation, it is difficult to say which values of these parameters will provide an optimal prediction. The default values given for these parameters may not be optimal in case of different datasets. To determine the best combination of the values of distinct parameters for the given dataset, hyperparameter optimization is carried out [7].

4. Results and Discussion

After preprocessing, both datasets are passed to machine learning pipelines. StandardScaler is used to scale all the features; standardized values are useful for tracking the data, which are difficult to compare otherwise due to different magnitudes, metrics, or circumstances [30]. The Python sklearn package is used to create machine learning pipelines; these pipelines are an ensemble of several transformers with a final estimator [26].

4.1. Results on Healthy Dataset. After cleaning the datasets, a total of 439 samples were finally selected for the healthy dataset. A split of 60 : 40 is tried initially, which results in the best MAD of 3.45 with RandomForestRegressor. After testing several splits, a random split of 80 : 20 is finally selected. A total of 87 samples are saved for testing the models. No hyperparameter tuning is done at this stage. The results of these models are shown in Table 3. The results for healthy testing data are provided in Table 4. It is clear from these results that Random Forest has produced the best score with a MAD of 2.51 on training data and 5.02 on independent data. The second best performance is shown by Gradient Boosting Regression for healthy training data.

These two best models were then selected for hyperparameter tuning. Random Forest Regressor and Gradient Boosting Regression were tuned using randomized searching and grid searching methods to improve the prediction score further. The untuned models performed well, but these had a low degree of generalization. The results on training and independent testing data after hyperparameter tuning are given in Tables 5 and 6, respectively.

Also, these results for random forest regressors on training and testing data are demonstrated in Figure 4. After tuning, the models had a good degree of generalization, and the MAD for testing dropped to 4.85 years for Random Forest Regression.

4.2. Results on Unhealthy Dataset. In the unhealthy dataset, 377 samples were selected and 15 samples were rejected. A split of 60 : 40 was tried initially, which resulted in the best MAD of 5.68 with Random Forest Regression; after testing several splits, a random split of 80 : 20 was finally selected. A total of 76 samples were saved for testing the models. No hyperparameter tuning was done at this stage. The results of these models for training data are shown in Table 7. The results for independent testing data are shown in Table 8. It has been observed from the results that Random Forest produced the best score with a MAD of 3.83 on training data and 9.53 on independent data. The untuned models performed well, but these had a low degree of general.

Like in the case of healthy data, the two best models were selected for hyperparameter tuning. Random Forest Regressor and Gradient Boosting Regression were tuned again using randomized searching and grid searching methods to improve the score further for unhealthy data. The results on training and testing data after hyperparameter tuning for the two best models are presented in Tables 9 and 10.

The results for Random Forest Regressor are shown in Figure 5. After tuning, the models had a good degree of generalization, and the MAD for testing was 9.67 years.

TABLE 3: Results of four algorithms on a healthy dataset on training split.

Training	Healthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.80	5.24	7.43
Support Vector Regression	0.70	6.63	9.08
Multiple Linear Regression	0.71	6.78	8.97
Random Forest Regression	0.96	2.51	3.48
Best result	Random Forest Regressor		

TABLE 4: Results of four algorithms on healthy dataset on the unseen independent split.

Testing	Healthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.77	5.28	7.67
Support Vector Regression	0.72	5.83	8.47
Multiple linear regression	0.78	4.92	7.59
Random Forest Regression	0.78	5.02	7.49
Best result	Random forest regressor		

TABLE 5: Results on healthy dataset on training split after hyperparameter tuning.

Training	Healthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.84	4.96	6.69
Random Forest Regression	0.87	4.51	6.08

TABLE 6: Results on healthy dataset on testing split after hyperparameter tuning.

Testing	Healthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.76	5.32	7.84
Random Forest Regression	0.81	4.85	7.01

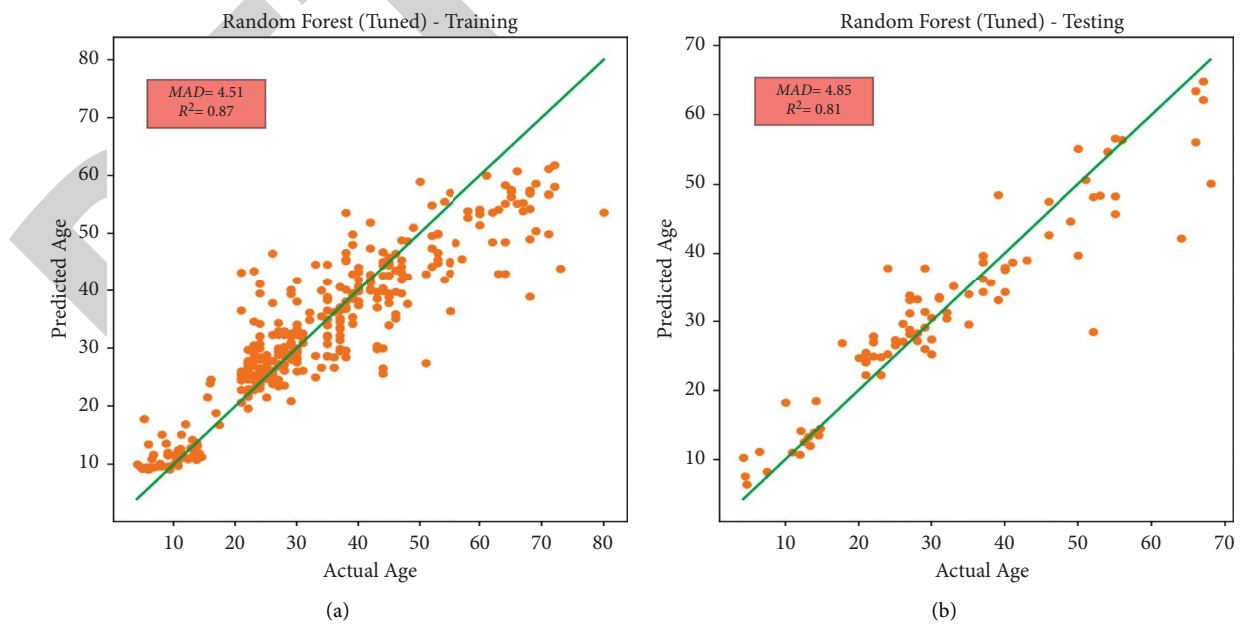


FIGURE 4: Results for healthy data with optimized Random Forest model: (a) training data; (b) testing data.

TABLE 7: Results of 4 algorithms on unhealthy dataset on training split.

Training	Unhealthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.75	8.0	10.68
Support Vector Regression	0.40	12.94	16.48
Multiple Linear Regression	0.56	11.49	14.10
Random Forest Regression	0.94	3.83	5.18
Best result	Random forest regressor		

TABLE 8: Results of four algorithms on unhealthy dataset on unseen independent split.

Testing	Unhealthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.53	10.40	13.45
Support Vector Regression	0.37	12.05	15.58
Multiple Linear Regression	0.46	11.52	14.40
Random Forest Regression	0.57	9.53	12.88
Best result	Random forest regressor		

TABLE 9: Results of unhealthy dataset on training split after hyperparameter tuning.

Training	Unhealthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.92	4.61	6.00
Random Forest Regression	0.92	4.75	6.18

TABLE 10: Results on unhealthy dataset on testing split after hyperparameter tuning.

Testing	Unhealthy dataset		
	R^2 -score	MAD	RMSE
Gradient Boosting Regression	0.62	10.28	13.17
Random Forest Regression	0.56	9.67	13.07

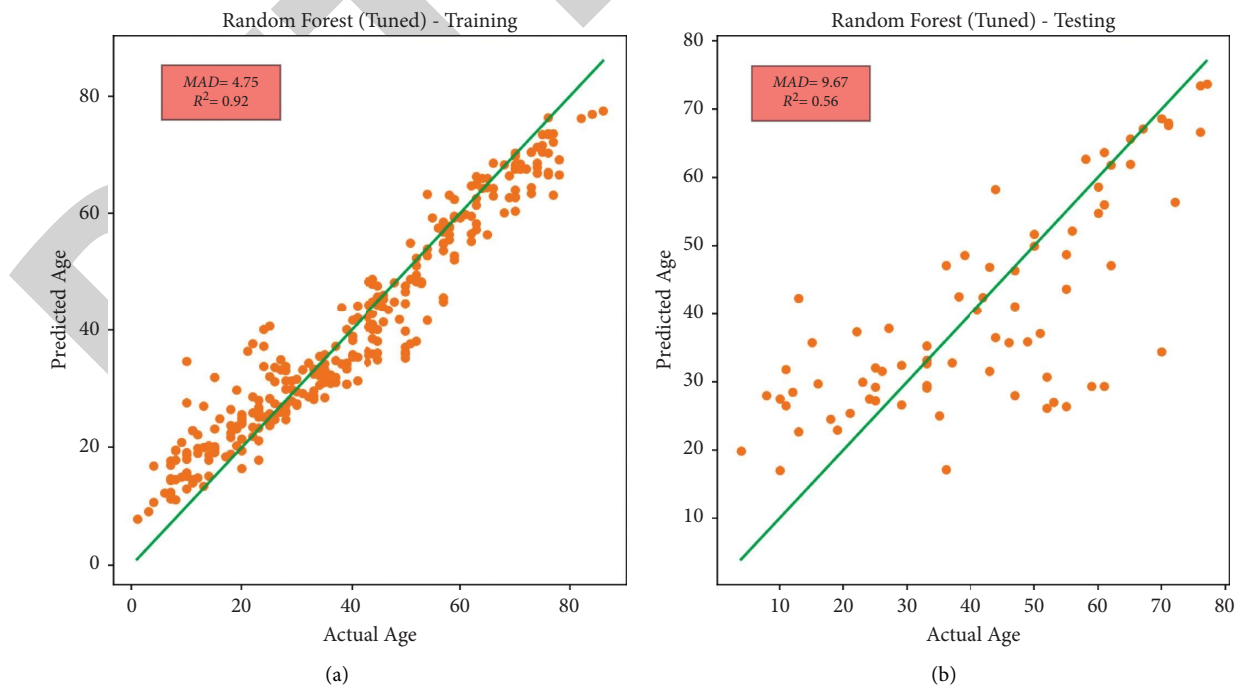


FIGURE 5: Results for unhealthy data with optimized Random Forest model: (a) training data; (b) testing data.

5. Conclusion

The utilization of DNAm data as the biomarker for the problems of age prediction is still a new growing research area that has gained significant attention from researchers all over the world. The present work has shown the potential of ensemble methods to create robust machine learning models that can be employed to predict human age based on DNAm data effectively. It has also been concluded that DNAm data are a good marker for predicting human age that can be used in forensics for medical investigation with a certain degree of assurance. From experimental results, it has been observed that the diseases that affected the human age adversely can easily be inferred by looking at the lower levels of correlation between DNA methylation markers and human age.

The work presented in this study has also shown a high degree of generalization, suggesting these models will be robust against unseen data samples. Our research provided strong evidence of how machine learning techniques can be used to predict human age from CpG data in an attempt to understand how the disease affects this correlation. In the future, the proposed work can be extended to the significance of diseases of human age; further investigation and research on demographic influence and gender effect on age can also be studied. Also, we are planning to include techniques like artificial neural network with more number of samples as training data in our future work.

Data Availability

Data and code can be made available upon reasonable request to the authors.

Consent

Not applicable.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by Taif University Researchers Supporting Project (TURSP-2020/114), Taif University, Taif, Saudi Arabia.

References

- [1] M. Dziechciaż and R. Filip, "Biological psychological and social determinants of old age: bio-psycho-social aspects of human aging," *Annals of Agricultural and Environmental Medicine: AAEM*, vol. 21, pp. 835–838, 2014.
- [2] N. Goel, P. Karir, and V. K. Garg, "Role of DNA methylation in human age prediction," *Mechanism of Ageing and Development*, vol. 166, pp. 33–41, 2017.
- [3] S.-E. Jung, K.-J. Shin, and H. Y. Lee, "DNA methylation-based age prediction from various tissues and body fluids," *BMB Reports*, vol. 50, no. 11, pp. 546–553, 2017.
- [4] S. Rodríguez-Rodero, J. L. Fernández-Morera, E. Menéndez-Torre et al., "Aging genetics and aging," *Aging Dis*, vol. 2, pp. 186–195, 2011.
- [5] C. H. E. Lau and O. Robinson, "DNA methylation age as a biomarker for cancer," *International Journal of Cancer*, vol. 148, no. 11, pp. 2652–2663, 2021.
- [6] M. Gasparetto, F. Payne, K. Nayak et al., "Transcription and DNA methylation patterns of blood-derived CD8+ T cells are associated with age and inflammatory bowel disease but do not predict prognosis," *Gastroenterology*, vol. 160, no. 1, pp. 232–244, 2021.
- [7] X. Li, W. Li, and Y. Xu, "Human age prediction based on DNA methylation using a gradient boosting regressor," *Genes*, vol. 9, no. 9, p. 424, 2018.
- [8] J. Naue, H. C. J. Hoefsloot, O. R. F. Mook et al., "Chronological age prediction based on DNA methylation: massive parallel sequencing and random forest regression," *Forensic Science International: Genetics*, vol. 31, pp. 19–28, 2017.
- [9] A. Vidaki, B. Daniel, and D. S. Court, "Forensic DNA methylation profiling-Potential opportunities and challenges," *Forensic Science International: Genetics*, vol. 7, no. 5, pp. 499–507, 2013.
- [10] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," *Comput Mater*, vol. 5, pp. 1–36, 2019.
- [11] A. Raza, S. Bardhan, L. Xu et al., "A machine learning approach for predicting defluorination of per- and poly-fluoroalkyl substances (PFAS) for their efficient treatment and removal," *Environmental Science and Technology Letters*, vol. 6, no. 10, pp. 624–629, 2019.
- [12] J. Zhang, L. Wang, L. Trasande, and K. Kannan, "Occurrence of polyethylene terephthalate and polycarbonate microplastics in infant and adult feces," *Environmental Science and Technology Letters*, vol. 8, no. 11, pp. 989–994, 2021.
- [13] P. Y. Lau and W. K. Fung, "Evaluation of marker selection methods and statistical models for chronological age prediction based on DNA methylation," *Legal Medicine*, vol. 47, Article ID 101744, 2020.
- [14] B. Liu, "BioSeq-Analysis: a platform for DNA, RNA and protein sequence analysis based on machine learning approaches," *Briefings in Bioinformatics*, vol. 20, pp. 1280–1294, 2018.
- [15] Z. Thong, J. Y. Y. Tan, E. S. Loo, Y. W. Phua, X. L. S. Chan, and C. K. C. Syn, "Artificial neural network, predictor variables and sensitivity threshold for DNA methylation-based age prediction using blood samples," *Scientific Reports*, vol. 11, 2021.
- [16] A. Aliferi, D. Ballard, M. D. Gallidabino, H. Thurtle, L. Barron, and D. Syndercombe Court, "DNA methylation-based age prediction using massively parallel sequencing data and multiple machine learning models," *Forensic Science International: Genetics*, vol. 37, pp. 215–226, 2018.
- [17] A. Vidaki, D. Ballard, A. Aliferi, T. H. Miller, L. P. Barron, and D. Syndercombe Court, "DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing," *Forensic Science International: Genetics*, vol. 28, pp. 225–236, 2017.
- [18] N. S. Jaddi and M. Saniee Abadeh, "DNA methylation-based age prediction using cell separation algorithm," *Computers in Biology and Medicine*, vol. 121, Article ID 103747, 2020.
- [19] S. R. Hong, K.-J. Shin, S.-E. Jung, E. H. Lee, and H. Y. Lee, "Platform-independent models for age prediction using DNA

Retraction

Retracted: Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. B. Alazzam, A. T. Al-Radaideh, N. Binsaif, A. S. AlGhamdi, and M. A. Rahman, "Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 1422963, 5 pages, 2022.

Research Article

Advanced Deep Learning Human Herpes Virus 6 (HHV-6) Molecular Detection in Understanding Human Infertility

Malik Bader Alazzam ¹, Ahmad Tawfig Al-Radaideh ², Nasser Binsaif ³,
Ahmed S. AlGhamdi ⁴ and Md Adnan Rahman ⁵

¹Faculty of Computer Science and Informatics, Amman Arab University, Amman, Jordan

²Department of MIS, College of Business, Jadara University, Irbid, Jordan

³Department of E-Commerce, College of Administrative and Financial Sciences, Saudi Electronic University, Riyadh, Saudi Arabia

⁴Department of Computer Engineering, Collage of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁵College of Business Administration (CBA), International University of Business Agriculture and Technology (IUBAT), Dhaka, Bangladesh

Correspondence should be addressed to Md Adnan Rahman; adnan.cba@iubat.edu

Received 16 November 2021; Revised 5 December 2021; Accepted 9 December 2021; Published 7 January 2022

Academic Editor: Deepika Koundal

Copyright © 2022 Malik Bader Alazzam 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.

To see if HHV-6 may be a cause of infertility, researchers looked at 18 men and 10 women who had unexplained critical fertility and had at least one prior pregnancy. HHV-6 DNA was discovered in both infertile and fertile peripheral blood mononuclear cells (PBMC) (12 and 14%, respectively); endometrial epithelial cells from 4/10 (40%) infertile women were positive for HHV-6 DNA; this viral DNA was not found in the endometrium of fertile women. When endometrial epithelial cells were cultivated, they produced viral early and late proteins, suggesting the existence of an infectious virus. Endometrial HHV-6 infection creates an aberrant NK cell and cytokine profile, resulting in a uterine domain that is not favorable to conception, according to the findings. To corroborate the findings, studies of extra fertile and barren women should be done. Semen samples were taken from 18 guys who visited the Government General Hospital Guntur's infertility department because they were having reproductive issues with their partners. Herpes virus DNA has been discovered in the sperm of symptomatic fertile and infertile male patients on rare instances. Furthermore, researchers must investigate the role of viral diseases in male infertility.

1. Introduction

After a year of unprotected sexual intercourse, 15% to 20% of couples in European nations experience infertility. In roughly 33% of infertility cases, both male and female factors are present, up to 49.5 percent, depending on the fairness of the viral DNA detection technique utilized. Human herpes virus 6 (HHV-6) is a beta herpes virus with a strong link to human cytomegalovirus. Were the first to isolate HHV-6 from peripheral blood lymphocytes (PBL) obtained from lymphoproliferative clutter patients. Cell tropism is particularly notable for T lymphocytes, which

was formerly dubbed human B-lymph tropic virus (HBLV). HHV-6A and HHV-6B are two hereditarily distinct strains of the virus. HHV-6 is a herpes virus that is related to HHV-7 and, to a lesser extent, human cytomegalovirus (HCMV). HHV-6 has been identified as a causative operator of exanthema subatomic (roseola), a common young exanthema, according to. HHV-6 illness is most common in the first two years of life. HHV-6 reactivates with immunological suppression, despite the fact that evident clinical illness is rare in adults. HHV-6 has been linked to a variety of human diseases, including multiple sclerosis (MS); however, the significance of these associations is unclear [1, 2].

The discovery of the HHV6 cell-surface receptor, as well as the confirmation of the hereditary groups of HHV-6A and-6B, is among the most recent discoveries in the subatomic science of HHV6. The discovery of the HHV-6 cell receptor, CD46, has provided yet another insight into HHV-6 cell tropism. In addition, the in vitro partnerships between HHV-6 and other viruses, particularly, the human immunodeficiency virus, and their significance for the in vivo situation, as well as the trans-activating limitations of a few HHV-6 proteins, are investigated. The clinical spectrum of HHV-6 is still being explored, and aside from being regarded as a significant pathogen among transplant recipients (as seen by the growing number of upcoming clinical tests), its role in focused sensory system disease is becoming increasingly obvious [3–5]. Finally, we offer a schematic depicting the many therapy options for HHV-6. The clinical spectrum of HHV-6 in immune-compromised people has occupied a large portion of the current literature on the virus. Reactivation of the beneficiary's strain, external illness in the contributor's strain, and reinfection with a different strain are all possibilities. In the bone marrow transplant population, a small number of individuals have been identified with interstitial pneumonitis caused by HHV-6, which is commonly associated with unit against host illness (GVHD) in the semen of asymptomatic infertile individuals [6].

2. Materials and Methods

2.1. Clinical Samples. Eighteen males who went to the reproductive center at the GGH Guntur, Andhra Pradesh (from January 2017 to December 2017), owing to couple reproductive concerns, provided samples (1 ml) to verify the HHV-6 and twelve blood serum samples to test the antibodies to herpes virus. Every patient gave their informed permission for the purposes of the current examination. There were no clinical symptoms of HSV infection in the men who took part in the study, and no evident reasons of infertility were discovered. They were all sterile for unknown reasons. Individuals with any underlying disorders, such as diabetes, that might potentially interfere with fertility were excluded from the study.

Ten female patients were engaged for tubal patency testing and endometrial samples were collected from them [7–9]. The participants in the study had to be between the ages of 28 and 36 and have a normal menstrual cycle (20–32 days) and a BMI of 17.5 to 25.5 kg/m². The menstrual cycles' secretory phase was planned (13 to 27 days). Tissue samples were obtained in HEPES-supplemented Dulbecco adjusted Eagle medium/Hams F-12 (DMEM/F-12).

2.2. Ethics Statement. A written consent was obtained from each patient, and ethics approval was obtained from the Ethical Committee of the medical centre.

All research studies involving human subjects necessitate the acquisition of research ethical approval. Before research subjects can be recruited or data collection can begin, this approval must be acquired.

2.3. Statistical Analysis. The data collected were analyzed with SPSS software version 15. The data were expressed as mean \pm SD. The differences of variables between two groups HHV (+) ve versus HHV (-) ve were compared using the Student's *t*-test.

2.4. Analysis of Semen and Blood. A sperm count, commonly known as a semen analysis, determines the number and quality of a man's sperm and semen. The thick, white fluid discharged from a man's penis at his sexual climax is known as sperm (orgasm). Ejaculation is the term for this process. Sperm is the genetic material-carrying cell in a man's body. When a sperm cell combines with an egg from a woman, an embryo (the earliest stage of the development of an unborn baby) is formed.

Masturbation in sterile containers was used to gather sperm samples after 48–72 hours of sexual abstinence. Within 1 hour of collection, samples were evaluated, and within 2 hours of collection, samples were processed for DNA extraction. Computer-aided sperm analysis was used to analyze the sperm concentrations and motility of 18 male patients, according to the 4th edition of the WHO Guidelines. Observation of eosin-thiazine-stained methanol-fixed smears of fresh ejaculate under a light microscope and evaluation according to rigorous Krueger criteria were used to assess sperm morphology. Semen specimens that matched all of the aforementioned criteria were considered normal. The presence of an unusually large number of white blood cells did not indicate that the sample was abnormal. All sperm samples with one or more abnormal semen values were considered abnormal [10–14].

2.5. PCR Amplification Reactions and Extraction of DNA from Semen Samples. The molecular laboratory's core tools include DNA extraction and polymerase chain reaction (PCR). This brief overview covers the numerous physical and chemical procedures for extracting DNA in order to obtain adequate quantities of high-quality DNA. PCR can be used to increase the amount of DNA in a sample.

To avoid contamination, all PCR test preparations were done in a "clean room" (no post-PCR DNA products) underneath a laminar flow hood. Each fresh semen sample was centrifuged for 30 minutes at 13,000 rpm after collection. The spermatozoa pellet was kept in the original Eppendorf tube, and the precipitate (seminal fluid) was transferred to a fresh one. In the event of a semen PCR test, sample DNA was collected using a DNA isolation from fluid Semen kit (QIAGEN; Cat no: 57704) according to the manufacturer's instructions. All participants' spermatogenesis and seminal fluids were tested for the presence of HHV-6 DNA using PCR with the appropriate set of primers (Table 1). Agarose gel in a 2 percent or 3 percent gel matrix, depending on the size of the PCR result, and then photography with an ultraviolet light transilluminator were used to examine the PCR results. Finally, to corroborate the initial amplification results, direct sequencing analysis of HHV 6 PCR products was done (Figure 1). The PCR products were purified using the QIAquick PCR purification Kit (Qian

TABLE 1: Demographic and clinical characteristics of women's cohorts (average values).

Parameters	Infertile (10)	Control (14)	<i>P</i> value*
Age (yrs)	33.4	33.8	0.93
Duration of infertility (yrs)	2.8	2.1	0.63
Length of menstrual cycle (days)	4.6	4.3	0.43
Follicle-stimulating hormone (FSH) (mUI/mL) (day 3)	9.1	7	0.41
Luteinizing hormone (LH) (mUI/mL) (day 3)	6.7	6.2	0.41
Estradiol (pg/mL) (day 3)	76.8	65.7	0.05
Thyroid-stimulating hormone (TSH) (uUI/mL)	2.4	1.9	0.44
Free thyroxine (FT4) (pg/mL)	1.8	1.1	0.85
Progesterone (pg/mL) (day 21)	11.67	13	0.53
Day (menstrual cycle) of sample collection	13.8	13	0.80

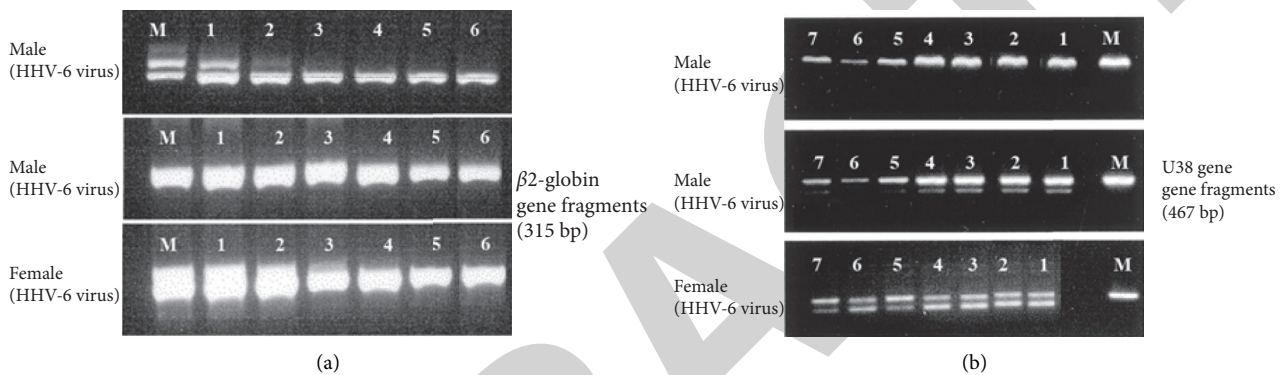
Student *t*-test.

FIGURE 1: PCR amplification for 2-globin gene (a) fragment (315 bp) and for U38 gene fragments (467 bp) (b).

Valencia, CA) and then eluted with water [7]. DNA sequences were examined and identified by uploading them to the ribosome database project website [2].

3. Results and Discussion

3.1. HHV-6 in Women's Clinical Specimens. A whole blood quantitative PCR DNA test for HHV-6 is the most feasible way to find out if someone has chromosomally integrated HHV-6 (ciHHV-6). If the amount is greater than 500,000 copies per ml without an acute illness, the person has ciHHV-6.

In the current clinical study, ten women with unexplained primary infertility who had no past pregnancies or pathological variables were compared to 14 bigger women who had at least one previous successful pregnancy. Table 1 shows no significant differences between the two groups, save for a small increase in estradiol levels in infertile males ($P = 0.05$; Student's *t*-test). Estradiol levels over a specific level may indicate that the hormone is suppressing the production of follicle-animating hormone (FSH) and that there are worries about richness. The amount of progesterone in the body did not change in any way. Table 2 summarizes the outcomes of each clinical case based on whether or not HHV-6 infection was present. In one of the individuals' PBMCs, HHV-6 DNA was detected. Surprisingly, HHV-6 viral DNA was found in 44 percent of PCOS women's endometrial epithelial cells, but not in control

women's cells ($P = 1.27 \times 10^{-4}$; Fisher correct test). Similarly, no HHV-6 DNA and HHV-6A infection is restricted to the endometrial epithelium of primary infertile women, according to stromal cells recovered through epithelial cell purification (Table 2). Because HHV-6 dynamic illness is common in a group of infertile women, search for sharp disparities in markers, based on HHV-6 contamination (Tables 3 and 4). HHV-6 positive infertile women had greater estradiol levels than HHV-6 negative infertile women ($P = 0.045$) (see Table 3, for further information.). Estradiol irregularities may have repercussions in HHV-6 infection, based on the discovery of a relationship. The proximity of HHV-6 infection was linked to estradiol dimensions ($P = 0.001$; $R^2: 0.89$; Spearman relationship). There were no clinical indicators that differed between the two spouses (Table 3). When we looked at immunological properties in the uterine state, we saw a shift in endometrial (e) NK cell-safe phenotype and cytokine levels. HHV-6 positive infertile women (8.06) had less CD56posCD16neg eNK cells than HHV-6 negative infertile women (23.82) ($P = 0.001$) (for further information, see Table 4). CD56bright and CD56dim eNK cells were distributed differently in HHV-6A positive and negative infertile women, with HHV-6 positive infertile women (3.72) having less CD56bright and CD56dim eNK cells than HHV-6 negative childless women (13.68) ($P = 0.001$) (for further information, see Table 4). Although there was no difference in the number of CD3+ lymphocytes between the two groups,

TABLE 2: Infertile women were split into two groups based on whether or not they had HHV-6 infection: infertile and control women's demographic and clinical characteristics (average).

Parameters	HHV-6 positive	HHV-6 negative	<i>P</i> value *	Control (14)
Age (yrs)	33.4	32.8	0.84	33.8
Duration of infertility (yrs)	2.9	2.8	0.56	2.1
Menstrual cycle length (days)	4.8	4.6	0.38	4.3
FSH (mUI/mL) (day 3)	9.5	9.2	0.37	7.0
LH (mUI/mL) (day 3)	7.0	6.8	0.37	6.2
Estradiol (pg/mL) (day 3)	79.9	77.5	0.04	65.7
TSH (uUI/mL)	2.5	2.4	0.39	1.9
FT4 (pg/mL)	1.9	1.8	0.76	1.1
Progesterone (pg/mL) (day 21)	12.1	11.8	0.48	13.0
Day (menstrual cycle) of sample collection	13.2	12.8	0.72	13.0

Student *t*-test.

TABLE 3: Infertile women were split into two groups based on whether or not they had HHV-6 infection: endometrial samples immunological parameters.

Immune cells	HHV-6 positive	HHV-6 negative	<i>P</i> value*	Control
NK CD56posCD16neg (N)	8.06	23.82	0.001	24.78
NK CD56brightCD16neg (N)	3.72	13.68	0.001	14.89
NK CD56dimCD16- (N)	4.44	10.18	0.017	9.87
CD3+ (N)	0.00	0.00	0.000	0.00
CD14+ (N)	0.46	0.52	0.612	0.48

Student *t*-test.

TABLE 4: Mean sperm count and motility in virally infected and noninfected semen samples.

	Number of specimens	Sperm count (million/ml)	<i>P</i> value	Motility (mean)	<i>P</i> value
Viral DNA positive	9	27.67	0.71	38.53	0.67
Viral DNA negative	12	30.24	0.67	61.47	0.61
HHV-6 positive	16	29.67	0.91	70.14	0.5
HHV-6 negative	2	27.47	0.14	29.86	0.37

there was a difference in the number of CD14+ monocytes ($P = 0.612$). The uterine flushing aspects of cytokines presented an alternate example in both HHV-6 positive and negative infertile women [3–6].

As a possible environmental cause of human infertility, viral infections have been connected. Although no single virus has been confirmed to induce female infertility, herpesviruses have been associated to male infertility in particular. In our study, furthermore, endometrial epithelial cells from infertile women exhibited high viral loads (about 4 copies of viral DNA per cell); no HHV-6A infection was found in stromal cells or PBMC, excluding the possibility of HHV-6 DNA being chromosomally integrated.

The frequency of the herpes virus in normal and abnormal sperm samples is shown in Table 4. According to the 4th edition WHO criteria, 9 of 18 sperm samples were classed as normal (normozoospermia). HHV6 viral DNA was found in 6 (33.3 percent) of normal sperm samples (Table 3). In 18 male individuals with aberrant zoospermia, viral DNA was also found (Table 5). The samples were further categorized into six oligozoospermic, five asthenozoospermic, three oligoasthenozoospermic, and four teratozoospermic subgroups in this group. In aberrant semen samples, HHV6 viral DNA was found in 4 (22.22

percent), 8 (44.44 percent), 2 (11.11 percent), 3 (16.67 percent), and 3 (16.67 percent) (Table 5). There was no significant statistical difference between the existence of each herpes virus and each aberrant grouping, according to statistical analysis. For at least one member of the herpes virus family, viral DNA was found by PCR in 16 (88.88 percent) of the 18 total semen samples (Table 5). The findings of viral DNA analysis for all 18 samples provided show that HHV-6 is present in 16 of them (88.88 percent). Previous research looked at the frequency of viral DNA in total sperm samples.

The mean sperm count and motility of virally infected and no infected semen samples are shown in Table 4. The mean sperm count after HHV-6 infection had no effect on mean sperm count. Kapranos et al. discovered that HHV1 infection was associated with decreased sperm count and motility. Previous research had revealed no link between HHV-6+ and HSV-6- DNA and poor sperm count, motility, or infertility.

3.2. The Difference between Normal and Abnormal Semen. Statistical analysis revealed no link between the presence of the herpes virus and the subsequent classification of samples as normal or abnormal semen, indicating that viral presence

Retraction

Retracted: Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] M. B. Alazzam, H. Mansour, F. Alassery, and A. Almulih, "Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 5759184, 7 pages, 2021.

Research Article

Machine Learning Implementation of a Diabetic Patient Monitoring System Using Interactive E-App

Malik Bader Alazzam ¹, Hoda Mansour ², Fawaz Alassery ³, and Ahmed Almulihi ⁴

¹Faculty of Computer Science and Informatics, Amman Arab University, Amman, Jordan

²College of Business Administration, University of Business and Technology, Jeddah, Saudi Arabia

³Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif, Saudi Arabia

⁴Department of Computer Science, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

Correspondence should be addressed to Malik Bader Alazzam; m.alazzam@aau.edu.jo

Received 27 October 2021; Revised 6 November 2021; Accepted 22 November 2021; Published 31 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Malik Bader Alazzam 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.

Lifestyle influences morbidity and mortality rates in the world. Physical activity, a healthy weight, and a healthy diet are key preventative health behaviours that help reduce the risk of developing type 2 diabetes and its complications, such as cardiovascular disease. A healthy lifestyle has been shown to prevent or delay chronic diseases and their complications, but few people follow all recommended self-management behaviours. This work seeks to improve knowledge of factors affecting type 2 diabetes self-management and prevention through lifestyle changes. This paper describes the design, development, and testing of a diabetes self-management mobile app. The app tracked dietary consumption and health data. Bluetooth movement data from a pair of wearable insole devices are used to track carbohydrate intake, blood glucose, medication adherence, and physical activity. Two machine learning models were constructed to recognise sitting and standing. The SVM and decision tree models were 86% accurate for these tasks. The decision tree model is used in a real-time activity classification app. It is exciting to see more and more mobile health self-management apps being used to treat chronic diseases.

1. Introduction

Diabetes is a life-altering illness that some people find difficult to cope with, which can lead to the emergence of depressive symptoms. People with diabetes are two to three times more likely than nondiabetics to suffer from depression, and only about 25 to 50 percent of those suffering from depression will be detected and treated [1]. Depression can be brought on by diabetes management obligations, but it can also be brought on by outside stressors. Factors like financial stress, food hardship, family separation, and immigration can cause depression in people with Hispanic ancestry who have type II diabetes. A patient may be reluctant to seek therapy because of cultural pressures such as the negative stigma associated with mental health in the Mexican culture. 7.3 percent of Hispanic people used mental

health services in 2008–2012, according to the National Survey on Drug Use and Health. This compares to white adults, who used mental health services at a rate of 16.6 percent (Substance Abuse and Mental Health Services Administration) [3]. If a person's diabetes care is neglected because of depression, it can lead to binge eating, missed doctor visits, and social isolation, all of which can worsen a person's health even more.

In order to perceive, feel, and act effectively, people need to be healthy. This is why health is important not just for the development of the individual but also for the environment to which they belong. The World Health Organization (WHO) defines health as "a state of complete physical, mental, and social well-being, and not merely the absence of disease or disability." A healthy body is always changing and adapting to the pressures and changes in the environment,

according to WHO, in order to maintain a balance within called homeostasis.

Taking care of one's health or one's well being by diagnosing, treating, and preventing sickness or other physical or mental impairments is known as healthcare. There is a study conducted in presenting primary, secondary, and tertiary care as well as in public health that is included [4]. In order to validate the appropriate healthcare distribution based on parameter monitoring and direct provisioning of scientific aid, it is critical to provide sufficient methods and tokens.

Information and communication technologies (ICTs) are being adopted by industries across the globe to improve transportation and increase competitiveness. This rule applies to all aspects of life, including health treatment. ITs have the potential to fundamentally alter the way healthcare is provided and managed. eHealth refers to the application of information and communication technologies (ICTs) in the delivery of fitness care services. Digital verbal interchange and information science are used in the health zone for clinical, educational, and administrative reasons at both the local site and at a distance according to the WHO definition of eHealth [5]. To put it another way, eHealth is the capability of ensuring that the correct fitness statistics are provided to the correct person at the correct time and place in a secure electronic structure in order to optimise fitness care delivery, research, training, and knowledge satisfaction and effectiveness.

Using the newest breakthroughs in information technology, remote tracking and monitoring collect patient data outside of the traditional healthcare setting. Simple remote patient monitoring systems that make use of consumer-grade equipment are the most effective. Most health monitoring systems make use of technology that is designed to be comfortable for patients [6].

2. Background

Diabetes self-management can be greatly improved with the use of smartphones and modern information technology. Exciting results from Kirwan and colleagues' research show that a diabetes self-management mobile app can reduce HbA1c levels in patients [7]. However, medical professionals, academics, and app developers must collaborate to create a better diabetes self-management mobile app, with consideration for the target audience's age [8].

Controlling blood sugar at a healthy level is difficult for diabetics for a variety of reasons. As a result of the increased risk of cardiovascular, renal, and neurological illnesses that are associated with high blood sugar levels [9], a study conducted by Basnet et al. [10] looked into the impact on type 1 diabetics using a smartphone app for diabetes self-management. Patients from Diabetes Australia in New South Wales and Queensland nominated participants.

Patients with type 1 diabetes who have had their diagnosis for at least six months and have glycosylated haemoglobin levels (HbA1c) greater than 7.5 percent must be eligible. Participants must also have access to a smartphone.

Figure 1 shows an existing method based on the cloud system.

There were two groups of participants. Using an iPhone app for diabetes self-management and help from a Certified Diabetes Educator (CDE) using the data generated by the app is one option for intervention. The other is a control group that continues to get the same level of treatment as before [11]. It is called "Glucose Buddy," and it has useful features like letting you manually enter your blood glucose levels, insulin dosages, nutrition, and viewing a personalised health information graph.

It is still a new field of healthcare, even though mobile health (mHealth) offers so many advantages to patients and healthcare providers alike. Many difficulties and challenges are still unresolved or unclear. There are many issues that need to be addressed, such as how reliable and efficient the service will be, how good the quality of the service will be, and how full the promised functionality will be for patients. In addition, as mHealth handles private information from patients, the topic of privacy security comes up frequently [12].

3. Related Work

Patients with diabetes who have been diagnosed by a doctor are said to have morbidity, but diabetics who have not been diagnosed do not. Adults aged 20 and over had a morbidity rate of 15% between 2013 and 2016, with Hispanic or Latino people of Mexican heritage having a morbidity rate of 19%. 83,564 people died in 2017 as a result of diabetes in the United States, or 25.7 fatalities per 100,000 persons, according to a National Center for Health Statistics report released in 2019 [13]. It is estimated that the Hispanic population has an average life expectancy of 81.8 years, which has remained stable over time due to underreporting because they are the most uninsured of any racial group and less likely to seek medical attention.

On the traditional and cloud computing environments, an overview of the IERF framework is shown in Figure 2 to improve classification accuracy and provide improved resource allocation. The University of California at Irvine (UCI) website was used to collect data for testing. There are two diabetes datasets included in this study: one from the Pima Indian community in Arizona and the other from the United States [14]. There is a description of the characteristics provided. The PID dataset includes 768 instances with 8 different attributes. It was given to the university as a gift in 1990.

4. Methodology

It combines the advantages of the genetic algorithm and rough set theory's relative reduct algorithm in a single model. Conditional and decision qualities are the two categories into which the attributes fall. The value of the variable R is set to 0 at the beginning of the programme. And the value 0 is assigned to the variable best. T contains R and another temporary variable, tmp, and contains the value that is now the best. Each attribute is removed one at a time, and the

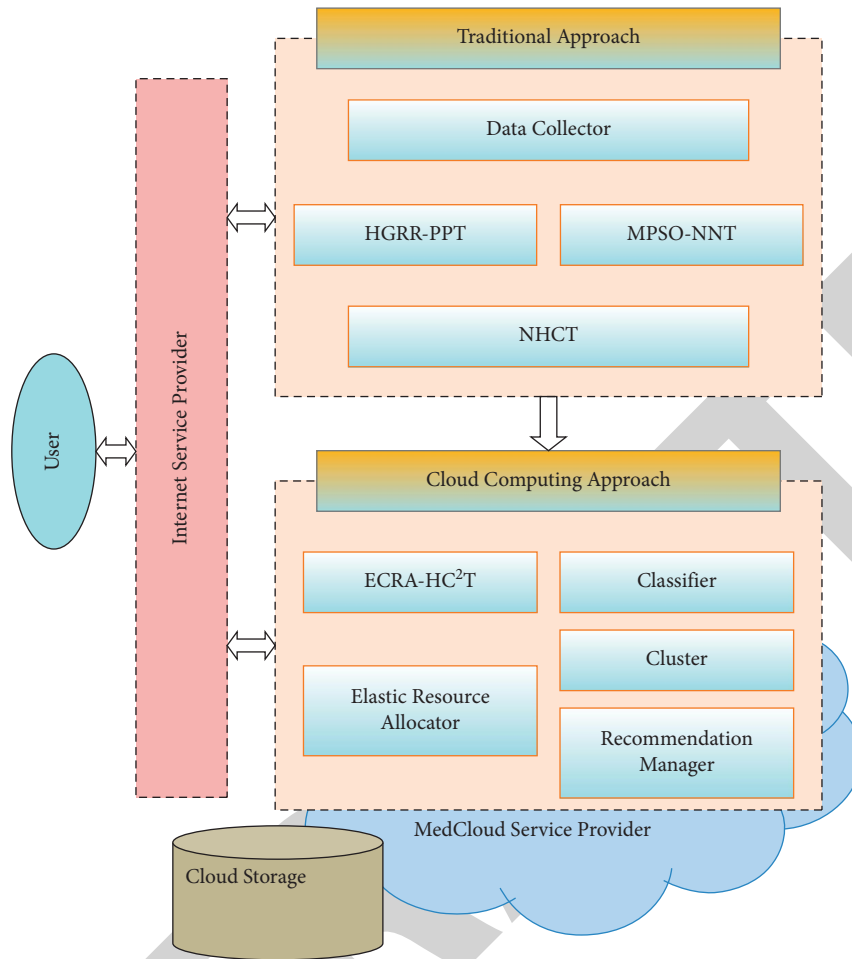


FIGURE 1: Overview of the existing method based on the cloud system.

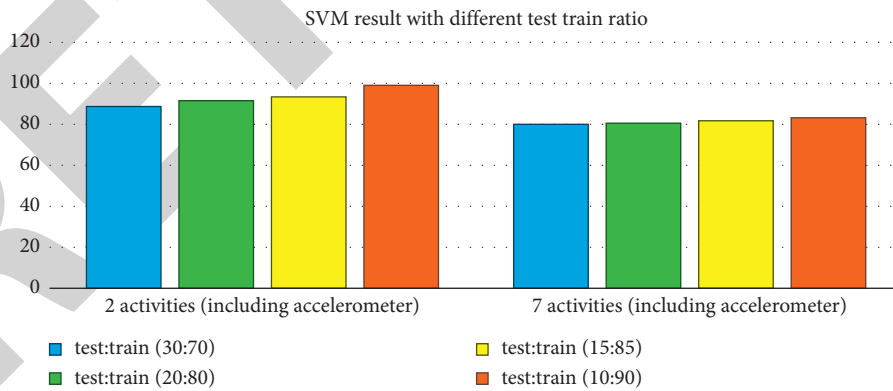


FIGURE 2: Two and seven SVM results on four test train ratios.

dataset is evaluated for consistency in the decision table. Attributes are eliminated, and the smaller dataset is saved if the decision table is consistent. Once the classification accuracy of the conditional attributes has been verified, the result is compared to the decision attributes [15]. The first generation of children is generated if the classification accuracy of the conditional attributes acquired is higher than the

classification accuracy of the decision attributes. Attributes are chosen from among the resulting children. The traits that have been chosen are subject to mutation and cross over.

T holds the resulting reduct set. A selection of the best decision attributes is kept in a variable called best. T gives R the reduct set it received, and R uses it. Once the optimal dataset is found, the process is repeated indefinitely. The

algorithm comes to an end once the optimal set of data has been obtained. The reduct set R is the final result.

After collecting and preprocessing the data, the dataset will be trained using some supervised learning algorithms like KNN, decision trees, and ANN. Then, the accuracy of each model's predictions will be compared, and the best model will be chosen to implement on the dataset in order to predict future occurrences [16].

There should be the provider side's key objective to create a straightforward approach for providers to monitor the education and communication with their patients from many angles. The features on the provider's end give medical professionals a slew of useful options [17–21]. As a result, patients can benefit from high-quality services and successful therapy. With smartphones as the primary device of both health care providers and patients, both sides of the equation require an internet connection. A user's data is transferred to the backend server when they utilise services such as adding a meal or uploading their blood pressure. The data are then returned to the frontend server when the server has finished processing it.

Clients and servers can exchange data by utilising the Hypertext Transfer Protocol (HTTP). Next, the server will handle HTTP requests and use SQL to query the database (Structured Query Language). Our applications frequently make use of GET, POST, PUT, and DELETE methods. Use the GET function to get data from a particular resource. Figure 3 shows an example of a client using the GET method to receive the day's meal information in the second phone screen capture. The server receives the GET service and processes it on its own before using SQL to retrieve the relevant data from our database's Meal table and present it to the client. The data are created and updated using POST and PUT methods. Think about when people add or change their daily meals, blood pressure, or sugar levels. The POST or PUT method is used to send data from the client to the server. With POST, you can produce data numerous times, however, with PUT, you can only make one copy of the same thing. That is the sole difference between them. If you are not familiar with SQL, the DELETE command is simpler to grasp because it simply removes specified data from the database.

5. Results

SVM and decision tree models used ratios of 10 : 90, 15 : 85, 20 : 80, 25 : 75, and 30 : 70, respectively. The decision tree and support vector machine models were tested on several ratios of tests to training data. Finally, two models were built. A sensor and accelerometer data produced it. For each sensor, the mean and median time differences were used to measure the accuracy of sensor-based activity detection. These 16 sensor data features correctly detected activity. To detect non-foot-movement activities, only sensor-based data were employed, but no accelerometer data were employed (e.g., sitting).

Various test train ratios were attempted during model training to obtain the most accurate model. 15% and 20% test to train yielded the best results (accuracies of 90–100

percent). Two sets of training sessions were used to build the machine learning models: seven activities employing sensor data and accelerometer.

The system correctly detected walking and running in 31 situations. One walking sample and twenty-six running samples were taken. Each test train ratio had five classification model observations on average.

Figures 4 and 5 show the results for four test train ratios (20, 5, 10, and 90). The decision tree and support vector machine models were put to the test using different test-training data ratios. In Figures 6 and 7, it is possible to state that two models were created. They were created using data from a sensor and an accelerometer. Each sensor has its own set of parameters. The accuracy of accelerometer and sensor data categorising walking and running was 96%. In 31 cases, the technology identified walking and running, one walking and 26 running samples.

The 15 sensor readings were 80% accurate.

Three of the seven samples correctly identified stair descent, three stair climbings, and one walking. The model had no false positives but missed several stair descending samples (1 : 0.43) for elliptical use. The samples were labelled as standing, walking, and sitting. Precision-recall ratio was 0.80 : 0.76 for non-elliptical use.

That means, the precision-to-recall ratio is 0 to 0. Because sitting does not create pressure sensor data, the analysis relies on accelerometer data. Without accelerometer data, sedentary movement cannot be recognised, as shown in Table 1.

The test data generated four decision-tree models. Among all 43 features, these models had the best walking and running detection results (100%). Activities detected with 90% to 100% accuracy with 43 characteristics. Using only sensor data, the 2-activity and 7-activity models obtained accuracy rates of 80–90% and 29%, respectively. Figure 6 shows the investigation's results integrated with a decision tree.

Based on model quality assessments, BeticTrack incorporated the MLM created from all 43 features. The model was then put to the test in real time to classify things. This categorisation was based on time. For example, a person performed the seven tasks over time, and the classification result was saved in the database. We tallied the total time spent on each activity to ensure that the categories were accurate. Figure 7 shows the classification after a two-minute stroll. The MLM classed most walking as elliptical (over the course of 1 minute and 45 seconds). This is a 10 second stroll upstairs.

The model correctly identified standing data, as shown in Figure 7, with a 4.1% error rate. This could be due to the dataset's size. The model's classifications could be enhanced by collecting more data from more people over time.

The shoe sole distinguishes between sitting and running with near-perfect precision. Sitting has no sensor data compared to the other 7 activities. During application execution, running data contain the truest events. If the sensor feels pressure, it accepts it. It simply has two values: 0 and 1. If pressed on the ground while running, it returns 1. For example, standing, walking, or using stairs yields zero data

[15 0 0 0 0 2]				
[0 28 0 0 0 0]				
[0 0 18 0 0 0]				
[0 0 3 10 1 0 1]				
[0 0 0 0 6 0 0]				
[1 0 0 0 3 3 0]				
[4 0 0 3 0 0 22]				
	precision	recall	F1-score	Support
1	0.75	0.88	0.81	17
2	1.00	1.00	1.00	28
3	0.86	1.00	0.92	18
4	0.77	0.67	0.71	15
5	0.60	1.00	0.75	6
6	1.00	0.43	0.60	7
7	0.88	0.76	0.81	29
Avg total	0.87	0.85	0.85	120
Accuracy of Linear kernel SVM : 0.850				

FIGURE 3: Analysis of accelerometer and sensor data for seven activities.

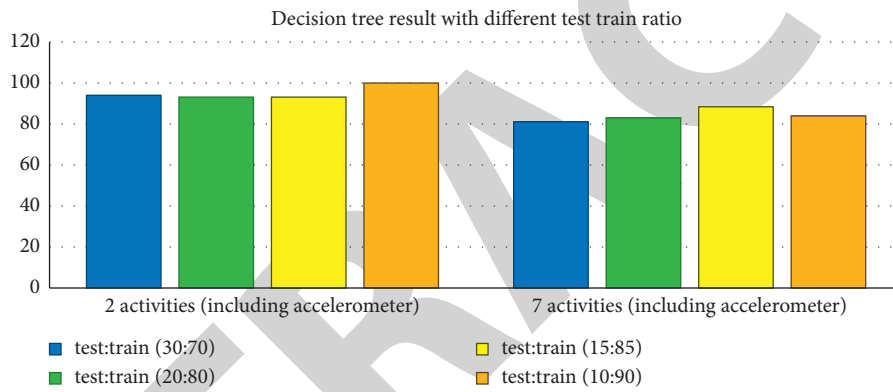


FIGURE 4: Four test train ratios with two and seven activity decision trees.

[31 1]				
[1 26]				
	Precision	Recall	F1-Score	Support
1	0.97	0.97	0.97	32
2	0.96	0.96	0.96	27
Avg/Total	0.97	0.97	0.97	59
Accuracy of Linear Kernel SVM: 0.966				

Activity	Walking	Running
Walking	31	1
Running	1	26

FIGURE 5: Classification of walking and running using accelerometers.

walking	Minutes: 3	Seconds: 14
running	Minutes: 0	Seconds: 0
sitting	Minutes: 4	Seconds: 8
standing	Minutes: 0	Seconds: 0
stairup	Minutes: 0	Seconds: 0
stairdown	Minutes: 0	Seconds: 0
elliptical	Minutes: 2	Seconds: 17

FIGURE 6: 4 min sitting, 3 min standing, and 2.5 min walking.

Activity	Error rate
Walking-----	2.3255813953488373
Running-----	3.0303030303030303
Sitting-----	3.2
Standing-----	4.123711340206185
Stair ascension-----	2.898550724637681
Stair descension-----	2.272727272727273
Elliptical-----	3.592814371257485

FIGURE 7: Decision tree analysis of error rate.

TABLE 1: Result of decision tree analysis.

No. of activities	Used accelerometer data	Classification accuracy
2	No	80%–90%
2	Yes	100%
7	No	29%
7	Yes	80–90%

from the pressure sensor. That is because the classification attempt failed. The use of elliptical data and classification of walking data overlaps. Eliminating elliptical data from classification may improve the classification of walking activity because walking and elliptical use are similar. This study's purpose was to collect data for five seconds straight to build features. Additional research can help construct a better categorisation model by identifying the best time period for this research.

6. Conclusion

The Patient and physician data usage are ensured through this study. There are many diabetes management smartphone apps. However, we built the system based on physician feedback and tried to precisely gather all diabetic treatment data. Patient self-management tools and doctor remote patient monitoring were the goals of this study. Wearable technology was employed to capture patient activity data. The wearable IoT insole records activity data with no further user effort. App learns sitting, walking, and running using SVM and decision trees. Mobile health and self-management solutions provide remote monitoring and contact without regard to time, location, or expense. The herb also aids diabetes. Each day's meal can be entered and reviewed by the patient. It enables patients to self-manage and track their development. These tools assist patients comprehend the importance of their choices. The long-term goal of this research is to test the application's effectiveness on type 2 diabetic patients. Based on trial results and patient feedback, future designs can be improved. Depending on the results of a pilot study with real patients, this application may become a standard tool for diabetic care. Assuming it works, this model could potentially treat other chronic diseases.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

Acknowledgments

We deeply acknowledge Taif University for supporting this study through Taif University Researchers Supporting Project Number (TURSP-2020/344), Taif University, Taif, Saudi Arabia. It was performed as a part of the Employment of Institutions.

References

- [1] M. Usman, K. Ahmad, A. Sohail, and M. Qaraqe, "The diabetic buddy: a diet regulator and tracking system for diabetics," in *Proceedings of the 2021 International Conference on Content-Based Multimedia Indexing (CBMI)*, pp. 1–4, Lille, France, June 2021.
- [2] R. Bateja, S. K. Dubey, and A. Bhatt, "Leveraging latest developments for delivering patient-centric healthcare to diabetic patients," in *Proceedings of the 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pp. 1201–1205, Noida, India, June 2020.
- [3] A. A. Hamad, A. S. Al-Obeidi, E. H. Al-Taiy, and D. Le, "Synchronization phenomena investigation of a new non-linear dynamical system 4d by gardano's and lyapunov's methods," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3311–3327, 2021.
- [4] L. M. Thivagar, A. A. Hamad, and S. G. Ahmed, "Conforming dynamics in the metric spaces," *Journal of Information Science and Engineering*, vol. 36, no. 2, pp. 279–291, 2020.
- [5] L. Wang, D. Jones, G. J. Chapman et al., "A review of wearable sensor systems to monitor plantar loading in the assessment of diabetic foot ulcers," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 7, pp. 1989–2004, 2020.
- [6] A. Tirkey and A. Jesudoss, "A non-invasive health monitoring system for diabetic patients," in *Proceedings of the 2020 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1065–1067, Chennai, India, July 2020.
- [7] D. Wang, J. Ouyang, P. Zhou, J. Yan, L. Shu, and X. Xu, "A novel low-cost wireless footwear system for monitoring diabetic foot patients," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 15, no. 1, pp. 43–54, 2021.
- [8] S. R. Benoit, I. Hora, F. J. Pasquel, E. W. Gregg, A. L. Albright, and G. Imperatore, "Trends in emergency department visits and inpatient admissions for hyperglycemic crises in adults with diabetes in the US, 2006–2015," *Diabetes Care*, vol. 43, pp. 1057–1064, 2020.
- [9] G. S. Kukreja, A. Alok, A. K. Reddy, and R. Nersisson, "IoT based foot neuropathy analysis and remote monitoring of foot pressure and temperature," in *Proceedings of the 2020 5th International Conference on Computing, Communication and Security (ICCCS)*, pp. 1–6, Patna, India, October 2020.
- [10] S. Basnet, R. Musaitif, A. Khanal et al., "Effect of potassium infusions on serum levels in children during treatment of diabetic ketoacidosis," *Journal of Pediatric Intensive Care*, vol. 9, pp. 113–118, 2020.
- [11] A. Usman, N. Mustafa, J. A. Dujaili, M. M. Bakry, and S. H. Gan, "Prevention of hypokalemia induced adverse cardiovascular effects in diabetic ketoacidosis: novel role of

Retraction

Retracted: A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. S. Al-Obeidi, S. Fawzi Al-Azzawi, A. Abdullah Hamad, M. L. Thivagar, Z. Meraf, and S. Ahmad, "A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3081345, 11 pages, 2021.

Research Article

A Novel of New 7D Hyperchaotic System with Self-Excited Attractors and Its Hybrid Synchronization

Ahmed S. Al-Obeidi,¹ Saad Fawzi Al-Azzawi,² Abdulsattar Abdullah Hamad ³,
M. Lellis Thivagar ³, Zelalem Meraf ⁴, and Sultan Ahmad ⁵

¹Specialty of Mathematics, Gifted School of Nineveh, Directorate of Education, Mosul, Iraq

²Department of Mathematics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

³School of Mathematics, Madurai Kamaraj University, Madurai, Tamilnadu, India

⁴Department of Statistics, Injibara University, Injibara, Ethiopia

⁵Department of Computer Science, College of Computer Engineering and Sciences, Prince Sattam Bin Abdulaziz University, P.O.Box. 151, Alkharj 11942, Saudi Arabia

Correspondence should be addressed to Zelalem Meraf; zelalemmeraf@inu.edu.et

Received 29 November 2021; Accepted 17 December 2021; Published 30 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Ahmed S. Al-Obeidi 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 study, a novel 7D hyperchaotic model is constructed from the 6D Lorenz model via the nonlinear feedback control technique. The proposed model has an only unstable origin point. Thus, it is categorized as a model with self-excited attractors. And it has seven equations which include 19 terms, four of which are quadratic nonlinearities. Various important features of the novel model are analyzed, including equilibria points, stability, and Lyapunov exponents. The numerical simulation shows that the new class exhibits dynamical behaviors such as chaotic and hyperchaotic. This paper also presents the hybrid synchronization for a novel model via Lyapunov stability theory.

1. Introduction

In 1963, Lorenz introduces the first known system of the 3D chaotic model, which has just one positive Lyapunov exponent and two quadratic nonlinearities. Subsequently, Rössler introduced another 3D chaotic model in 1976 which also includes seven terms, with one quadratic nonlinearity. Several well-known paradigms of the 3D chaotic models are chaotic Chua's circuit, Liu model, and the Pan model [1–10].

In 1979, the first four-dimensional (4D) model with two positive Lyapunov exponents (LEs) including real variables is performed by Rössler, and various 4D hyperchaotic models have been discovered in the previous works. These models are distinguished to own two +ve LEs and the dimension of the hyperchaotic model is related to the number

of +ve LEs so that the minimum dimension for the hyperchaotic model is four. To increase the number of +ve LEs, the dimension of the model must be increased. Recently, there is great interest in construction of 5D models with three +ve LEs as the hyperchaotic Hu model 2009 [11, 12].

Due to its increased unpredictability and randomness, the chaotic model with a higher dimension is beneficial compared to the low dimension and has a superior performance compared to the standard 3D, 4D, and 5D models. To date, only a few studies on the subject have been increased, and many articles have been dedicated to the construction of new high-dimensional (6D) models with four +ve LEs [13, 14] and (7D) models with five +ve LEs [15, 16]

In 2018, Yang et al. construct a 6D model which contains 16 terms; three terms are nonlinearities and are described by [17]

$$\begin{cases} \dot{x}_1(t) = a(x_2 - x_1) + x_4 + rx_6, \\ \dot{x}_2(t) = cx_1 - x_2 - x_1x_3 + x_5, \\ \dot{x}_3(t) = -bx_3 + x_1x_2, \\ \dot{x}_4(t) = dx_4 - x_1x_3, \\ \dot{x}_5(t) = -hx_2 + x_6, \\ \dot{x}_6(t) = k_1x_1 + k_2x_2. \end{cases} \quad (1)$$

The above system has four positive Lyapunov exponents:

$$\begin{cases} LE_1 = 0.4302, \\ LE_2 = 0.2185, \\ LE_3 = 0.1294, \\ LE_4 = 0.0775, \\ LE_5 = -0.0001, \\ LE_6 = -12.5222, \end{cases} \quad (2)$$

where $(x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t))^T \in R^6$ is the real state variables of the model (1), $abd \neq 0$, a, b, c are constant parameters, and d, h, r, k_1, k_2 are the control parameters.

To construct a hyperchaotic model, it is required to increase the dimension of a model. Based on state feedback control, we can add linear and nonlinear control (state variable) to the standard model [11–13].

The first pioneering study was introduced by Pecora and Carrol in 1990 for chaos synchronization of the above-mentioned model which has received a lot of attention from many areas such as encryption [17], FPGA implementation [18], optimization [19–23], electronic circuits [24], and Engineering [25]. There have been various schemes for synchronization phenomena as complete synchronization [5, 7], antisynchronization [26], hybrid synchronization [27], projective synchronization [28], and generalized projective synchronization [3]. There are several reasons for this study. One is that a few works exist in the 7D model. The second reason led us to look for another method called the linear method. It is believed that the HS with another approach (linearization) can open the way for other kinds of synchronization phenomena.

2. The New 7D Hyperchaotic Model

A novel model of high-dimensional (7D) system presents via adding nonlinear controller x_7 ; a 7D hyperchaotic model is constructed, which is described as

$$\begin{cases} \dot{x}_1(t) = a(x_2 - x_1) + x_4 + rx_6 - x_7, \\ \dot{x}_2(t) = cx_1 - x_2 - x_1x_3 + x_5, \\ \dot{x}_3(t) = -bx_3 + x_1x_2, \\ \dot{x}_4(t) = dx_4 - x_1x_3, \\ \dot{x}_5(t) = -hx_2 + x_6, \\ \dot{x}_6(t) = px_1 + qx_2, \\ \dot{x}_7(t) = x_1x_2 - kx_7, \end{cases} \quad (3)$$

where $(x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t))^T \in R^7$ is the real variables of (3), $a, b, c, d, h, r, k_1, k_2$ are the constant real parameters, and k is the parameter which determines the dynamical behavior. Fix $a = 10, b = 8/3, c = 28, d = 2, h = 9.9, r = 1, p = 1, q = 2$, and $k = 13.5$; model (3) has a hyperchaotic attractor as explained in Figure 1. The new model includes 19 terms with four nonlinearities.

2.1. Equilibrium and Stability. Equal the right-hand side to zero, such that

$$\begin{cases} a(x_2 - x_1) + x_4 + rx_6 - x_7 = 0, \\ cx_1 - x_2 - x_1x_3 + x_5 = 0, \\ -bx_3 + x_1x_2 = 0, \\ dx_4 - x_1x_3 = 0, \\ -hx_2 + x_6 = 0, \\ px_1 + qx_2 = 0, \\ x_1x_2 - kx_7 = 0. \end{cases} \quad (4)$$

Solving system (4) leads to obtaining one origin point, and the Jacobian of (3) is

$$J(O) = \begin{bmatrix} -a & a & 0 & 1 & 0 & r & -1 \\ c & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d & 0 & 0 & 0 \\ 0 & -h & 0 & 0 & 0 & 1 & 0 \\ p & q & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k \end{bmatrix}. \quad (5)$$

The model is dissipative or nonconservative since sign of diverges is negative under the typical parameters; its divergent volume is given by

$$\text{div } V = \sum_{i=1}^7 \frac{\partial \dot{x}_i}{\partial x_i} = -(a + 1 + b - d + k). \quad (6)$$

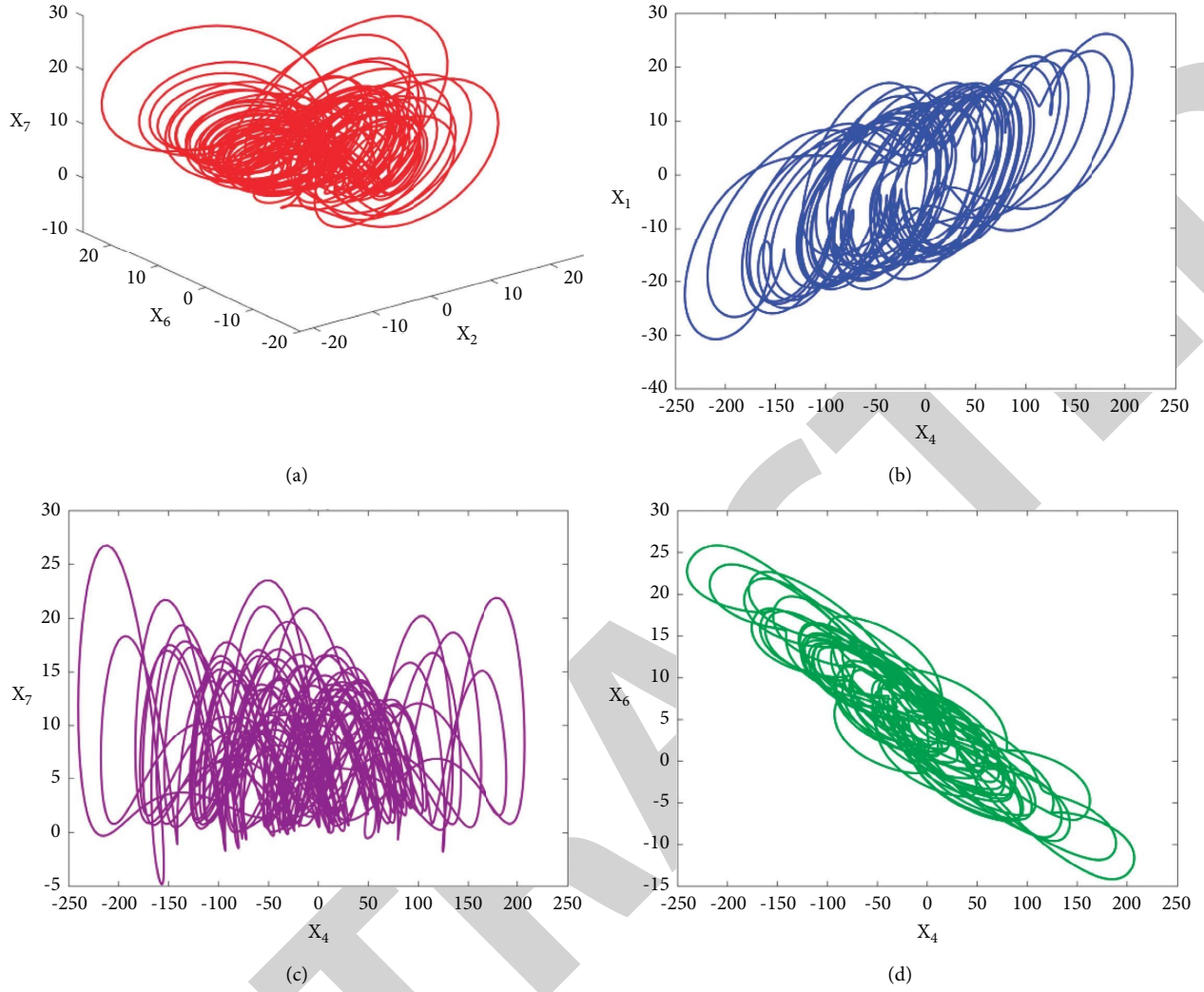


FIGURE 1: The attractors of new model: (a) $x_2 - x_6 - x_7$ space, (b) $x_7 - x_2$ plane, (c) $x_4 - x_7$ plane, and (d) $x_4 - x_6$ plane.

Using $|J(O) - \lambda I| = 0$, $I_{7 \times 7}$ is the polynomial equation and roots at $(a, b, c, d, h, r, p, q, k) = (10, 8/3, 28, 2, 9.9, 1, 1, 2, 12)$, respectively,:

$$p(\lambda) = \lambda^7 + \frac{71}{3}\lambda^6 - \frac{1191}{10}\lambda^5 - \frac{49529}{15}\lambda^4 - \frac{1867}{2}\lambda^3 + \frac{48935}{3}\lambda^2 - \frac{13332}{5}\lambda + \frac{12768}{5},$$

$$\left\{ \begin{array}{l} \lambda_1 = 2, \\ \lambda_2 = -12, \\ \lambda_3 = -\frac{8}{3}, \\ \lambda_4 = 11.4755, \\ \lambda_5 = -22.6229, \\ \lambda_{6,7} = 0.0737 \pm 0.3850i. \end{array} \right.$$

(7)

It is clear that some roots are with positive real parts; therefore, the point O is unstable. Therefore, (3) has self-excited attractors (if the model possesses unstable equilibrium points, then it is called a system with self-excited attractors) [20, 29–35].

2.2. Analysis of Lyapunov Exponents. The simulation was implemented via Wolf Algorithm and MATLAB software 2020, with parameters $a = 10, b = 8/3, c = 28, d = 2, h = 9.9, r = 1, p = 1, q = 2$ and control parameter $k = 13.5$, and the new model has five +ve Lyapunov spectra under initial conditions $(0.1, 0.2, 0.3, 0.3, 0.2, 0.1, 0.4)$, and the corresponding five exponents are

$$\begin{cases} LE_1 = 0.4783, \\ LE_2 = 0.1688, \\ LE_3 = 0.0925, \\ LE_4 = 0.0501, \\ LE_5 = 0.0001, \\ LE_6 = -12.3702, \\ LE_7 = -13.5845, \end{cases} \quad (8)$$

$$\sum_{i=1}^7 LE_i = -25.1647.$$

Figure 2 displays these exponents with step = 0.5 and tend = 200. To show the effect of the control parameter k on the proposed model, fix $a = 10, b = 8/3, c = 28, d = 2, p = 1, q = 2, h = 9.9, r = 1$ and vary parameter k . Table 1 demonstrates the new class changes into chaotic or

hyperchaotic, and some corresponding parameters k are shown in Figure 3.

3. HS of the New 7D Hyperchaotic Model

Let us model (3) is the drive as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \\ \dot{x}_7 \end{bmatrix} = \underbrace{\begin{bmatrix} -a & a & 0 & 1 & 0 & r & -1 \\ c & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & d & 0 & 0 & 0 \\ 0 & -h & 0 & 0 & 0 & 1 & 0 \\ p & q & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -k \end{bmatrix}}_{A_1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{B_1} \begin{bmatrix} -x_1 x_3 \\ x_1 x_2 \\ -x_1 x_3 \\ x_1 x_2 \end{bmatrix}, \quad (9)$$

where $A_1, B_1,$ and C_1 are the parameters and nonlinear part of (3), respectively. The response model is

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \\ \dot{y}_4 \\ \dot{y}_5 \\ \dot{y}_6 \\ \dot{y}_7 \end{bmatrix} = A_2 \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} + \left(B_2 C_2 + \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \end{bmatrix} \right), C_2 = \begin{bmatrix} -y_1 y_3 \\ y_1 y_2 \\ -y_1 y_3 \\ y_1 y_2 \end{bmatrix}, \quad (10)$$

and let $U = [u_1, u_2, u_3, u_4, u_5, u_6, u_7]^T$ be the nonlinear controller to be constructed:

- (i) If $A_1 = A_2$ and $B_1 = B_2$, then we refer to the identical model
- (ii) If $A_1 \neq A_2$ or/and $B_1 \neq B_2$, then refer to the non-identical model (different)

The two models can be synchronized as $e_i = y_i - \alpha x_i$, where

$$\alpha = \begin{cases} 1; & i = 1, 3, 5, 7 \text{ (odd)}, \\ -1; & i = 2, 4, 6 \text{ (even)}, \end{cases} \quad (11)$$

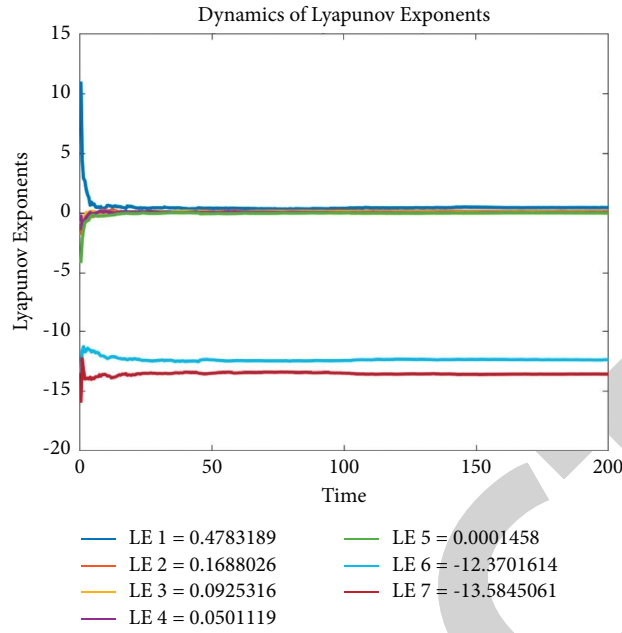


FIGURE 2: Lyapunov spectrum of the new 7D model.

TABLE 1: Dynamics of (3) versus control parameter k .

k	LE_1	LE_2	LE_3	LE_4	LE_5	LE_6	LE_7	Signs of LE
0.18	0.0306	0.0060	-0.1588	-0.3153	-1.3681	-2.1789	-7.8652	(+, ≈ 0 , -, -, -, -)
0.55	0.4753	0.1636	0.0082	-0.0094	-0.5403	-1.6395	-10.6733	(+, +, ≈ 0 , ≈ 0 , -, -, -)
0.74	0.6136	0.1414	-0.0008	-0.0584	-0.7603	-1.3147	-11.026	(+, +, 0, -, -, -)
0.85	0.4863	0.0857	-0.032	-0.0005	-0.8137	-1.2039	-11.1011	(+, +, +, 0, -, -, -)
0.88	0.5951	0.1517	-0.0008	-0.0402	-0.8021	-1.3026	-11.146	(+, +, 0, -, -, -)
1.01	0.5266	0.9952	0.0388	0.0001	-0.8955	-1.171	-11.2735	(+, +, +, 0, -, -, -)
12.99	0.3734	0.1941	0.1386	0.0470	-0.0005	-12.1641	-13.2435	(+, +, +, +, 0, -, -)
13.5	0.4783	0.1688	0.0925	0.05011	0.0001	-12.3701	-13.5845	(+, +, +, +, 0, -, -)

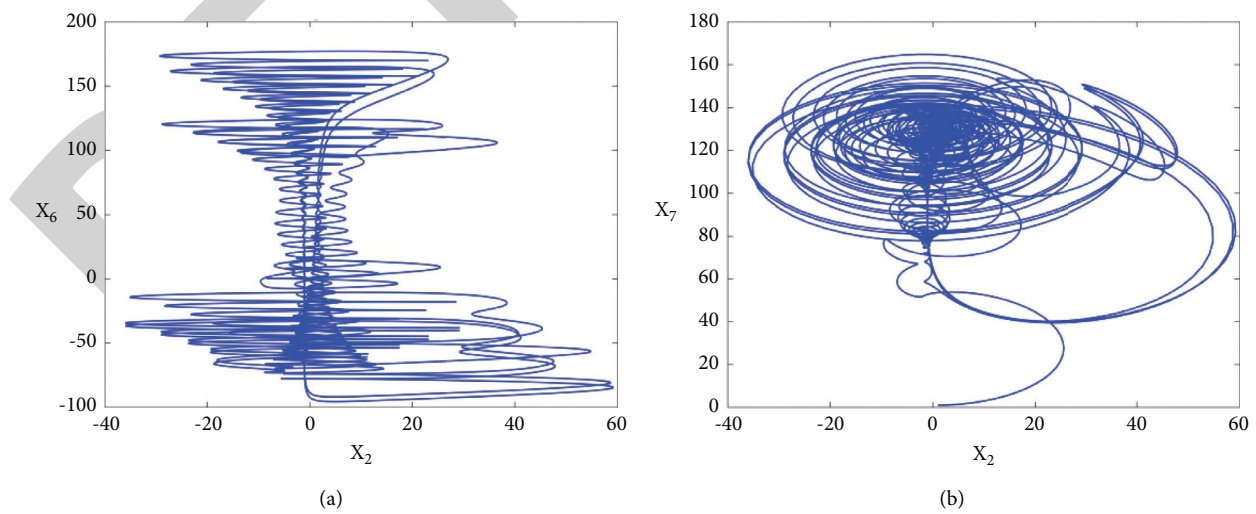


FIGURE 3: Continued.

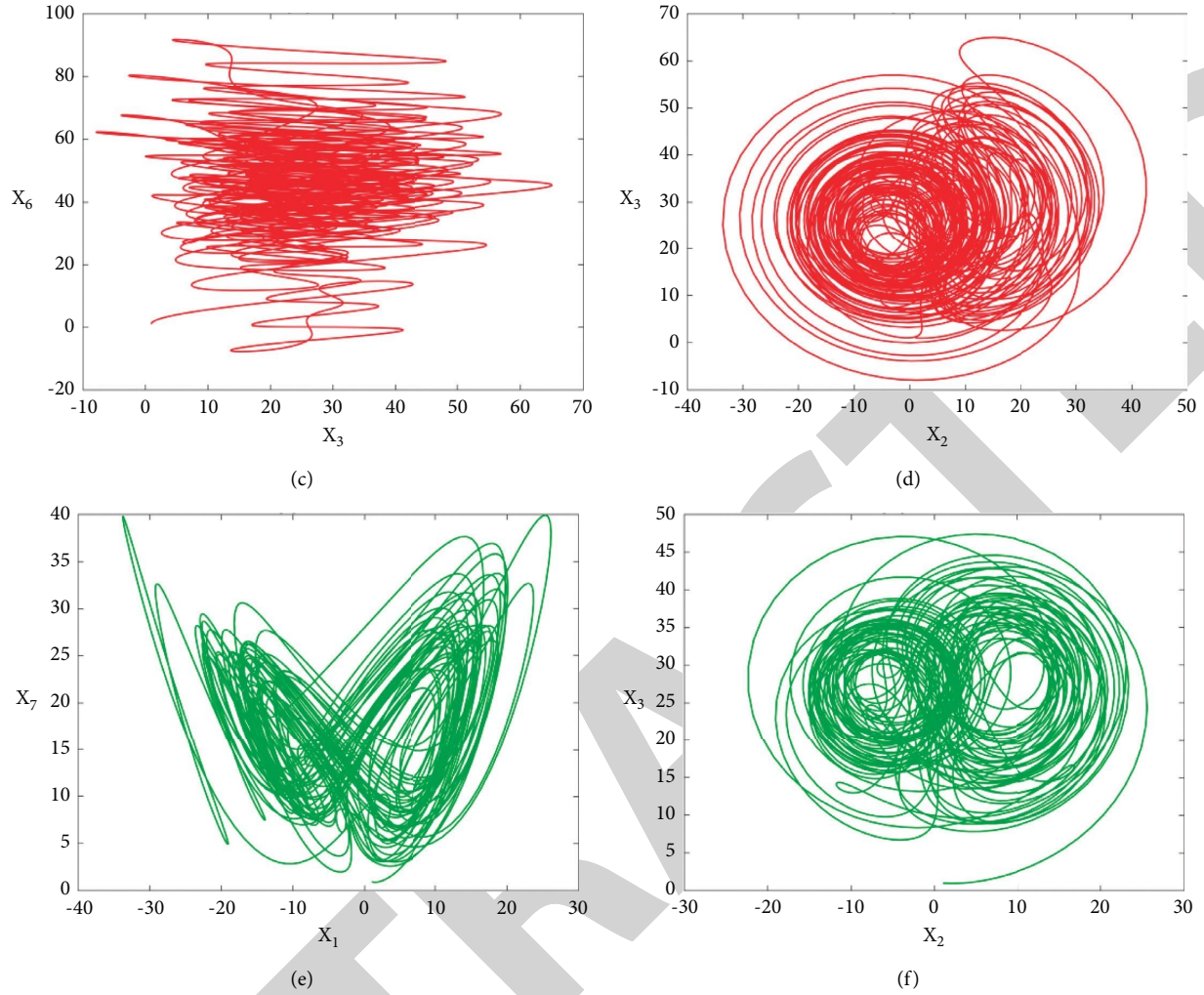


FIGURE 3: Typical dynamical behaviors of (3) at different control parameters k . (a) $k=0.5$. (b) $k=0.5$. (c) $k=0.8$. (d) $k=0.8$. (e) $k=4.25$. (f) $k=4.25$.

and satisfied that $\lim_{t \rightarrow \infty} e_i = 0$. Subtracting and adding of (10) from (9), we have the error dynamics as

$$\begin{aligned}
 \dot{e}_1 &= ae_2 - 2ax_2 - ae_1 + e_4 - 2x_4 + re_6 - 2rx_6 - e_7 + u_1, \\
 \dot{e}_2 &= ce_1 + 2cx_1 - e_2 + e_5 + 2x_5 - y_1e_3 + x_3e_1 - 2y_1x_3 + u_2, \\
 \dot{e}_3 &= -be_3 + e_1e_2 - x_2e_1 + x_1e_2 - 2x_1x_2 + u_3, \\
 \dot{e}_4 &= de_4 - y_1e_3 + x_3e_1 - 2y_1x_3 + u_4, \\
 \dot{e}_5 &= -he_2 + 2hx_2 + e_6 - 2x_6 + u_5, \\
 \dot{e}_6 &= pe_1 + 2px_1 + qe_2 + u_6, \\
 \dot{e}_7 &= -ke_7 + e_1e_2 - x_2e_1 + x_1e_2 - 2x_1x_2 + u_7.
 \end{aligned} \tag{12}$$

Theorem 1. Models (9) and (10) are globally and asymptotically HS via the nonlinear control U of equation (12) which is designed as follows:

$$\begin{cases}
 u_1 = 2ax_2 + 2x_4 - re_6 + 2rx_6 - ce_2 - x_3e_2 - pe_6, \\
 u_2 = -ae_1 - 2cx_1 - 2x_5 + 2y_1x_3 - x_1e_3 - qe_6 - x_1e_7, \\
 u_3 = y_1e_2 - e_1e_2 + x_2e_1 + 2x_1x_2 + y_1e_4, \\
 u_4 = -2de_4 - e_1 - x_3e_1 + 2y_1x_3, \\
 u_5 = -2hx_2 + 2x_6 - e_5, \\
 u_6 = -2px_1 - e_5 - e_6, \\
 u_7 = e_1 - e_1e_2 + x_2e_1 + 2x_1x_2.
 \end{cases} \tag{13}$$

Proof. Inserting the above control in (12), we obtain

$$\begin{cases} \dot{e}_1 = ae_2 - ae_1 + e_4 - e_7 - ce_2 - x_3e_2 - pe_6, \\ \dot{e}_2 = ce_1 - e_2 + e_5 - y_1e_3 + x_3e_1 - ae_1 - x_1e_3 - qe_6 - x_1e_7, \\ \dot{e}_3 = -be_3 + x_1e_2 + y_1e_2 + y_1e_4, \\ \dot{e}_4 = -de_4 - y_1e_3 - e_1, \\ \dot{e}_5 = -he_2 + e_6 - e_5, \\ \dot{e}_6 = pe_1 + qe_2 - e_5 - e_6, \\ \dot{e}_7 = -ke_7 + x_1e_2 + e_1. \end{cases} \quad (14)$$

The characteristic equation and roots are as

$$\lambda^7 + \frac{89}{3}\lambda^6 + \frac{6529}{10}\lambda^5 + \frac{238931}{30}\lambda^4 - \frac{1182143}{30}\lambda^3 + \frac{291515}{3}\lambda^2 + \frac{1944721}{15}\lambda + \frac{1163288}{15} = 0,$$

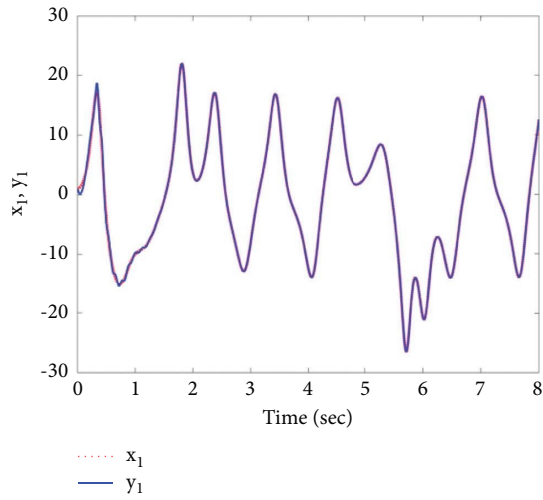
$$\begin{cases} \lambda_1 = -\frac{8}{3}, \\ \lambda_2 = -11.9657, \\ \lambda_3 = -2.0021, \\ \lambda_{4,5} = -1.1984 \pm 1.4399i, \\ \lambda_{6,7} = -5.3177 \pm 17.8223i. \end{cases} \quad (15)$$

Clearly, all roots are with negative real parts; the linearization approach achieved the HS between (9) and (10).

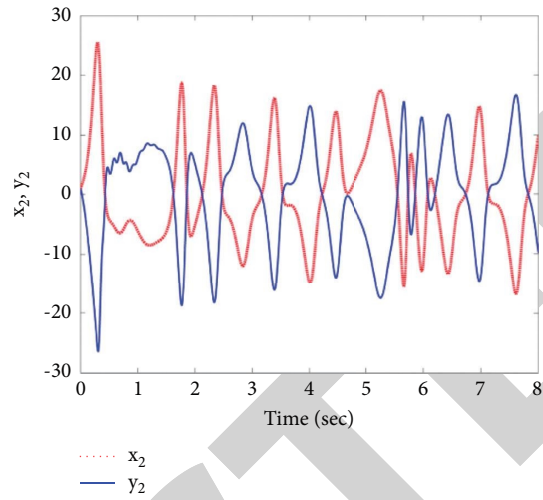
Now, in second approach, we construct the auxiliary (Lyapunov) function as $V(e_i) = e^T P e$, i.e.,

$$V(e_i) = [e_1, e_2, e_3, e_4, e_5, e_6, e_7]^T \underbrace{\begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5/99 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \end{bmatrix}}_P \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix}. \quad (16)$$

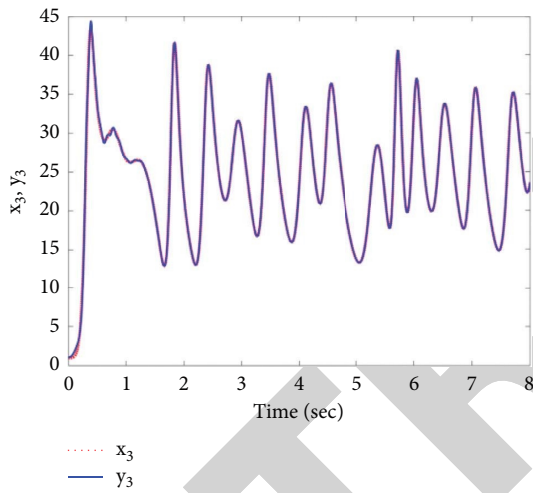
The derivative of the above function $V(e_i)$ is



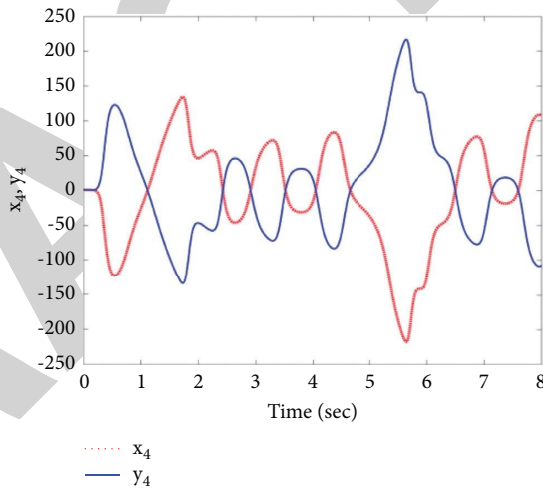
(a)



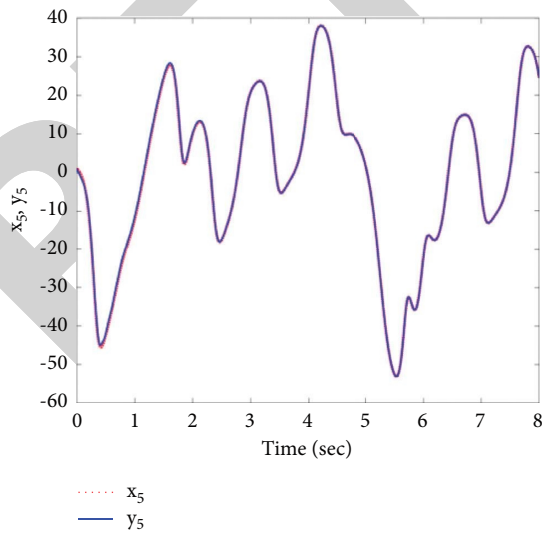
(b)



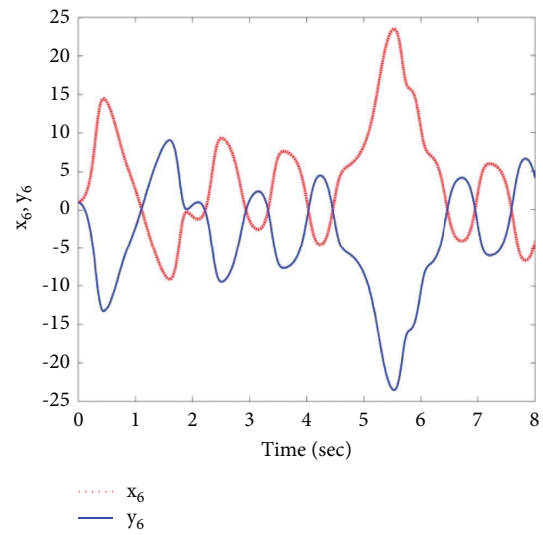
(c)



(d)

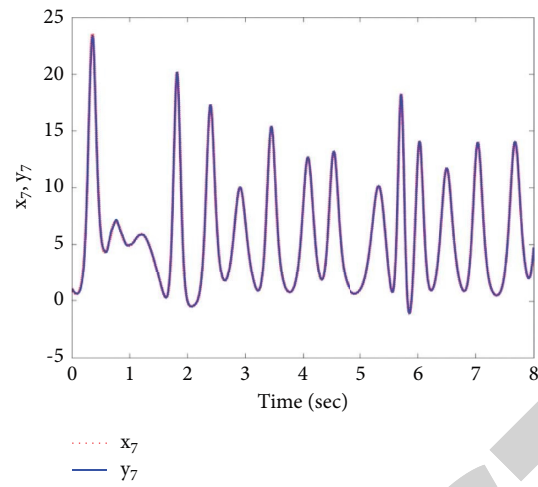


(e)



(f)

FIGURE 4: Continued.



(g)

FIGURE 4: HS between models (9) and (10) with nonlinear control (13).

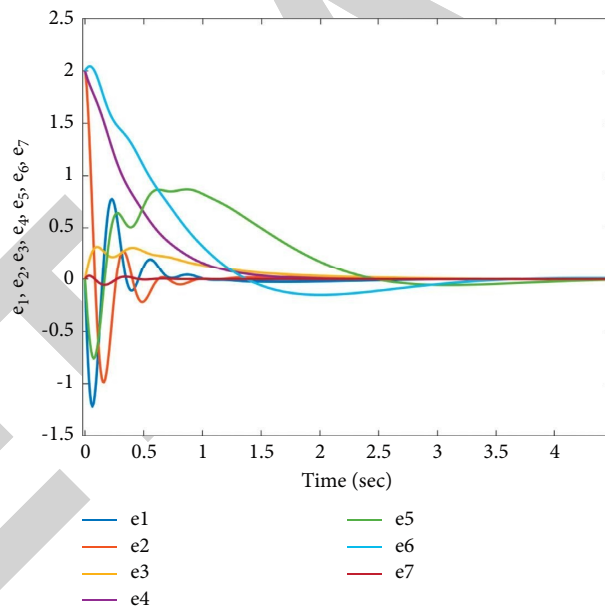


FIGURE 5: The convergence of models (12) with nonlinear controllers (13).

$$\dot{V}(e_i) = e_1 \dot{e}_1 + e_2 \dot{e}_2 + e_3 \dot{e}_3 + e_4 \dot{e}_4 + \frac{10}{99} e_5 \dot{e}_5 + e_6 \dot{e}_6 + e_7 \dot{e}_7,$$

$$\begin{aligned} \dot{V}(e) = & e_1 (ae_2 - ae_1 + e_4 - e_7 - ce_2 - x_3 e_2 - pe_6) + e_2 (ce_1 - e_2 + e_5 - y_1 e_3 + x_3 e_1 - ae_1 - x_1 e_3 - qe_6 - x_1 e_7) \\ & + e_3 (-be_3 + x_1 e_2 + y_1 e_2 + y_1 e_4) + e_4 (-de_4 - y_1 e_3 - e_1) + \frac{10}{99} e_5 (-he_2 + e_6 - e_5) + e_6 (pe_1 + qe_2 - e_5 - e_6) \\ & + e_7 (-ke_7 + x_1 e_2 + e_1), \end{aligned}$$

$$\dot{V}(e_i) = -[e_1, e_2, e_3, e_4, e_5, e_6, e_7]^T \underbrace{\begin{bmatrix} 10 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 8/3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10/99 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 12 \end{bmatrix}}_Q \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix}, \quad (17)$$

where $Q = \text{diag}(10, 1, 8/3, 2, 10/99, 1, 12)$, so $Q > 0$. Consequently, $\dot{V}(e_i) < 0$ on R^7 . The nonlinear controller realized the HS between models (9) and (10).

For simulation results, the initial values are $(15, 2, 0, -2, -3, 0)$ and $(-15, -10, -8, 6, 0, -4)$ to illustrate the HS that happened between (9) and (10) numerically. Figures 4 and 5 check these results numerically, respectively. \square

4. Conclusions

In this paper, a novel class 7D model with a self-excited attractor and multiple positive Lyapunov exponents has been proposed via a state feedback controller. Furthermore, some features of dynamical behaviors such as equilibria points, stability, and Lyapunov exponents are investigated, as well as hybrid synchronization between two new identical models, are rigorously derived and studied by designing a suitable controller, based on nonlinear control strategy with two analytical methods: Lyapunov's and linearization approach. The new system may have a good application in the field of encryption and nonlinear circuits.

Data Availability

The data underlying the results presented in the study are available within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] G. Zhang, F. Zhang, X. Liao, D. Lin, and P. Zhou, "On the dynamics of new 4D Lorenz-type chaos systems," *Advances in Difference Equations*, vol. 2017, no. 1, 13 pages, 2017.
- [2] K. Adel Abed, "Controlling of jerk chaotic system via linear feedback control strategies," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, no. 1, pp. 370–378, 2020.
- [3] Z. S. Al-Talib and S. F. Al-Azzawi, "Projective synchronization for 4D hyperchaotic system based on adaptive nonlinear control strategy," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 19, no. 2, pp. 715–722, 2020.
- [4] Y.-W. Wang and Z.-H. Guan, "Generalized synchronization of continuous chaotic system," *Chaos, Solitons & Fractals*, vol. 27, no. 1, pp. 97–101, 2006.
- [5] H. Chen, "Global chaos synchronization of new chaotic system via nonlinear control," *Chaos, Solitons & Fractals*, vol. 23, no. 4, pp. 1245–1251, 2005.
- [6] G. Hu, "Generating hyperchaotic attractors with three positive Lyapunov exponents via state feedback control," *International Journal of Bifurcation and Chaos*, vol. 19, no. 02, pp. 651–660, 2009.
- [7] Q. Yang and C. Chen, "A 5D hyperchaotic system with three positive Lyapunov exponents coined," *International Journal of Bifurcation and Chaos*, vol. 23, no. 06, Article ID 1350109, 2013.
- [8] H. Semercioglu, "The new balance of power in the southern caucasus in the context of the nagorno-karabakh conflict in 2020," *ReS-Research Studies Anatolia Journal*, vol. 4, no. 1, pp. 49–60, 2021.
- [9] Q. Yang, D. Zhu, and L. Yang, "A new 7D hyperchaotic system with five positive Lyapunov exponents coined," *International Journal of Bifurcation and Chaos*, vol. 28, no. 05, Article ID 1850057, 2018.

Retraction

Retracted: Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] K. Twarish Alhamazani, J. Alshudukhi, S. Aljaloud, and S. Abebaw, "Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 3941978, 8 pages, 2021.

Research Article

Implementation of Machine Learning Models for the Prevention of Kidney Diseases (CKD) or Their Derivatives

Khalid Twarish Alhamazani ¹, **Jalawi Alshudukhi** ¹, **Saud Aljaloud** ¹,
and Solomon Abebaw ²

¹University of Ha'il, College of Computer Science and Engineering, Department of Computer Science, Ha'il, Saudi Arabia

²Department of Statistics, Mizan-Tepi University, Tepi, Ethiopia

Correspondence should be addressed to Solomon Abebaw; solomonabebaw@mtu.edu.et

Received 29 November 2021; Revised 9 December 2021; Accepted 14 December 2021; Published 30 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Khalid Twarish Alhamazani 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.

Chronic kidney disease (CKD) is a global health issue with a high rate of morbidity and mortality and a high rate of disease progression. Because there are no visible symptoms in the early stages of CKD, patients frequently go unnoticed. The early detection of CKD allows patients to receive timely treatment, slowing the disease's progression. Due to its rapid recognition performance and accuracy, machine learning models can effectively assist physicians in achieving this goal. We propose a machine learning methodology for the CKD diagnosis in this paper. This information was completely anonymized. As a reference, the CRISP-DM® model (Cross industry standard process for data mining) was used. The data were processed in its entirety in the cloud on the Azure platform, where the sample data was unbalanced. Then the processes for exploration and analysis were carried out. According to what we have learned, the data were balanced using the SMOTE technique. Four matching algorithms were used after the data balancing was completed successfully. Artificial intelligence (AI) (logistic regression, decision forest, neural network, and jungle of decisions). The decision forest outperformed the other machine learning models with a score of 92%, indicating that the approach used in this study provides a good baseline for solutions in the production.

1. Introduction

Chronic kidney disease (CKD) is one of the leading causes of death in recent years, according to a report by the Global Burden of Disease [1]. One in every seven persons has CKD, one of the undiscovered illnesses that have the greatest influence on patients' quality of life and increase the chance of death significantly. The general system of social security in health (SGSSS) has taken chronic kidney disease (CKD) into account [2], as a high-cost pathology for generating a powerful economic impact on the finances of the system, causing a dramatic effect on the quality of life of the patient and their family, including employment repercussions. To reduce the high mortality of CKD, research should be deepened and directed to the initial stages of the disease, analyzing its risk group, with the help of laboratory tests, seeking that patients do not reach the final

stages such as dialysis, transplantation, or death [3]. Through automatic learning, the aim is to find a valuable contribution so that an early classification of the disease can be carried out in its initial stages through the results of clinical laboratories, taking advantage of the great potential of automatic learning in the analysis and classification of the data. It is necessary that the technical help tools that are based on data can support the decision-making process in the initial diagnoses quickly, with high precision, and at low cost. With them, the time required for diagnosis is reduced, allowing the patient to receive treatment for the disease before it progresses to a stage of no return.

Machine learning can be broadly divided into supervised learning, unsupervised learning, and reinforcement learning [4]. Supervised learning is the most common form of machine learning used in medical research [5]. Each instance of supervised learning contains an input object (usually a

vector) and the desired output value (also known as a supervised signal) [3]. Usually, the algorithms applied for supervised learning are decision trees, naive bayes classification, least squares regression, logistic regression, and vector support machine (SVM) methods (Classifier Sets). Recent studies show that deep neural networks have achieved comparable high performance at the expert level in natural and biomedical image classification tasks [6]. This, coupled with the ability to generate assumptions [7], the adaptability to heterogeneous data set analysis, and open-source deep learning programs that are widely disseminated, makes deep learning play an essential role in promoting medical development [8].

This research work aims to design and implement a machine learning model that, based on data from clinical laboratories, allows predicting the possible diagnosis of CKD in its initial stages, helping reduce the mortality rate and costs for the health system.

2. Methodology

In the development of this project, the CRISP-DM® model [9–15] is used, which is the broadest reference guide used in the development of analytical and mining projects to data collected from clinical laboratories. For this, each of the proposed stages will be implemented.

2.1. Phase I. Understanding the Business. This phase is divided into four tasks that will help better understand the business, as shown in Figure 1.

2.1.1. Determination of Business Objectives. Chronic kidney disease (CKD) is a type of kidney disease in which there is gradual loss of kidney function over a period of months to years. Initially, there are generally no symptoms; later, symptoms may include leg swelling, feeling tired, vomiting, loss of appetite, and confusion. Complications include an increased risk of heart disease, high blood pressure, bone disease, and anaemia. CKD is associated with a decrease in kidney function related to age and is accelerated in hypertension, diabetes, obesity, and primary kidney disorders. CKD is a global health problem with a high morbidity and mortality rate, and it induces other diseases. As there are no obvious symptoms during the early stages of CKD, patients often do not notice the disease, this being the main feature, eventually leading to a complete loss of kidney function. Early detection of CKD allows patients to receive timely treatment to improve the progression of this disease. As it has been proposed in the objectives of the work, the aim is to develop an automatic learning model for the prediction in the diagnosis of CKD and to contribute to the reduction of significant complications in the disease such as dialysis processes, kidney transplantation, or reaching death. The main criterion of success for this project, with the help of machine learning, is to identify the behaviors or behavior patterns in the initial stages of CKD to improve the quality of life of patients.

2.1.2. Assessment of the Situation. The idea for the approach of this project arises from the current situation regarding the increase in the confirmatory diagnosis of CKD, and lack of treatment or the user's ignorance of its pathologies leads to irreversible kidney failure in the final stages of CKD, such as dialysis for life, financially affecting the health system, as it is a costly treatment that generates the most significant amount of absorption of the resources available for health in Iraq. This could be reduced by using tools such as machine learning to classify ERC from the initial stages. Although the application of machine learning in healthcare and other areas is favorable, the field of kidney disease has not yet exploited its full potential [16–25].

2.1.3. Determination of the Data Mining Objectives. As referenced in the general objective, the technical terms of this project are to design, implement, and deploy a machine learning model that, based on data from clinical laboratories, allows to classify the possibility of a diagnosis of CKD. Through the analysis of laboratory studies that are low-cost for health entities, these data reduce the mortality rate and costs of the health system.

The medical history and the laboratory tests indicate identifying symptoms or signs that can be used as constitutive variables of the problem in CKD patients on a large scale since a large amount of data can be handled without inconvenience. With the initial data, a description and exploration of these are made, verifying that they can be used or have the minimum information to perform the classification, through the analysis of these data and obtain the patients with an incidence of CKD. With the data obtained, a training set is molded. Several tests are carried out that define or determine the most appropriate technique(s) for the classifier and that the results are practical and efficient. With the defined classifier, the predictive models are trained and validated to establish the model with the highest precision for the data, selecting the one that offers the best results. Predictive models often run calculations during ongoing transactions, for example, to assess the risk or opportunity for a particular patient in a way that provides insight into the treatment decision-making.

2.1.4. Production of the Project Plan. The project plan can be found in the schedule annex-project work plan. It describes all the necessary steps, from the problem statement and data collection to its analysis.

2.2. Phase II: Study and Understanding of the Data. This section describes the initial data obtained, such as the number of records and fields per record and their identification, each field's meaning, and the initial format's description, as shown in Figure 2.

2.2.1. Collection of Starting Data. The data set used for this project was obtained thanks to the Baghdad Renal Clinic, Iraq, and its manager and legal representative, nephrologist Dr. Ahmad. They allowed and authorized the treatment of

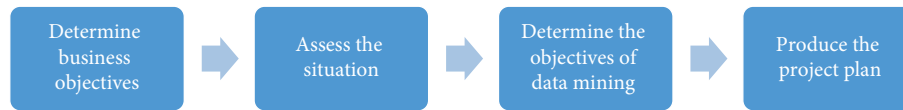


FIGURE 1: Business understanding tasks.

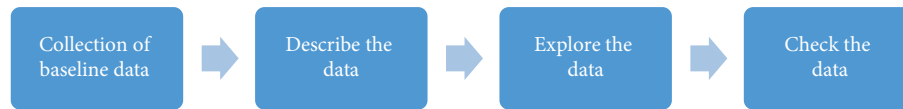


FIGURE 2: Study tasks and understanding of the data.

these data. The dataset contains 373,770 anonymized samples. In this data set, each sample has 17 variables or predictive characteristics (11 numerical variables and six categorical (nominal) variables).

2.2.2. Explore the Data. To explore the data, a database was created on the Azure platform. The data are imported and connected to the Power BI visualization tool to perform and visualize all the data contained in the data set more harmoniously. Initially, descriptive statistics of the variables that make up the data are carried out. In Table 1, the main characteristics of the variables of the files used are identified; this result was obtained from Python commands within Power BI.

Before starting to process the data set, a set of visualizations is made to help better understand the characteristics of the information being worked with and their correlation. First, the four input characteristics with names “Duration,” “Pages,” “Actions,” and “Value” are displayed in a historical format; it can be seen graphically which values comprise the minimums and maximums and in which intervals the highest density of records is there. As can be seen in the data, the base diseases of CKD are hypertension and diabetes. The data show us that there are many patients with this disease, but hypertension, also known as a silent disease, is highlighted. Considering the obtained results, the highest prevalence of CKD is in women due to their longer life expectancy and reaching the age of risk of CKD (older adult). It is estimated that a quarter of chronic patients present with low hemoglobin in the initial stages of the disease. Epidemiological studies place excess weight as a risk factor for the development of CKD along with other factors. As creatinine rises in your blood average, the percentage of kidney function goes down. For the above reason, it is a significant variable for the diagnosis of kidney disease. As evidenced, the age and weight variables present a high positive correlation, which is not present between hypertension and diabetes mellitus, and no correlation is observed between these variables. There is a low positive correlation for the variables of weight and hypertension, as it also happens between the variables of hypertension with age and diabetes mellitus with age. That is, hypertension and diabetes mellitus are diseases that do not depend on age factor; it can be caused due to family history.

2.2.3. Verification of the Quality of Data. In this section, data verification was performed to determine the consistency of the field values and the amount of distribution of the null values and to find out-of-range values that can generate noise for the process. This verification process was carried out on the entire data set received. In the fields where no records were found, the empty fields were changed to a Null value.

2.3. Phase III: Data Analysis and Selection of Characteristics. By having the information from the data, the focus is on identifying the variables to be used. During the data review, a total of 58 variables were found; within these, three variables were identified that were eliminated to preserve the privacy, security, and sensitivity of the patients. Of the 55 remaining variables, those who did not have a sample were eliminated since they would not have any determination if they did not provide information in training. Afterward, a classification was made between the subjective and objective variables. After performing the medical analysis, 38 variables that are not relevant for CKD prediction were eliminated.

With the identification of the scale used to classify patients with CKD, the variables to be used to evaluate a patient according to their degree of severity were determined. The initial list of 58 variables after performing the medical analysis allows us to arrive at a set of 17 variables, which required the expertise of the nephrologist in the elimination process. Within this entire process, these criteria will be used to measure the severity of kidney disease. It is essential to be clear that the judgment of the expert (in this case, the doctor) is crucial to make a final decision about the condition of a patient; for this, the expert will take into account the essential antecedents such as hypertension and diabetes. When the data set is ready, the import is carried out on the experiment that will be carried out in Azure Machine Learning Studio, where the algorithms will be evaluated. Selecting the numeric and non-numeric attributes allows limiting the columns available for a later operation. By having the attributes selected, it is indicated which column will have the values to be categorized or predicted; in the case of this evaluation, the variable stage is used. As indicated in Table 2, it is the one that describes the stage of renal failure in a patient.

Having identified the characteristics of the data set with the most significant predictive power, statistical tests are performed with the Pearson correlation test to determine which columns are more predictive. For this task, the filter-based

TABLE 1: Statistics of the variables.

.	Average	Std.dev	Minimum	25%	50%
Municipality	30328.27	28219.83	5001.00	8001.00	13001.00
Age	73.51	12.01	1.00	67.00	75.00
HTA	1.00	0.00	1.00	1.00	1.00
Albunia	2.50	1.50	0.00	1.00	3.00
Sugar	2.00	1.22	0.00	1.00	2.00
UREA_SANGRE	143.39	86.16	1.00	70.00	139.00
Creatinine	5.78	4.89	0.10	1.70	3.80
Sodium	134.95	20.23	100.00	117.00	135.00
Potassium	4.75	1.30	2.50	3.60	4.80
Hemoglobin	14.40	1.31	12.10	13.40	14.30
DM	0.00	0.00	0.00	0.00	0.00
Weight	77.75	10.00	60.00	70.50	77.20
Stadium	3.14	0.44	3.00	3.00	3.00

TABLE 2: Stadium classification.

Stadium	Description
1	Kidney damage with normal or high filtration
2	Kidney damage with mild decrease in kidney function
3	Kidney damage with a moderate decrease in kidney function
4	Kidney damage with a severe decrease in kidney function
5	Renal failure

feature selection module that Azure machine learning brings is used, as seen in Figures 3 and 4, which provides multiple feature selection algorithms, including correlation methods identifying the ordinal and non-ordinal variables.

2.4. Phase IV: Modeling. As the next step, the selection process of the machine learning technique that will be used to develop the classifier of patients, this is when the object of this research begins.

According to the effectiveness of the algorithms most used in studies for the diagnosis of diseases, the following algorithms are chosen:

- (i) Logistic regression (LR)
- (ii) Artificial neural networks (ANN)
- (iii) Forest of decisions (RF)
- (iv) Jungle of decisions (DJ)

By having these algorithms, the data set is divided randomly into two parts: training sets and test sets. The experiment is carried out with the split module, (Figure 5) using 70% of the data as training data and 30% is reserved to evaluate the model's efficiency.

Then, the chosen models are trained using the train model tool, an Azure tool used to train a model in a supervised way with the training data set as input to the image model training experiment Figure 6. To these models, the score model is added which is used in Azure for score predictions of a trained model. In each combination of the trained model and the test data, the evaluate model classification model is used to calculate the results confusion matrix Figure 7.

When carrying out this first evaluation of the chosen algorithms, it is verified that the data for the models are

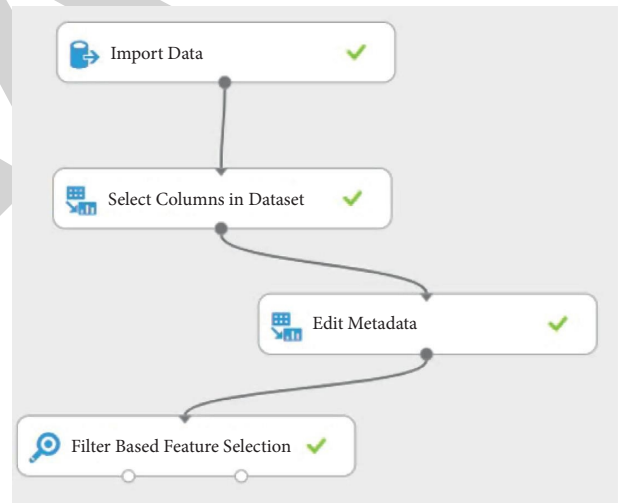


FIGURE 3: Image data preparation experiment.

unbalanced. For example, in general, precision is the probability that an individual is correctly classified by a test, that is, the sum of the true positives plus the true negatives divided by the total number of individuals evaluated gives acceptable values as observed in Table 3. Still, when verifying the matrices, the values always tend to stage 3.

The SMOTE module is added to carry out the balancing task, a statistical technique that allows increasing the number of cases in a balanced way, taking samples from the characteristic space for each target class and its closest neighbors.

2.5. Phase V: Evaluation. The most critical measure of classification algorithms is their accuracy. As observed in the comparative results based on the precision Table 4, it is

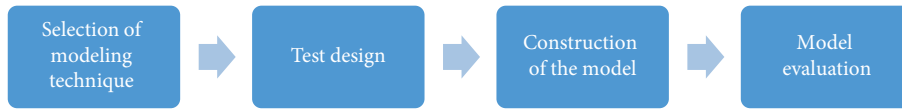


FIGURE 4: Modeling.

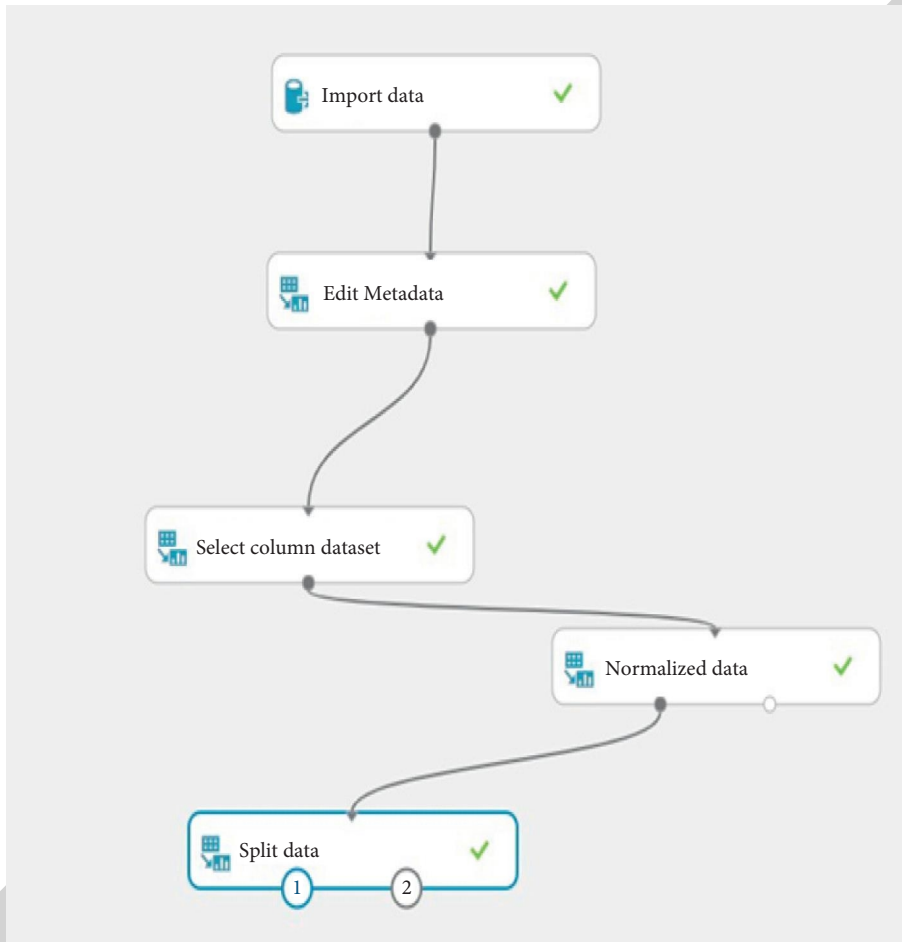


FIGURE 5: Image data division experiment.

found that the four classification algorithms were compared. Multiclass logistic regression (LR) achieves the lowest accuracy, 68%, which implies a poor classifier of results.

However, the jungle multiclass decision (DJ) and multiclass neural network (ANN) algorithms work well and show competitive performance with each other.

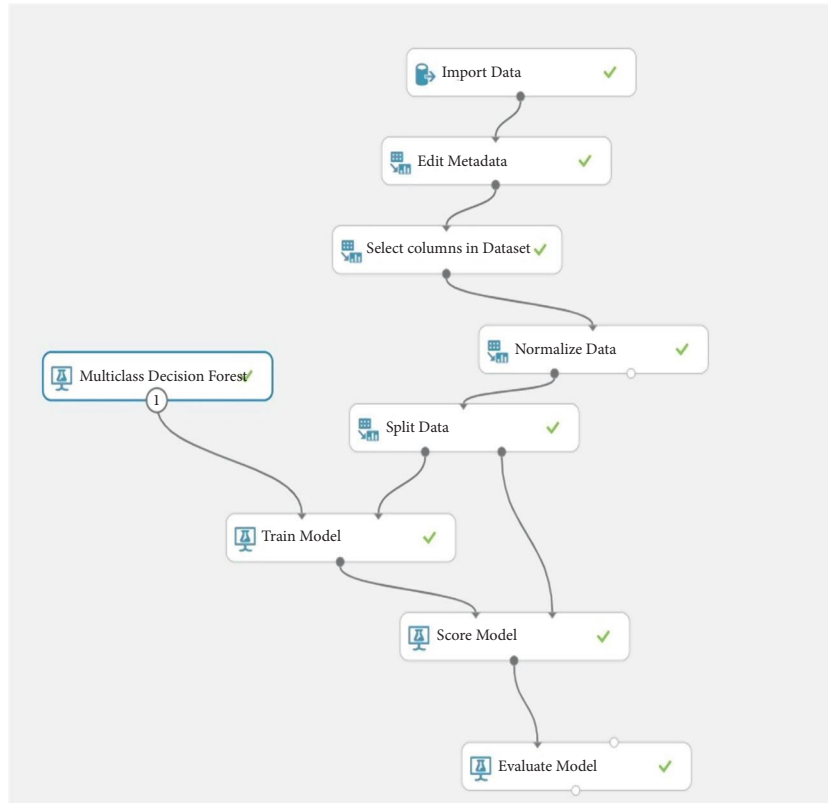


FIGURE 6: Image model training experiment.



FIGURE 7: Evaluation (obtaining results).

TABLE 3: Performance comparison between models.

	Multiclass red neuronal	Decision forest multiclass	Multiclass logistic regression	Multiclass jungle of decisions
General accuracy	0.784	0.831	0.773	0.835
Medium accuracy	0.892	0.887	0.886	0.890
Micro-precision averaged	0.784	0.831	0.773	0.835
Macro-accuracy averaged	0.446	0.691	0.374	0.817
Micro-sensitivity averaged	0.784	0.831	0.773	0.835
Macro-sensitivity averaged	NA	0.617	NA	0.537
Artificial neural networks matrix				
		3	4	5
Current class	3	0.971	0	0.029
	4	0.972	0.001	0.028
	5	0.829	0	0.17
Decision forest matrix				
		3	4	5
Current class	3	0.939	0.033	0.027
	4	0.494	0.417	0.089
	5	0.406	0.1	0.494
Logistic regression matrix				

TABLE 3: Continued.

	Multiclass red neuronal	Decision forest multiclass	Multiclass logistic regression	Multiclass jungle of decisions
		3	4	5
Current class	3	0.94	0.014	0.046
	4	0.799	0.121	0.08
Jungle matrix of decisions	5	0.595	0.035	0.37
		3	4	5
Current class	3	0.987	0	0.013
	4	0.778	0.153	0.069
	5	0.507	0.023	0.471

TABLE 4: Final accuracy and completeness values.

	Accuracy	Completeness
Decision forests	0.922	0.921
Neural networks	0.806	0.800
Logistic regression	0.689	0.689
Decision jungle	0.754	0.750

Although they have reached an accuracy of 75% and 80%, respectively, they fail to show superior performance to the multiclass decision forests (RF) algorithm, which reached an accuracy of 92% that indicates the effective performance in the classification of the ERC dataset used.

3. Conclusions

This study explored how a learning model can be used to classify the possibility of a CKD diagnosis.

Consequently, the indirect agreement with the objectives of the project, the CRISP-DM methodology was adapted to the context of the problem so that the different logically organized stages were taken; data collection, pre-processing, learning, evaluation, and selection, which allowed the construction of a model capable of classifying the possibility of a diagnosis of CKD with an accuracy of 93%. As evidenced in the results, the decision forests algorithm has obtained quite optimal results, where predictions of 93% have been obtained. Data preparation is a fundamental step in the process and the absolute precision of the model is directly dependent on this phase. Thanks to the models, we can see how changing the characteristics affects the search for the target value with a simple change of column selection or improvements in the data. The innovation of this work results from the design adjusted to the environment of the health system in Iraq and the pathology of the CKD in our country, with a methodology adapted to the case study and a production architecture proposal for the model with Microsoft Azure tools of form that allows satisfying in the future the scalability of the solution. Furthermore, this methodology could apply to clinical data of other diseases and pathologies inaccurate medical diagnoses. The development of this project allowed the author to acquire more excellent knowledge through both practical and theoretical work about current techniques for the development of machine learning.

This study has limitations, so there is a room for future research. The study did not have a significant data sample due to the restrictions of medical data and its legal effects in Iraq. Continuing with the expansion of the database (increasing the number of examples for each variable) would reduce the limited generalization error for the model and, at the same time, allow the severity of the disease to be detected. This model can be refined with increasing data size and quality. It also opens the space for a variety of studies from other disciplines, such as economic studies around the impact of obtaining a diagnosis in less time to treat the disease in early stages, reducing costs in the health system. In addition, a variety of sociological and clinical studies on the consequences of early CKD management brings about the quality of life of patients and their families. Although the validity of this research is internal, since the document corpus is private and cannot be published for other works, it will help interested professionals with machine learning to carry out their studies in the classification area.

Data Availability

The data underlying the results presented in the study are available in the manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] A. Vijayalakshmi and V. Sumalatha, "Survey on diagnosis of chronic kidney disease Using Machine learning algorithms," in *Proceedings of the 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)*, pp. 590–595, Thoothukudi, India, December 2020.
- [2] S. Sengan, O. I. Khalaf, P. Vidya Sagar, D. K. Sharma, L. Arokia Jesu Prabhu, and A. A. Hamad, "Secured and privacy-based IDS for healthcare systems on E-medical data using machine learning approach," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–11, 2022, p.
- [3] A. H. Hameed, E. A. Mousa, and A. A. Hamad, "Upper limit superior and lower limit inferior of soft sequences," *International Journal of Engineering and Technology*, vol. 7, no. 4.7, pp. 306–310, 2018, p.

Retraction

Retracted: Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] R. R. Althar, D. Samanta, M. Kaur, A. A. Alnuaim, N. Aljaffan, and M. Aman Ullah, "Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8522839, 19 pages, 2021.

Research Article

Software Systems Security Vulnerabilities Management by Exploring the Capabilities of Language Models Using NLP

Raghavendra Rao Althar ^{1,2}, Debabrata Samanta ³, Manjit Kaur ⁴,
Abeer Ali Alnuaim ⁵, Nouf Aljaffan⁵ and Mohammad Aman Ullah ⁶

¹Data Science Department, CHRIST Deemed to Be University, Bangalore, Karnataka, India

²QMS, First American India Private Ltd., Bangalore, Karnataka, India

³Department of Computer Science, CHRIST Deemed to Be University, Bangalore, Karnataka, India

⁴School of Electrical Engineering and Computer Science, Gwangju Institute of Science and Technology, Gwangju, Republic of Korea

⁵Department of Computer Science and Engineering, College of Applied Studies and Community Services, King Saud University, P.O. Box 22459, Riyadh 11495, Saudi Arabia

⁶Department of Computer Science and Engineering, International Islamic University Chittagong, Chittagong, Bangladesh

Correspondence should be addressed to Mohammad Aman Ullah; aman_cse@iiuc.ac.bd

Received 25 October 2021; Revised 28 November 2021; Accepted 9 December 2021; Published 27 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Raghavendra Rao Althar 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.

Security of the software system is a prime focus area for software development teams. This paper explores some data science methods to build a knowledge management system that can assist the software development team to ensure a secure software system is being developed. Various approaches in this context are explored using data of insurance domain-based software development. These approaches will facilitate an easy understanding of the practical challenges associated with actual-world implementation. This paper also discusses the capabilities of language modeling and its role in the knowledge system. The source code is modeled to build a deep software security analysis model. The proposed model can help software engineers build secure software by assessing the software security during software development time. Extensive experiments show that the proposed models can efficiently explore the software language modeling capabilities to classify software systems' security vulnerabilities.

1. Introduction

Software has become a core part of human life and plays a prominent role in day-to-day activities. With advances in technology, loopholes are also getting created at a rapid pace. Software development organizations consider the security of the software as a prominent part of their focus area for customers. However, there are extensive knowledge sources available across the industry within and outside the organization. Due to the continuously changing priorities of the organization, security knowledge takes a back seat. There is a need for devising a proper mechanism to assimilate all the knowledge prevalent in the industry and provide it to the software development team in a controlled way as and when

they need it. It requires an intelligent way of putting the information together, learning from them continuously, and using the learnings in operations. There is a need for a smart knowledge management system for the security of the software systems. A smart knowledge management system will be the system that will have the capability to integrate the data from various sources. And this integrated source of required information is made available for the one interested at the right point of time in software development processes. The system will have to be fed in with the events from the outside organization and within the organization. All the data processed within the software development value stream will also be leveraged for this knowledge system. Customer conversations with the technology team will hold

prominence as they may have the software system's explicit and implicit security needs. This paper is focusing on streamlining the software engineering process by leveraging the artificial intelligence approach. Security concerns of the software engineering processes are targeted. This focus area aligns with the journal's scope, as the journal intends to bridge the gap between engineering, artificial intelligence, and neuroscience.

Leveraging the latest data science advancements for this problem area will bring efficiency in managing this domain. Most of the data will be a natural language, so NLP (natural language processing) approaches will be explored. Data related to a conversation with customer and software requirements management team is natural language data. Good practices documented in the security management platforms in the industry provide natural language inputs. Within the software development landscape, there are natural language data like security scan-related outcomes. These rich data forms make NLP a go-to framework for modeling. All these data sources have hidden patterns for the possible security flaws that would creep into the system. With some good language processing capabilities emerging in the field, their relevance has to be studied. The research area taken up involves building a comprehensive knowledge management system to facilitate secured software development. The system will be banking on the data sources from the industry, the company, and the data processed during conversations between customer and software development teams. The paper's focus is to explore some of the key constructs associated with this system envisioned.

Furthermore, the experiment is taken to the next level by exploring NNLM (neural network language model) and BERT (bidirectional encoder representations from transformers). NNLM is a neural network-based language model that focuses on learning the distributed representation from language. This focus helps reduce the complexity of modeling a language due to a large number of features. BERT is a specialized state-of-the-art model for language modeling. BERT is studied in different formats for data modeling. BERT as vectorizer, tokenizer, and other capabilities of BERT are explored. DCNN (deep convolutional neural network) is the following approach explored, followed by BERT's capability to set up a question answering system. Exploration is essential as the overall architecture of this knowledge management system provides the necessary knowledge to the software development team to be handy in this setup. TensorFlow and PyTorch implementation aspects are discussed, later moving on to the possibilities of learning from the source code that holds some core knowledge about the software systems. As a roundup, some of the research gaps and prospects are explored.

The remaining paper is organized as follows. In Section 2, scope of the work is presented. In Section 3, the literature review is presented. Constructs of language models are discussed in Section 4. Data set considered for experiments is demonstrated in Section 5. The experimental setup, which considers various models, is presented in Section 6. Experiments exploration is presented in Section 7. BERT exploration is illustrated in Section 8. Key constructs of the

security knowledge management system are elaborated in Section 9. The learning from source code is demonstrated in Section 10. The concluding remarks are given in Section 11.

2. Scope of Work

Real-world data consists of unbalanced data and unlabeled ones. Unbalanced data includes those data where a particular class of data is predominant over the others. Deep transfer learning plays a prominent role in the space of NLP in these situations. Pretrained models have made the processing of these problems much more efficient and effective. In this study, we explored the security landscape in software development to ease the life of software developers and other team members by providing security-critical information.

3. Literature Review

3.1. Text Analysis. In [1], the effort involved in labeling the text data was reduced using supervised learning. Kohonen self-organizing map (SOM) was employed for labeling the data. Accuracy of classification was validated using a decision tree, Naive Bayes, support vector machine, and classification and regression tree algorithms. In [2], the text classification approaches were surveyed for unstructured mining data. The strengths, weaknesses, opportunities, and threats (SWOT) were explored to know the trend of their usage. Software security vulnerabilities are loopholes in the software system that can compromise the data within the system. Some of the examples can be missing authentication for an important function and missing data encryption.

In [3], text data mining was done using the back and forth matching (BFM) algorithm to make the pattern matching a faster process. In [4], the optimized named entity recognition (NER) approaches were explored for expectation-maximization with semisupervised learning approaches. In [5], authors identified different requirements for linguistic analysis such as linguistic rules, incorporation of NLP, and so on. In [6], authors consider the need of experimenting with NLP or machine learning or other text analysis approaches as a separate focus area, beyond their attempt to run a combination of these techniques. Formalizing the structure of the software requirements can help standardize the conditions outlined for the organization to ensure that security-related requirements get their focus right at the beginning of the software development, which can potentially reduce the cost and effort.

3.2. Software Processes Analysis. Technical debt is an essential consideration in this paper. Some studies identified the influencing factors for technical debt in software systems. These are gathered during the software development process when there is an attempt to balance the customer's strategic and short-term needs.

The anomaly detection method was introduced in [7], which leveraged the optimized mechanism of routing the raw features of the problem area inside a Boltzman machine algorithm. In [8], a software fault detection and correction modeling framework was proposed in software testing. In

[9], a software effort estimation approach was proposed for the agile software development model. Artificial neural network feedforward, backpropagation neural network, and Elman neural network were employed. In [10], a variety of nonfunctional requirements were considered for modeling and to set up an appropriate pipeline to update the criteria and send it across to the next phase [11].

In [12], the security controls were identified using automated decision support. These controls can be relevant to any specific system. In [13], a security requirement elicitation approach based on problem frames was proposed. It considered the incorporation of security into the software development process at an early stage. It helps the developers gather information regarding security requirements in an efficient way. A security catalog was prepared to identify the security requirements, and threats were accessed using abuse frames.

In [14], machine learning techniques were assessed to identify the software requirements for stack overflow. It showed that latent Dirichlet allocation (LDA) was used widely to identify the software requirements. In [15], a tool was devised to discover the vulnerabilities based on the features of software components. The software component features represented the domain knowledge in other software development domains. Since the approach prescribed targets to predict the vulnerabilities in a new component, there is potential to leverage the history of the vulnerabilities for these software components in the production. There is also the potential of taking these solutions and integrating them into the development environment for ease of usage to software developers. The overall emphasis is to make software systems secure right from their inception. It can cut down on the need for the significant investment made by companies for the security of the software.

3.3. Machine Learning. The fixing time of a security issue within the project significantly impacts the overall development process. Therefore, it is essential to fix the issue within a timeline. In [16], machine learning models were used to predict the fixing time. In [17], long short-term memory (LSTM) technique was used to classify the spam. This technique can learn abstract features automatically. In this, the text was changed into semantic word vectors using ConceptNet and WordNet. After that, spam was detected using LSTM from the data. In [18], the accuracy of the K-nearest neighbor (K-NN) algorithm was improved in

classification tasks. However, it takes a longer time in large data sets but provides significant accuracy as compared to others. In [19], the security requirements mentioned in the software requirement specification document were mined. These requirements were classified as data integrity, cryptography, access control, and authentication using a J48 decision tree. After that, prediction models were developed for each security requirement. The pretrained models can also identify the wrongly classified requirements in the document to provide better insight to the requirements engineer. The further refinement will give users the classified information on requirements and explain why a particular classification was chosen. It is a challenging issue with a neural network that needs a better approach to make an interpretable model.

4. Constructs of Language Models

It is essential to understand the construct of the language model that effectively applies to solve the problems associated with software security issues. Problem related to software security can be solved by leveraging the natural language data that is available across companies and industries. Effective language modeling capabilities can help derive the information hidden in these data. Further from this, security-related information can be leveraged by the software development team as and when they need it. Language models are the basis of the models that used in this paper. Language models find their roots in the N-gram modeling approaches. N-gram modeling uses the thought process of assessing the probability of a given the word using its history [20]. For example, the probability of the next word in the phrase “Jack and Jill went up the” to be “hill.”

$$P(\text{hill}|\text{Jack and Jill went up the}). \quad (1)$$

One of the approaches used to compute this probability is relative frequency. By taking the corpus of language as a base, how often the word “hill” follows the phrase can be calculated as follows:

$$P(\text{hill}|\text{Jack and Jill went up the}) = \frac{C(\text{Jack and Jill went up the hill})}{C(\text{Jack and Jill went up the})}, \quad (2)$$

where C represents the count of occurrence of the phrase. The chain rule of probability is applied to words to obtain the following expression:

$$\begin{aligned} P(W_{1:n}) &= P(w_1)P(w_2|w_1)P(w_3|w_{1:2}) \dots P(w_n|w_{1:n-1}) \\ &= \prod_{k=1}^n P(w_k|w_{1:k-1}), \end{aligned} \quad (3)$$

where w denotes the word, n represents a word count, and k represents the length of the sequence. To simplify the complexity of dependencies on the word, Markov

assumptions are employed [21]. This assumption emphasizes that the probability of future prediction can be done with just a few instances from history. In the case of N-gram,

it will be enough to look at $n - 1$ previous words. Maximum likelihood estimation is used for calculating the probabilities of N-grams as follows:

$$P(w_n | w_{n-1}) = \frac{c(w_{n-1}w_n)}{w_{n-1}}. \quad (4)$$

5. Data

This paper takes the data set from a software development team that works on an insurance domain project. All the customer requirement-related data and internal software development process-related data such as test cases and defects are taken and labeled as security- and nonsecurity-related data classes. Thus, the data set contains the text and corresponding labels. Data set contains natural language data obtained from customer requirements specifications, test cases, defects, and other software development work maintained by the software development team. Data set is divided into training, testing, and validation in the ratio of 50%, 40%, and 10%, respectively. The software development requirements management experts labeled the data that are associated with customer requirements. Software development technical leads are involved in labeling the data associated with software development work, such as defects and test cases. Data set is randomly split into training, testing, and validation using the train-test split library of python. We have considered various fractions of training and testing data. It is found that when the training data set fraction is 50%, then the model does not suffer from over- and underfitting issues. There are 31,342 data points, with 3,082 security-related ones and 28,260 nonsecurity-related data points. Therefore, data augmentation techniques such as back translation [22] and easy data augmentation [23] are used to balance the data set. Since some of the deep learning approaches are explored, basic text cleaning methods are only applied. NLTK is used for text cleaning purposes.

Preprocessing of the data includes tokenization of the text to create a vocabulary. Text is first converted to sequence of words and then converted to sequence of numeric IDs. Tokenization of the text involves the conversion of the text into numerical representation so that these representations can be passed into machine learning or deep learning models for modeling purposes. Text sequence padding is also done to normalize the text sequence length. Data are represented as a vector sequence for further modeling.

6. Experiments Set Up

In the first experiment, CNN is explored in text classification [24–26]. Algorithm 1 demonstrates the various steps involved in text classification using CNN. Firstly, text data is tokenized using the TensorFlow Keras preprocessor. Text sequences are transformed into a sequence of numeric IDs. So there will be three sets of text sequences for training, testing, and validation.

Sentence length distribution is visualized to check the length distribution. The maximum sequence length for padding can be kept at 250, as most of the long sequences fall

within a sequence length of 250. TensorFlow Keras preprocessor is used to create text sequences of 250 lengths for all three sets of the data set.

In the next stage, a pretraining-based fastText embedding matrix is explored [27–29]. fastText is an open-source lightweight library that helps learn the text representation in the language models. fastText is configured as a matrix of the data that it has already learned during its pretraining, in the form of embeddings of the numerical representation. The pretrained model, “wiki-news-300d-1M-subword.vec.zip,” provides 1 million word trained vectors from the information of Wikipedia. fastText works are similar to word2vec, where each word is considered for its bag of character-based N-grams. Pretrained word embedding architecture is built in a standard way. An embedding size of 300 is chosen for this pretrained model that is constructed on 300 dimensions. TensorFlow Keras-based CNN model architecture is built (see Algorithm 1). Three sets of Conv1D and max-pooling layers are built using 256, 128, and 64 filters in each of the Conv1D layers and a pool size of 5 for the max-pooling layer. The activation function applied is “ReLU” (rectified linear unit) [30, 31]. The architecture has three sets of dense and dropout layers, with a dropout set at 25%. Binary cross-entropy loss [32, 33] and Adam optimizer [34] are configured for model compilation. Model architecture is run on the training data set. Though the model is configured to be run for epochs count of 100 and batch size of 128, with early stopping, the model reached optimum accuracy on 7th epoch, with validation accuracy of 95.95%. Model performance is evaluated on the test data set; it showed an accuracy of 71.89%. A weighted average of precision is 0.88; recall is 0.72; and F1-score is 0.77. Without much fine-tuning of the parameters, the model would provide an accuracy of 71.89%. Therefore, these architectures can be further fine-tuned for the insurance data to achieve better accuracy.

The experiment is further topped up with a bidirectional LSTM and attention layer (see Algorithm 2). The embedding layer is retained with a pretrained FastText model. Tokenization, vectorization, and padding are conducted similarly to the earlier part of the experiment. The output from the last layer of the long short-term memory gated recurrent unit (LSTM GRU) is fed into the global attention layer sequence.

$$e_t = a(h_t) = \tanh(W h_t + b), \quad (5)$$

$$\alpha_t = \text{softmax}(e_t) = \frac{\exp(e_t)}{\sum_{k=1}^T \exp(e_k)}, \quad (6)$$

$$c = \sum_{t=1}^T \alpha_t h_t. \quad (7)$$

Vectors from the hidden sequence are passed on to a learning function (h_t), including a product vector. c is the final context vector, and T is the total time steps for the input sequence. Attention layer architecture is based on the TensorFlow Keras attention mechanism for temporal data and masking. TensorFlow is a machine learning library, and Keras is the high-level API of TensorFlow. Attention

Input: security- and nonsecurity-related text with labeling
 Process:

- (1) Tokenization of text to create vocabulary:
`t = tf.keras.preprocessing.text`
`Tokenizer (oov_token = "<UNK>")`
- (2) Conversion of text to sequence of words further to sequence of numeric IDs: `train_sequences = t.texts_to_sequences (normalized training text)`
- (3) Sentence length distribution visualization
- (4) Text sequence padding:
`tf.keras.preprocessing`
`sequence.pad_sequences ()`
- (5) FastText-based embedding matrix construction
- (6) Model architecture construction:
`tf.keras.models.Sequential ()`
- (7) Training and validation
- (8) Model performance evaluation on test data

Output:
 Accuracy: 71.89%
 Precision: 0.88
 Recall: 0.72
 F1-score: 0.77

ALGORITHM 1: Text classification using CNN.

Input: security- and nonsecurity-related text with labeling
 Process:

- (1) Step 1 to 4 as in Algorithm 1
- (2) Global attention layer architecture construction
- (3) Entire sequence is sent to global attention layer instead of sending the last output from GRU cell ((5))
- (4) Learning function is fed with hidden sequence vectors ((6))
- (5) Production of a probability vector α_t
- (6) Weighted average of outcomes of above two steps results in a context vector ((7))
- (7) Attention layer definition
- (8) FastText-based embedding matrix construction using "wiki-news-300d-1M-subword.vec"
- (9) Building LSTM-based sequential model architecture: `bigru = tf.keras.layers.Bi-directional (); model = tf.keras.models.Model (inputs = inputs, outputs = outputs)`
- (10) Training and validation
- (11) Model performance evaluation on test set

Output:
 Accuracy: 84.33%
 Precision: 0.91
 Recall: 0.84
 F1-score: 0.87

ALGORITHM 2: Classification using bidirectional LSTM and attention layer.

mechanism is the cutting-edge approach for leveraging the learning from essential parts of the language rather than trying to learn everything. TensorFlow and Keras's attention mechanism provides tool for implementing this capability in the model. Constructs used in this experiment are taken from [35]. FastText-based embedding is built as in the first part of the experiment. Core model architecture is built with LSTM to form the sequential models. LSTM is better than RNN in remembering the long sequence of data. LSTM manages input at the current time step, with the output of the previous LSTM unit and memory of cell state in the previous unit.

Bidirectional LSTMs help feed both forward and backward sequences of the content. The output provided by each of these is combined at every time step. By considering the past and future sequences, a better context of the text is retained. The architecture encompasses embedding as input, bidirectional LSTM GRU with 256 units, attention layer, and three sets of dense and dropout layers (see Algorithm 2). Two hundred and fifty-six units for thick layer and 0.25 dropout rate are also used alternatively. "Relu" is the activation function used for intermediate layers; in the final layer, "sigmoid" is used. Binary cross-entropy loss and "Adam" optimizer are also considered. Model is trained with

an early stop mechanism for 100 epochs feeding in training and validation. A batch size of 128 is retained; best results are reached at the 6th epoch, with a validation accuracy of 98.69%. The model achieves an accuracy of 84.33% with weighted average precision of 0.91, recall of 0.84, and F1-score of 0.87.

In the next phase of the experiment, Google's Universal Sentence Encoder (USE) is explored. It can encode high-dimensional vectors with varying lengths into a standard size. Figure 1 shows the working of sentence encoding.

In this experiment, USE is implemented on TensorFlow 1.0. USE model is loaded from the TensorFlow hub [36]. The USE-based embedding layer is constructed and fed into the model architecture with two dense and dropout layers (see Algorithm 3). The dense layer has 256 units and a ReLU activation layer, with a dropout value of 0.25. The sigmoid activation function is used to classify the results. Binary cross-entropy and Adam optimizers are also used. Model training is implemented by considering the early stopping. It achieves an optimal validation accuracy of 97.32% in the first epoch. It has performed an average accuracy of 92.61%, with average recall, precision, and F1-score as 93.0%, 95.0%, and 93.0%, respectively.

7. Experiments Exploration

Software requirements modeling is one of the prominent parts of this work. Most of the time, the focus is on software correctness in the development that can lead to performance issues later in the development process. In [37], a comprehensive study of all the work done towards modeling the performance of the software across the software development lifecycle is presented. In [16], the factors that impact the time for fixing security issues with linear regression methods are assessed.

The focus of [38] is to look at machine learning applications for software vulnerabilities management and various data mining approaches. Vulnerabilities discovery models with software metrics, vulnerable code pattern identification, and anomaly detection methods are explored. In [39], the focus is given on technical debt being used as a base for security issues modeling with machine learning. In [40], a focus is presented on applying a set of hybrid codes covering static and dynamic variables that characterize input validation and patterns of input sanitization code, and these are expected to be prominent indicators of vulnerabilities in web applications. There is a good complement between static and dynamic program analyses; both techniques extract the proposed variables in a scalable and accurate method.

Building on the base constructs experimented with within the previous section, this section explores more advanced architectures. NNLM is explored in this section. Sentence encoding is experimented with NNLM, which also provides fixed-length vectors for the documents. The exact format of data processing is continued in this section as well. NNLM model is fetched from the TensorFlow hub from <https://tfhub.dev/google/tf2-preview/nnlm-en-dim128/1> [41]. Keras layer is built from the hub with an output shape of 128. Model architecture is built with Keras layers from

TensorFlow that will have pairs of dense and dropout layers. The dense layer will have 128 units and "ReLU" activation; The value for the "dropout" is chosen to be 0.15, with loss function of "binary cross-entropy"; and 'adagrad' optimizer is determined (see Algorithm 4). Model is fit with the train data for 100 epochs and a batch size of 128. At the end of 100 epochs, the model reached a validation accuracy of 90.17%. The model demonstrated 90.16% accuracy on test data with an average precision rate of 0.81, recall of 0.9, and F1-score of 0.86.

BERT is explored in the next part of the experiment. Simple BERT architecture is shown in Figure 2. BERT undergoes semisupervised learning on a large amount of online data such as Wikipedia. Language understanding capabilities are inherent in these pretrained models that can be leveraged to many of the tasks associated with language processing. These pretrained models can be refined to different data sets with less effort to use their capability concerning the specific domain of application. BERT expects the text content to be tokenized, and in lowercase, Hugging Face's <https://huggingface.co/> [42].

BERTTokenizer, which is part of transformers, is used. The word piece tokenizer approach segments the words into a subword level like any other NLP (natural language processing) task. The tokenizer is formed with the pretrained model, "bert-base-uncased." Data used for this experiment need to be preprocessed with an approach similar to that of the original method used in the BERT pretrained model. Pre-processing includes, the conversion of data in to lower case, tokenization, breaking of word pieces, word to index mapping using BERT vocabulary, adding the separator and end of sequence tokens, and finally appending, 'mask' and 'segment' tokens.

The model architecture here consists of the BERT that helps process the input data of text based on its prelearned language capability; on top of this, a feed-forward neural network with softmax is built for customized classification tasks. Models layers are made with TensorFlow Keras layers. A maximum sequence length of 250 was selected for the data used in this experiment, as most of the sentences fall within the range of 250 lengths of words. Input ID, mask, and segment are created with the hidden state made from the pretrained model "bert-base-uncased." The further part of architecture will have dense and dropout layers of two, each alternating, followed by a dense output layer. Two hundred fifty-six units are used for dense layers, and a 0.25 dropout rate is considered. The "ReLU" activation layer is used in dense layers with "sigmoid" activation for the dense output layer. For a model compilation, Adam optimizer is used with a learning rate value $2e-5$ and $1e-8$ epsilon value, and binary cross-entropy loss and accuracy metrics are used.

The utility function that is created for converting the text input data to BERT features is used. The same training set used until this point is used here. To recap, this data is a customer requirement in an insurance domain, and test cases and defects data were created as part of software development operations on this insurance company. Data is labeled as security and nonsecurity data based on expert input. Data is prepared in a similar way to the earlier part of

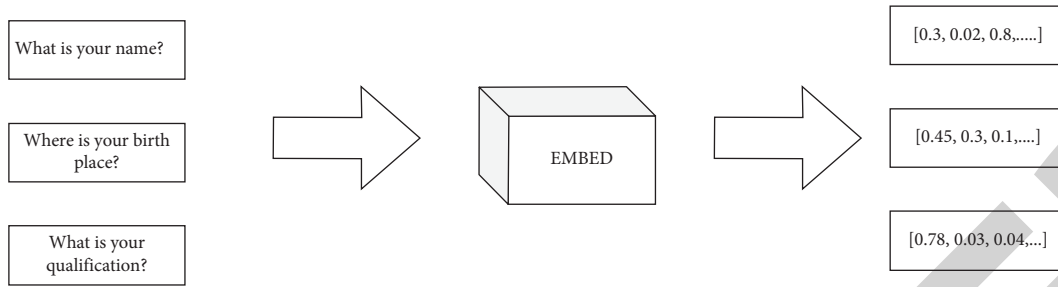


FIGURE 1: Working of sentence encoding.

Input: security- and nonsecurity-related text with labeling
 Process:

- (1) Data set split to 50% for training, 10% for validation, and 40% for testing
- (2) TF 2.0 eager execution is disabled as the Google has not updated USE model for compatibility with TF 2.0 (`tf.compat.v1.disable_eager_execution ()`)
- (3) USE model to be loaded from TF Hub: `embed = hub.Module (module_url, trainable = True)`
- (4) USE embedding layer to be built
- (5) Model architecture to be constructed:
`model.compile (loss = "binary_crossentropy," optimizer = "Adam," metrics = ["accuracy"])`
- (6) Latest version of Google's USE that supports TF 2.0 will be utilized in the further phase of the research
- (7) Training and validation to be conducted for 100 epochs and batch size of 128, with an early stopping approach
- (8) Trained model weights are loaded for inference
- (9) Model performance evaluation on the test data set

Output:
 Accuracy: 92.61%
 Precision: 0.95
 Recall: 0.93
 F1-score: 0.93

ALGORITHM 3: Classification using USE with TensorFlow 1.0.

Input: security- and nonsecurity-related text with labeling process:

- (1) Data set split to 50% for training, 10% for validation, and 40% for testing
- (2) Function for preprocessing the corpus
- (3) Basic text preprocessing
- (4) Embedding layer of NNLM is built: `model = "https://tfhub.dev/google/tf2-preview/nnlm-en-dim128/1"`
- (5) Model architecture is constructed: `model = tf.keras.models.Sequential ()`
- (6) Training and validation is conducted for 100 epochs and 128 batch size with an early stopping method
- (7) Model performance is evaluated on test data

Output:
 Accuracy: 90.16%
 Precision: 0.81
 Recall: 0.90
 F1-score: 0.86

ALGORITHM 4: Neural network language model.

the experiments. Training data is converted to training feature IDs, training feature masks, and training feature segments using the BERT tokenizer constructed earlier. Now the built model is trained using the training and validation part of data passing in the IDs, masks, and segments. Three

epochs and a batch size of thirteen are used with early stopping (see Algorithm 5). The model reached a validation accuracy of 99.11% within the second epoch. The refined model is stored for further usage. Test data is also processed similarly, and predictions are run on the test data. It gave an

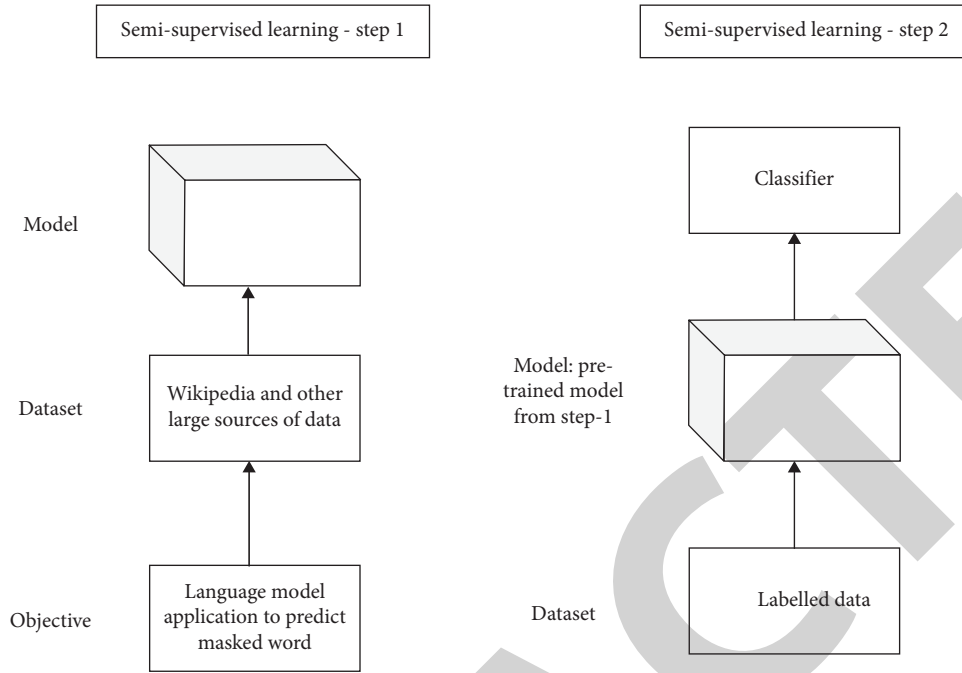


FIGURE 2: BERT architecture.

accuracy of 91.3%, average precision of 0.92, recall of 0.91, and F1-score of 0.88. Since the current focus is to build a pipeline of the knowledge management system for software vulnerabilities management, these experiments are not explored further to make it more customized to the insurance company data targeted here.

Though pretrained models provided ease of modeling with no need to train from scratch, they ended up being of massive size. To tackle this, distilBERT will be explored next. DistilBERT compresses the BERT with the knowledge distillation technique. The distillation technique is a way of reducing the complexity of the BERT architecture to make it concise and small during the pretraining phase. During the distillation, 40% of the size is reduced, retaining 97% of its language understanding and making it 60% faster than BERT. The teacher-student training approach is used where the student is trained to replicate the model output that of the teacher. Hugging Face uses KL-divergence loss to train distilBERT. The approach reduced the number of parameters to almost half compared to BERT retaining the performance. BERT tokenization is conducted similar to the experiment's last part; data preparation and model architecture setup are also the same. Sixty-six million parameters are generated in this case, unlike BERT, which had 109 million parameters. Input features of training and validation set of data are also prepared similar to the last part; tokenizer used is "distilbert-base-uncased." In the case of distilBERT, three epochs are run with a batch size of 20 and an early stopping mechanism (see Algorithm 6). The model reached a validation accuracy of 99.07% at the second epoch. The test data model demonstrated 94.77% accuracy, average precision of 0.95, recall of 0.95, and F1-score of 0.94.

With these basic building blocks, the AI-based knowledge management system for software security

vulnerabilities can be conceptualized. The system will have three prominent parts to it. Software security model architecture is shown in Figure 3. Customer conversation modeling will take in all the content generated as part of the conversation with the customer and use it to learn the security needs for software. The second part will be industry landscape modeling that will depend upon all the knowledge sources in the industry and leverage the same to understand security from an industry expert's point of view. In the third part, software landscape modeling, all the information within software development processes is leveraged to build security knowledge for the software development team. Exploration done so far will help construct the customer landscape and industry landscape modeling.

Deep transfer learning constructs have been explored in the experiment so far that can help build this knowledge system for the insurance-based company's software development team. Pretrained word embeddings of deep learning models are explored, covering FastText with CNNs or bi-directional LSTMs with attention layer, universal embeddings with sentence encoders, and neural network language model. Transformers are explored, covering BERT and DistilBERT. TensorFlow 2.0 is used as a base to leverage its capabilities.

The attention mechanism is an interesting thought process and discussed in [16, 43]. The author calls the attention they have used in their architecture "scaled dot-product attention."

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \quad (8)$$

where "Q" stands for queries, "K" stands for keys, and "V" stands for values. Attention function is specialized with

```

Input: security- and nonsecurity-related text with labeling process:
(1) Tokenization with BERT: tokenizer =
    transformers.BertTokenizer
    from_pretrained ("bert-base-uncased")
(2) Data preparation for the need of BERT, including lower casing of text, tokenizing, word split to word pieces, word to index
    matching with vocabulary file of BERT, add special tokens, and adding mask and segment tokens to each input
(3) Model architecture building with TF:
    model.compile (optimizer = tf.optimizers
    Adam (learning_rate = 2e - 5, epsilon = 1e - 08), loss = "binary_crossentropy," metrics = ["accuracy"])
(4) Maximum sequence length set to 250
(5) TF Keras layers are built for model compilation
(6) Text converted to BERT input features:
    create_bert_input_features (tokenizer,
    train_text, max_seq_length)
(7) Data set split to 50% for training, 10% for validation, and 40% for testing
(8) Create function for BERT input features creation
(9) Feature IDs, feature masks, and feature segments are created for training and validation
(10) Model is trained and validated
(11) Test review data are converted in to BERT input features
(12) Model performance is evaluated with test data: from sklearn.metrics import
    confusion_matrix,
    classification_report, accuracy_score
Output:
Accuracy: 91.39%
Precision: 0.92
Recall: 0.91
F1-score: 0.88
    
```

ALGORITHM 5: BERT for tokenization and feature creation.

```

Input: security- and nonsecurity-related text with labeling
Process: Except for distilBERT tokenizer being used rest of the steps are same as Algorithm 5
tokenizer = transformers.DistilBertTokenizer
from_pretrained('distilbert-base-uncased')
Output:
Accuracy: 94.77%
Precision: 0.95
Recall: 0.95
F1-score: 0.94
    
```

ALGORITHM 6: DistilBERT for tokenization.

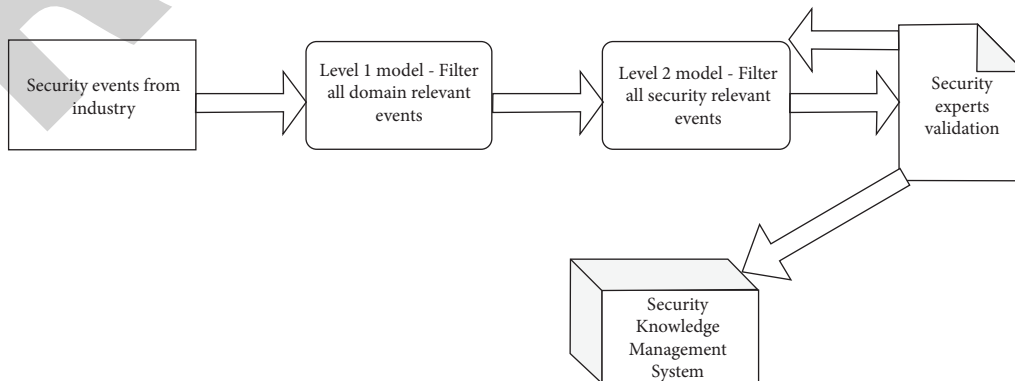


FIGURE 3: Software security model architecture.

mapping queries to a set of key-value pairs, further to an output. All the entities query, key, values, and output are vectors. Input encompasses queries and d_k , which are keys of dimension, and d_v , which are the dimension values. Authors also have explored the multihead attention mechanism as follows:

$$\text{Multihead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_n)W^o, \quad (9)$$

where $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$.

Multihead attention facilitates the model to access information jointly from a variety of representations. “ W ” here is the parameter matrices; “ W^O ” represents the parameter matrix for the output; and “head” is an attention head. The author also uses a position-wise feedforward network in their architecture. This network is used in encoder and decoder layers that are applied to each position separately and identically.

$$\text{FFN}(x) = \max(0, xW_1 + b_1)W_2 + b_2, \quad (10)$$

where “FFN” is feedforward network, “ W ” stands for weight, bias is given by “ b ,” and “ x ” is the input.

The next phase of the experiment industry landscape and customer landscape can be modeled by building the following pipeline. Any events about the software processes that include security and nonsecurity will feed into this knowledge system; the first level of the task would be to screen the relevant events for the software teams to understand the need for software security. All the events that are not associated with security have to be removed. The next part of the pipeline should categorize the events further into security-related events and are closely associated with the insurance domain, which is the focus domain of this research. As the system is in the initial stage of development, security experts can be involved in screening the outputs generated by the model to act as an input for the model to calibrate itself.

8. BERT Exploration

In [44], researchers have shown that sentimental analysis of software artifacts can improve various software engineering activities. BERT has been leveraged for language modeling natural language text in Stack Overflow posts, subjected to sentiment analysis. In the view of leveraging on the wealth of knowledge hidden in the open-source software ecosystem, authors in [45] have proposed software entity recognition methods based on BERT word embedding. In [46], the author proposes automatic text classification based on BERT to develop an environment for the internet of things. The approach used here uses text-to-dynamic character level embedding with BERT. Furthermore, Bi-LSTM and CNN output features are utilized to leverage the CNN for local feature extraction and bi-LSTM for its capability on memory to link back the extracted feature.

In [47], the author focuses on utilizing BERT’s self-attention mechanism for bidirectional context representation for predicting the resolution time of the bugs, which can

contribute towards a better estimation of the software maintenance effort and cost.

BERT provides capabilities such as tokenizer for data preprocessing, creating word embedding, and constructing the question answering system. These parts have to be validated with the data from the insurance domain where the knowledge management system is built. BERT is the latest specialized approach; leveraging its capabilities will help exceed other methods for the targeted use cases in this space. As per the approach followed in earlier sections, the experiment will be built step by step to explore each capability of BERT. In the first part, the BERT tokenizer’s data processing capability has been experimented. TensorFlow 2.0 is leveraged here as well. The training data set composed in previous sections of the experiment will be used in this set of experiments. Data will have text derived from customer requirements, test case descriptions, and defect descriptions for insurance domain-related software development processes. The text would have security- and nonsecurity-related content; the other column is a label that classifies security and nonsecurity classes of the text by the experts. Security-related contents are those where information related to software security is available. This security-related information will help understand the security controls that are needed in the software systems. Nonsecurity-related content is the rest of the information related to software system development. Text preprocessing is done to remove any nonalphanumeric content, “https://” from any URLs (Uniform Resource Locator), and punctuations and white spaces, if any. BERT layer has to be created so that metadata associated with the same can be assessed; vocabulary size is metadata. The tokenizer is designed with features of a vocabulary file and lower casing. Vocabulary file is derived from BERT layer that is constructed from Keras layer from TensorFlow hub. The text embedding feature used is “bert_en_uncased_L-12_H-768_A-12” [48], with a trainable parameter set to “false” as the feature will be used as is without further training it for this experiment sake. The tokenizer is used to convert tokens to IDs bypassing the sentences of the cleaned data into the same.

Further data set creation will involve adding padded batches of sentences; each batch of the sentence is padded independently. Padding helps optimize the padding tokens. Their length sorts sentences, and then padding is applied and then shuffled. Data set are transformed as tensors using the “from_generator” method of TensorFlow (tf_data.Dataset). A batch size of 32 is chosen for performing padding. Model building involves DCNN with the following architecture. Table 1 shows the training specification of DCNN architecture.

DCNN model construct is expressed in Table 2. The further architecture will have layers built as follows.

All the layers formed in Table 2 are merged by passing in the training data. DCNN is fit with the training data. Test data produces an accuracy of 97.44%. DCNN model can be used to pass in the text and predict whether the text is security related or nonsecurity related, with a confidence level on the scale of 0 to 1 (see Algorithm 7).

TABLE 1: DCNN architecture training specification.

Parameter	Value
Vocabulary size	Based on the vocabulary size of the tokenizer
Embedding dimension	200
Number of filters	100
Feedforward network units	256
Number of classes	2
Dropout rate	0.2
Epochs	1
Training	False

TABLE 2: DCNN model construct.

Layer	Configuration
Embedding	Vocabulary size, embedding dimension
Convolutional 1D bigram	Kernel size = 2; padding = valid; activation = ReLU
Convolutional 1D trigram	Kernel size = 3; padding = valid; activation = ReLU
Convolutional 1D fourgram	Kernel size = 4; padding = valid; activation = ReLU
Pooling	GlobalMaxPool1D
Dense	Activation = ReLU
Dropout	0.2
Last dense (for 2 classes)	Units = 1; activation = sigmoid
Last dense (for more than 2 classes)	Units = number of classes; activation = softmax
Loss (for 2 classes)	Binary cross-entropy
Loss (for multiclass)	Sparse categorical cross-entropy

<p>Input: security- and nonsecurity-related text with labeling</p> <p>Process:</p> <ol style="list-style-type: none"> (1) Data preprocessing and tokenization to create a BERT layer: FullTokenizer = bert.bert_tokenization FullTokenizer bert_layer = hub.KerasLayer ("https://tfhub.dev/tensorflow/bert_en_uncased_L-12_H-768_A-12/1," trainable = False) (2) For data set creation purposes, padding of the data batches to be done, to bring all the training sequence to a consistent length (3) Test and train data set batches are created (4) DCNN model building as per the specifications provided in the tables above (5) Training the model with the specifications provided in Table 2 (6) DCNN model compilation: DCNN(tf.keras.Model) DCNN = DCNN (vocab_size, emb_dim, nb_filters, FFN_units, nb_classes, dropout_rate) (7) Fit the model with training data (8) Model evaluation with test data <p>Output: Accuracy: 97.44%</p>

ALGORITHM 7: BERT with DCNN model.

The second part of this entire exploration approach would remain the same except for the additional step of creating word embeddings. Embedding creation includes the creation of IDs, masks, and segments. IDs are created by using the "convert_tokens_to_ids" function from the tokenizer. Masks are created on tokens using NumPy by padding to tokens, and segments are created appending the tokens into segments. An accuracy of 99.58% was the result with the BERT tokenizer and embedding being applied.

In the third part, architecture is refined to build a question answering module using BERT. Some of the

libraries used for this experiment are as follows. BERT question answering system library specifications is presented in Table 3.

The SQuAD (Stanford Question Answering Dataset) is a well-known data set to explore question answering systems. Input metadata creation is the first step, which takes training data in JSON (JavaScript Object Notation) file and vocabulary file as input. Training data consists of a paragraph of information, which provides context, a question from that text, and answers for those questions. TensorFlow record is created from these files. The training data set is created using

TABLE 3: BERT question answering system library specifications.

Modules	Library
Tensorflow	Tensorflow hub
official.nlp.bert.tokenization	FullTokenizer
official.nlp.bert.input_pipeline	create_squad_dataset
official.nlp.data.squad_lib	generate_tf_record_from_json_file
official.nlp	Optimization
official.nlp.data.squad_lib	read_squad_examples
official.nlp.data.squad_lib	FeatureWriter
official.nlp.data.squad_lib	convert_examples_to_features
official.nlp.data.squad_lib	write_predictions

the “create_squad_dataset” library, passing the TensorFlow record of the data produced. “BertSquadLayer” is constructed with TensorFlow Keras dense layers. The complete model of BERT will have a class of “BERTSquad” that will take in the BERT layer from Keras layer of hub (https://tfhub.dev/tensorflow/bert_en_uncased_L-12_H-768_A-12/1). The trainable parameter is kept as “true.” BERT application function of this class will take in input word IDs, input masks, and input type IDs. As part of the training phase following configuration is set. Table 4 shows BERT question answering system training phase specifications.

These are computationally intensive experiments; parameters are optimized based on the need of these experiments. Once the pipeline of this knowledge management system is organized, it can be further customized to improve effectiveness. The training data set is also made lighter to facilitate an easy start for the experiment. Optimizer is created using the function “create_optimizer,” and SQuAD loss function is created to compute loss. BERT SQuAD model is compiled passing in optimizer and SQuAD loss function. Training is conducted passing in training data sets in batches as inputs and targets. BERT SQuAD layer and loss function is passed in, and gradients are applied on optimizer for training purposes. At the end of one epoch of training, training loss was reduced from 5.94 to 2.50. Next is the evaluation phase, where the development data set is used. Evaluation examples are created using the “read_squad_examples” library. TensorFlow record format of this evaluation data set is generated. BERT tokenizer that was created earlier is used. The function is created to add features into the evaluation feature list that is created (see Algorithm 8). Once the evaluation feature is created, the TensorFlow record is generated for the same. Then SQuAD is created from this evaluation data set. Thereafter, prediction is utilized to build the dictionary-like collection. The “NamedTuple” function of the “collections” library is used to create a dictionary-like collection. It will generate batched output in a timely way. Collection, evaluation examples, and evaluation features are passed into the “write_prediction” function of “official.nlp.data.squad_lib,” which generates prediction output files.

The evaluation script has a function for normalizing the answers, generating an F1-score, defining the exact match score, comparing the results with ground truth, and evaluating the data set with the predicted values. The evaluation script must be run with the development data by passing in

predictions generated in previous steps. Step here produces a result with an F1-score of 77.26 and an exact match of 66.91%. The experiment was conducted with limited resources to understand the approach. Once the construct of the pipeline is worked out, these approaches can be further refined and customized to target domain data that is software development practices of the insurance domain.

Based on the comparative analysis, Table 5 provides the comparison of the accuracy of the models. Table 5 also explains the key processes involved in the modeling, which provides information about the complexity of the approaches. Now, this setup can be utilized to predict any of our data. Question and context text must be concatenated with a separator token after tokenization like it was done for training content. Utilities required for this phase of prediction are the BERT layer from the Keras layer of the hub, which takes in the pretrained BERT model like in earlier phases. A comparison of all the experiments is provided in Table 5. Vocabulary file is generated from this BERT layer, including the lower casing function. Both vocabulary file and lower casing function are passed into “FullTokenizer” to generate tokenizers. Other utilities include white space recognizer, text to words converter, tokenizing each word, and keeping track of the tokens and words. Processing is continued with creation of IDs, masking, and segment creation from tokens, and finally a function is created to take a question and context as input, and return a dictionary with three elements as expected by the model. Expected output are context words, correspondence between context tokens to context word IDs, and length of the question tokens. Answers can now be predicted by passing in question and context to create an input dictionary run on BERT SQuAD trained earlier. Some more refinement can be done to organize the interpretation for the reading answers as output for the input provided in context and question. The model was used to pass in the context and question related to software applications security from the target domain, and it produced the answers for the given context.

Question answering systems can play a handy role in the intended knowledge management system for software security in the insurance domain, the primary focus area.

CNN is one of the prominent areas of exploration in the domain explored in this paper. Some of the exciting work done with CNN is as follows. Exploration in [49] and [50] aims at devising an improvised CNN-based approach to improve the bug localization task in software engineering. In

TABLE 4: BERT question answering system training phase specifications.

Parameter	Value
Training data size	88 641
Number of training batches	500
Batch size	1
Number of epochs	1
Initialized learning rate	$5e-5$
Warmup steps	10% of the number of training batches

Input: SQuAD data set process:

- (1) Data preprocessing to create input meta data from SQuADdatasetinput_meta_data = generate_tf_record_from_json_file()
- (2) Building BERT SQuAD layer:
BertSquadLayer (tf.keras.layers.Layer)
- (3) BERT questions answering system training phase configuration
- (4) Create a SQuAD loss function for further computation
- (5) Training and evaluation

Output:

F1-score: 77.26

Exact match: 66.91%

ALGORITHM 8: SQuAD dataset process.

TABLE 5: Comparison of all the experiments.

Experiment	Process	Output
CNN and FastText embedding	CNN-based processing	Accuracy: 71.89%; precision: 0.88; recall: 0.72; F1-score: 0.77
Bidirectional LSTM with FastText embedding	Bidirectional GRU or LSTM with global attention	Accuracy: 84.33%; precision: 0.91; recall: 0.84; F1-score: 0.87
USE model	USE pretrained model with TF 1.0	Accuracy: 92.61%; precision: 0.95; recall: 0.93; F1-score: 0.93
NNLM	NNLM-based sentence encoder, with pretrained model	Accuracy: 90.16%; precision: 0.81; recall: 0.90; F1-score: 0.86
BERT	BERT tokenization and TF Keras modeling	Accuracy: 91.39%; precision: 0.92; recall: 0.91; F1-score: 0.88
DistilBERT	DistilBERT-based preprocessing of data	Accuracy: 94.77%; precision: 0.95; recall: 0.95; F1-score: 0.94
BERT	Data preprocessing and tokenization with BERT	Accuracy: 97.44%

[49, 51], a bidirectional LSTM algorithm is proposed based on CNN and independent RNN for malicious web page identification. Word2vec is used for training URL word vector feature for modeling. Input sequence illustration in BERT represents in Figure 4.

Attention models are another critical component explored in this paper. Attention layers are techniques used in a neural network for processing the input, which facilitates the process of focusing on a specific aspect of complex information. In the work of Software System Security Vulnerabilities Management, modeling of customer conversation, industry knowledge bases, and knowledge gathered in software development processes are done. Since the focus is on the security dimension of the content, attention models will help build the required focus around security. Some of the other interesting results using attention models are as follows. The inability of the software defect prediction at the granular level of code hinders the possibility of

providing detailed information to developers. In [51], the ensemble learning techniques and attention mechanism is utilized to provide comprehensive information to developers pointing at the suspect line of code and method-level defect prediction. The lack of focus on nonfunctional requirements like usability and security is handled ad hoc and results in cost. In [52], modeling of nonfunctional requirements across the software product life cycle is explored. Pseudocode generation is one of the essential aspects of software engineering, as it involves a lot of effort. In [53], there is an effort to treat pseudocode generation tasks as a language translation task that involves programming language translation to natural language description using the neural machine translation model and the attention seq2seq model. Bag of words with Naive Bayes approach shows the accuracy of 90%, precision of 77.7%, and recall of 93%. Classification of unsupervised learning approach shows the accuracy of 73%, precision of 69.3%, and recall of 77.7%. Table 5 shows a series

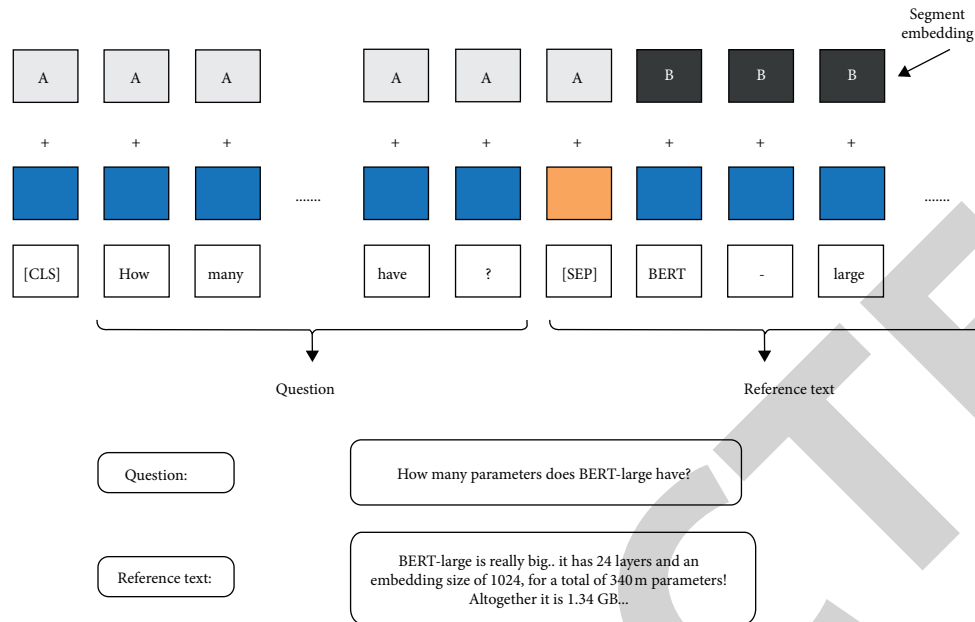


FIGURE 4: Input sequence illustration in BERT.

of experiments conducted in this work that shows improved performance on security issues classification as the approach was tuned for data from the insurance domain. It started with accuracy of 71.89%, with CNN and FastText embedding. Performance went up to 84.33% with bi-directional LSTM and FastText embedding. USE model made up to the accuracy of 92.61%; NNLM model made up to the accuracy of 90.16%; and BERT model would further refine the accuracy up to 97.44%. Experiments are started with conventional CNN with pretrained FastText embeddings. Pretrained FastText embeddings are further combined with bidirectional LSTMs and attention layers. Bidirectional LSTMs provide the ability to read the inputs from both directions and enhance learning. USE approach provides the ability to encode texts of any length in higher dimensions vector and is useful to try. NNLM model provides the ability to encode the sentence with a fixed length of vectors with pretrained layers for predicting the next word. BERT and DistilBERT are tried at the end, as they are the cutting-edge approaches and provide the ability to fine-tune the pretrained model to suit the targeted domain data. DistilBERT also has played a key role in reducing the size of pretrained models. The system proposed in this work focus on effectively handling security aspects in software development processes. The central theme of this exploration is modeling the data from customer conversation, information on public websites from the organization that works on educating the industry on security vulnerabilities, and data internal to software development processes. Most of this information is in natural language data, so it is beneficial to explore NLP-related approaches to tackle the problem. With the latest developments in the NLP space, some of the advancements around language models fit well to solve some of the areas targeted in this work. Advanced NLP will help pool all the

unstructured data in and around software development and learn patterns from them. This aspect can be leveraged to solve some of the challenges software development processes face around the quality of work, productivity, and process maturity. In the further set of experiments, a 95% confidence interval will be considered for accuracy, F1-score, recall, and precision. This set of experiments being the first phase to evaluate the suitability of the models in the selected domain; all the metrics are reviewed without specification of the confidence interval. Area under curve (AUC) also will be included as a metric for evaluation in the next set of experiments.

9. Key Constructs of Security Knowledge Management System

As part of building a knowledge system that takes in data and creates valuable information, some of the critical properties to be accounted for are discussed here. As part of BERT, labeling the training data and tokenization results in loss of traceability of the original word and the token. As the answer to the question is explored, it is crucial to figure out the position of these words that form the response from the context statement. One of the approaches would be to identify the answer string with a unique identifier that can be easily identified from various identifiers. TensorFlow and PyTorch being the approaches available, for application of transformers, Hugging Face provides the PyTorch interface. Pytorch sticks to API (pplication programming interface) provision, not worrying about the internal workings of the approach, whereas TensorFlow provides insight into the inner workings of the approach, but that may sidetrack from the primary intention of this approach. The distinction between the workings of TensorFlow and PyTorch is made

based on the earlier experience of working with PyTorch. BERT also provides various classes that intend to perform different tasks like token classification, question answering, next sentence prediction, and so on. Choosing the maximum length of the sequence is critical to the trade-off for the computational expense involved. If the maximum length is 512, it takes 4x longer to train than the length of 256 and 16x more time for 128. Hugging Face has used 384 lengths for the sequence in its implementation.

Truncation of the sequence results in a loss of answers for some of the questions. This is because, generally, answers are the last part of the sequence of sentences. In case if sentences are cut off during truncation, there is a higher possibility of losing the answers for some of the questions.

One possible way to handle this would be to truncate the context sequence from the beginning instead of the end. This truncation would involve significant effort, and it needs to be traded off for the returns from this. One of the ways would be to skip the questions where the context was truncated. Ninety-seven training examples were lost due to truncation from among 87,502 samples of the SQuAD. Fine-tuning of the BERT model involves predicting a category to call out if the token is identified correctly for the start of the span and another for the end of the span. Fine-tuning of the model is an attempt to increase the expertise of the model from being good at the text to the capability of question answering.

10. Learning from Source Code

In [54], the authors proposed an adaptive deep code search method for training once and then reusing the same in new code bases. This approach optimizes the need for training the codebases every time for search purposes. Though there are many programming data available in online sources like Stack Overflow, there is a lack of sound natural language processing-related approaches to extract code tokens and software-related named entities. In [55], the named entity recognition approach is proposed on code for named entity recognition. In [56], there is an exploration of topic modeling to mine unstructured software engineering data. This work surveyed topic modeling applications and identified an increasing need to focus on tasks associated with comprehension of source code and software history. In [57], code flaws and vulnerabilities modeling are focused on using the deep learning-based long short-term memory model, focusing on learning semantic and syntactic features of the code. In [58], a systematic literature review for multilanguage source code analysis is presented. This study helps explore the focus areas for development like static source code analysis, refactoring, detection of cross-language links, and other vital areas. In this study, statistical modeling of the source code is one of the focus areas. This statistical learning and representation of the source code can be based on these applications like static source code analysis, refactoring, and detection of cross-language links. Here, cross-language means multiple software programming languages. Language modeling study will be leveraged to model the natural language-related data that gets accumulated during the process of software development. And source code itself has naturalness in it which

is more structured compared to the natural language text. This provides further opportunities for modeling the source code.

In the software development landscape, software source code forms a critical component that potentially holds quite a pattern that can provide insight into potential software security vulnerabilities that can creep into the system. Devising a mechanism to model source code will add to the intelligence built via a software security knowledge management system. The idea would be to derive the pattern hidden in the program by modeling the property of the code. Code2vec is the recent popular approach that attempts to learn the distribution representation of the code. The thought process is detailed in [59], where authors try to present a neural model where the code is represented as a continuous distribution of vectors. This approach will help model inherent semantics property that is at the core of code semantics. Code is in the form of a collection of paths that is part of abstract syntax tree representation. Method name prediction is attempted in this work consuming the vector representation of the body. Twelve million methods of code are used as input for this modeling. In this approach, code snippets provide information to be represented as a bag of context, and a vector representing the context, the value of the same need to be learned. Work here revolves around finding a meaningful way to break up the code into smaller significant blocks and then aggregating these building blocks to arrive at meaningful predictions. Figure 5 expresses the overview of the code2vec.

The general approach shows that a program needs to be taken in as input and broken into meaningful parts, converted to vectors that can then be aggregated into predictions. While the features are devised for this modeling approach, there is a trade-off between the effort involved for the model to learn and the effort involved in building those features. AST approach helps balance these factors by banking on the syntax associated with the code. AST paths are set as vectors here, which will include token vectors and path vectors. Tokens are the entities that are connected by the specific path in AST. All the tokens vectors and path vectors are combined as a matrix for which the tanh function is applied to bring in element-wise nonlinearity to rescale all the vectors between -1 and 1 . A fully connected layer will provide a path context that encapsulates all the patterns in this part of the code sequence. The challenge would be to take in all the path contexts and aggregate the same into a code vector. There are three approaches to aggregate the path context: take the most crucial path context, take an average of all, and take the final one to bring attention and compute the weighted average of all the path context. As part of attention, vectors learn the semantic meaning of the path context and attention needed by the path context. The learned attention vector is the randomly initiated component that learns in parallel in the network. Path context vector and attention vector are combined with dot product to obtain scalar normalized with softmax to get the score that sums to one. Path context vectors are then multiplied by the normalized scalar values summed up to get the code vector. Code vector is the weighted average of the input vectors continuously learned and updated by the network during the

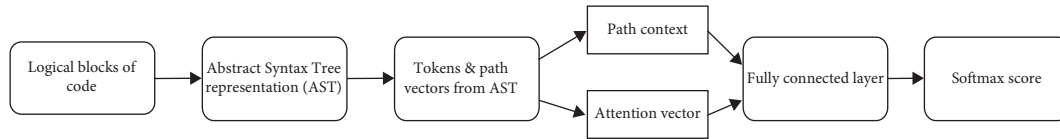


FIGURE 5: Overview of the code2vec.

TABLE 6: Model configuration for BatchProgramClassifier.

Parameter	Value
Hidden dimension	100
Encoding dimension	128
Labels	104
Epochs	15
Batch size	64
Maximum tokens	Based on word2vec embeddings
Embedding dimension	Based on word2vec embeddings

learning phase. Target method name vectors are also trained with 14 million methods examples; these are used as a reference for validating the predictions done by the network.

In [60], the author proposes a neural source code representation with AST. The author here focuses on the fundamentals of extending code modeling for further analysis. Information retrieval methods would consider the code as text and miss the semantics understanding of the source code. ASTs are leveraged by the authors as the better approach to represent the code. ASTs, by nature, end up being long data sets way they represent the code. Instead of working on the entire ASTs, this proposed model works on split AST smaller statement trees.

Furthermore, these are encoded into vectors, which will retain lexical and syntactic knowledge of statements. These statement vectors are modeled with bidirectional RNN, which will use the naturalness of these representations and help optimally represent the code in vector format. Authors have applied this model of code representation for solving code classification problems and code clone detection problems. Code classification approach will be explored to fit into the framework being explored here for smart requirements management. Authors have provided their implementation scripts in <https://github.com/zhangj111/astnn>.

The implementation proposed for the code classification will help model the code data and explore any vulnerabilities in the code. For this purpose, either training data can be built to train further this architecture of specific code data from the projects of the insurance company targeted for this security management framework for software development. Or implementation can be reused on local data as it is already trained on the C# language, the target language for this framework experimented with within this paper.

Implementation for code classification has a class written called pipeline that will parse the source code; split the data into training, developing, and testing; construct the dictionary; and train a word embedding. Further block

sequences are generated for index representation, and data is processed for training purposes. Pipeline.py also uses the prepare_data.py for some of the functions. Further prepare_data.py uses class AST node from tree.py script. The class used here helps configure AST from the source code. In the train.py script, the author has configured a training approach. Model.py is referred to in the script train.py for running the model-related script. Train.py reads in train, test, and validation data runs the word2vec on the data and creates the embeddings. Parameters in Table 6 are configured for the model, BatchProgramClassifier.

The authors have reported an accuracy of 98.2% for the code classification task; this needs to be validated for the insurance company project data targeted for building the software security vulnerability modeling framework.

11. Conclusion

This paper has explored various building blocks that can be helpful build a comprehensive vulnerabilities management system for software development processes. The security of the software systems is the topmost priority for the software organization. Therefore, it is essential to leverage all the information generated across the industry and within the company. Various data science methods have been explored to build a knowledge management system that can assist the software development team to ensure a secure software system is being developed. Various approaches in this context have been explored using data of insurance domain-based software development. These approaches could facilitate an easy understanding of the practical challenges associated with actual-world implementation. The capabilities of language modeling and their role in the knowledge system were also discussed. The source code has been modeled to build a deep software security analysis model. The proposed model can help software engineers build secure software by assessing the software security during software development time. Extensive experiments have

been drawn by considering the various machine learning and deep learning models. It has been observed that the proposed deep learning-based models can efficiently explore the software language modeling capabilities to classify software systems' security vulnerabilities. Distill BERT and BERT have shown the good capability to model the insurance domain data to learn the security loopholes in the software development processes. Experiments also have demonstrated the capability of attention models in software security modeling.

Data Availability

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

Ethical Approval

This article does not contain any studies with human participants performed by any of the authors.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

The authors extend their appreciation to the researchers supporting project number RSP-2021/314, King Saud University, Riyadh, Saudi Arabia.

References

- [1] D. Barman and N. Chowdhury, "A novel semi supervised approach for text classification," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1147–1157, 2020.
- [2] V. Dabas, H. Parul Kumar, and A. Kumar, "Text classification algorithms for mining unstructured data: a swot analysis," *International Journal of Information Technology*, vol. 12, no. 4, pp. 1159–1169, 2020.
- [3] M. O. Al-Faruk, K. A. Hussain, M. A. Shahriar, and S. M. Tonni, "BFM: a forward backward string matching algorithm with improved shifting for information retrieval," *International Journal of Information Technology*, vol. 12, no. 2, pp. 479–483, 2020.
- [4] G. S. Lehal and H. Sintayehu, "Named entity recognition: a semi-supervised learning approach," *International Journal of Information Technology*, vol. 1, no. 7, 2020.
- [5] T. Li, "Identifying security requirements based on linguistic analysis and machine learning," in *Proceedings of the 2017 24th Asia-Pacific Software Engineering Conference (APSEC)*, pp. 388–397, IEEE, Nanjing, China, December 2017.
- [6] R. Malhotra, A. Chug, A. Hayrapetian, and R. Raje, "Analyzing and evaluating security features in software requirements," in *Proceedings of the 2016 International Conference on Innovation and Challenges in Cyber Security (ICICCS-INBUSH)*, pp. 26–30, IEEE, Greater Noida, India, February 2016.
- [7] G. H. de Rosa, M. Roder, D. F. Santos, and K. A. Costa, "Enhancing anomaly detection through restricted boltzmann machine features projection," *International Journal of Information Technology*, vol. 13, no. 1, pp. 49–57, 2021.
- [8] I. J. Saraf, "Generalized software fault detection and correction modeling framework through imperfect debugging, error generation and change point," *International Journal of Information Technology*, vol. 11, no. 4, pp. 751–757, 2019.
- [9] S. Bilgaiyan, S. Mishra, and S. Bilgaiyan, "Effort estimation in agile software development using experimental validation of neural network models," *International Journal of Information Technology*, vol. 11, no. 3, pp. 569–573, 2019.
- [10] R. Raje and A. Hayrapetian, "Empirically analyzing and evaluating security features in software requirements," in *Proceedings of the 11th Innovations in Software Engineering Conference*, pp. 1–11, Hyderabad, India, 2018.
- [11] R. R. Althar and D. Samanta, "The realist approach for evaluation of computational intelligence in software engineering," *Innovations in Systems and Software Engineering*, vol. 17, pp. 17–27, 2021.
- [12] S. Bettaieb, S. Y. Shin, M. Sabetzadeh, L. Briand, G. Nou, and M. Garceau, "Decision support for security-control identification using machine learning. in international working conference on requirements engineering: foundation for software quality," in *Proceedings of the 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 3–20, Springer, Orlando, FL, USA, 2019.
- [13] E.-K. Sherif and H. El-Hadary, "Capturing security requirements for software systems," *Journal of Advanced Research*, vol. 5, no. 4, pp. 463–472, 2014.
- [14] A. Ahmad, C. Feng, M. Khan et al., "A systematic literature review on using machine learning algorithms for software requirements identification on stack overflow," *Security and Communication Networks*, vol. 2020, Article ID 8830683, 19 pages, 2020.
- [15] S. Neuhaus, T. Zimmermann, C. Holler, and A. Zeller, "Predicting vulnerable software components," in *Proceedings of the 14th ACM Conference on Computer and Communications Security*, pp. 529–540, New York, NY, USA, 2007.
- [16] L. B. Othmane, G. Chehrizi, E. Bodden, P. Tsalovski, and A. D. Brucker, "Time for addressing software security issues: prediction models and impacting factors," *Data Science and Engineering*, vol. 2, no. 2, pp. 107–124, 2017.
- [17] G. Jain, M. Sharma, and B. Agarwal, "Optimizing semantic lstm for spam detection," *International Journal of Information Technology*, vol. 11, no. 2, pp. 239–250, 2019.
- [18] S. K. Sahu, P. Kumar, and A. P. Singh, "Modified k-nn algorithm for classification problems with improved accuracy," *International Journal of Information Technology*, vol. 10, no. 1, pp. 65–70, 2018.
- [19] R. Jindal, R. Malhotra, and A. Jain, "Automated classification of security requirements," in *Proceedings of the 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 2027–2033, Jaipur, India, 2016.
- [20] Wikipedia, *n-Gram*, <https://en.wikipedia.org/wiki/N-gram>, 2021.
- [21] Wikipedia, "Markov property," 2021, <https://en.wikipedia.org/wiki/N-gram>.
- [22] Y. Zhang, Y. Li, Y. Zhu, and X. Hu, "Wasserstein gan based on autoencoder with back-translation for cross-lingual embedding mappings," *Pattern Recognition Letters*, vol. 129, pp. 311–316, 2020.
- [23] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Text data augmentation for deep learning," *Journal of Big Data*, vol. 8, no. 1, pp. 1–34, 2021.

- [24] D. Sarkar, "dipanjans/deep_transfer_learning_nlp_dhs2019," 2021, https://github.com/dipanjanS/deep_transfer_learning_nlp_dhs2019.
- [25] D. Singh and V. Kumar, "Image dehazing using moore neighborhood-based gradient profile prior," *Signal Processing: Image Communication*, vol. 70, pp. 131–144, 2019.
- [26] G. Hu, S.-H. K. Chen, and N. Mazur, "Deep neural network-based speaker-aware information logging for augmentative and alternative communication," *Journal of Artificial Intelligence and Technology*, vol. 1, no. 2, pp. 138–143, 2021.
- [27] Fasttext, *English Word Vectors*, <https://fasttext.cc/docs/en/english-vectors.html>, 2021.
- [28] H. S. Basavegowda and G. Dagnew, "Deep learning approach for microarray cancer data classification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 22–33, 2020.
- [29] Y. Xu and T. T. Qiu, "Human activity recognition and embedded application based on convolutional neural network," *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 51–60, 2021.
- [30] B. Gupta, M. Tiwari, and S. S. Lamba, "Visibility improvement and mass segmentation of mammogram images using quantile separated histogram equalisation with local contrast enhancement," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 73–79, 2019.
- [31] D. Singh and V. Kumar, "A novel dehazing model for remote sensing images," *Computers & Electrical Engineering*, vol. 69, pp. 14–27, 2018.
- [32] D. Jiang, G. Hu, G. Qi, and N. Mazur, "A fully convolutional neural network-based regression approach for effective chemical composition analysis using near-infrared spectroscopy in cloud," *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 74–82, 2021.
- [33] S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 55–65, 2020.
- [34] D. Singh, V. Kumar, M. Kaur, M. Y. Jabarulla, and H.-N. Lee, "Screening of covid-19 suspected subjects using multi-crossover genetic algorithm based dense convolutional neural network," *IEEE Access*, vol. 9, pp. 142566–142580, 2021.
- [35] C. Raffel and D. P. Ellis, "Feed-forward networks with attention can solve some long-term memory problems," 2015, <https://arxiv.org/abs/1512.08756>.
- [36] T. Hub, *Universal-Sentence-Encoder*, <https://tfhub.dev/google/universal-sentence-encoder-large/3>, 2021.
- [37] P. I. S. Balsamo, A. Di Marco, and M. Simeoni, "Model-based performance prediction in software development: a survey," *IEEE Transactions on Software Engineering*, vol. 30, no. 5, pp. 295–310, 2004.
- [38] H. R. Shahriari and S. M. Ghaffarian, "Software vulnerability analysis and discovery using machine-learning and data-mining techniques: a survey," *ACM Computing Surveys*, vol. 50, no. 4, pp. 1–36, 2017.
- [39] M. Stavvas, "Technical debt as an indicator of software security risk: a machine learning approach for software development enterprises," *Enterprise Information Systems*, vol. 1, no. 43, 2020.
- [40] L. C. B. Shar, L. Khin, and H. B. K. Tan, "Web application vulnerability prediction using hybrid program analysis and machine learning," *IEEE Transactions on Dependable and Secure Computing*, vol. 12, no. 6, pp. 688–707, 2014.
- [41] T. Hub, "tf2-preview/nnlm-en-dim128," 2021, <https://tfhub.dev/google/tf2-preview/nnlm-en-dim128/1>.
- [42] H. Face, "The ai community building the future," 2021, <https://huggingface.co/>.
- [43] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," 2017, <https://arxiv.org/abs/1706.03762>.
- [44] E. Biswas, M. E. Karabulut, L. Pollock, and K. Vijay-Shanker, "Achieving reliable sentiment analysis in the software engineering domain using bert," in *Proceedings of the 2020 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pp. 162–173, IEEE, Adelaide, Australia, 2020.
- [45] C. Sun, M. Tang, L. Liang, and W. Zou, "Software entity recognition method based on bert embedding," in *Proceedings of the International Conference on Machine Learning for Cyber Security*, pp. 33–47, Springer, Cham, Germany, 2020.
- [46] W. Li, S. Gao, H. Zhou, Z. Huang, K. Zhang, and W. Li, "The automatic text classification method based on bert and feature union," in *Proceedings of the 2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 774–777, IEEE, Tianjin, China, 2019.
- [47] C. Mele and P. Ardimento, "Using bert to predict bug-fixing time," in *Proceedings of the 2020 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, pp. 1–7, IEEE, Bari, Italy, 2020.
- [48] T. Hub, *bert_en_uncased_l-12_h-768_a-12*, https://tfhub.dev/tensorflow/bert_en_uncased_L-12_H-768_A-12/1, 2021.
- [49] H. H. Wang, L. Yu, S. W. Tian, Y. F. Peng, and X. J. Pei, "Bidirectional lstm malicious webpages detection algorithm based on convolutional neural network and independent recurrent neural network," *Applied Intelligence*, vol. 49, no. 8, pp. 3016–3026, 2019.
- [50] S. Ke, C. Jingfei, G. Liu, Y. Lu, and X. Wei, "Convolutional neural networks-based locating relevant buggy code files for bug reports affected by data imbalance," *IEEE Access*, vol. 7, pp. 131304–131316, 2019.
- [51] J. Xu, J. Li, T. Zhang, Q. Du, and X. Li, "Software defect prediction and localization with attention-based models and ensemble learning," in *Proceedings of the 27th Asia-Pacific Software Engineering Conference (APSEC)*, pp. 81–90, Singapore, 2020.
- [52] Q. L. Nguyen, "Non-functional requirements analysis modeling for software product lines," in *Proceedings of the 2009 ICSE Workshop on Modeling in Software Engineering*, pp. 56–61, Vancouver, Canada, 2009.
- [53] S. Xu and Y. Xiong, "Automatic generation of pseudocode with attention seq2seq model," in *Proceedings of the 25th Asia-Pacific Software Engineering Conference (APSEC)*, pp. 711–712, Nara, Japan, 2018.
- [54] C. Ling, Z. Lin, Y. Zou, and B. Xie, "Adaptive deep code search," in *Proceedings of the 28th International Conference on Program Comprehension*, pp. 48–59, Seoul, South Korea, 2020.
- [55] J. Tabassum, M. Maddela, W. Xu, and A. Ritter, "Code and named entity recognition in stackoverflow," 2020, <https://arxiv.org/abs/2005.01634>.
- [56] X. Sun, X. Liu, J. Hu, B. Li, Y. Duan, and H. Yang, "Exploring topic models in software engineering data analysis: a survey," in *Proceedings of the 2016 17th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, pp. 357–362, IEEE/ACIS, Shanghai, China, 2016.
- [57] H. K. Dam, T. Tran, T. T. M. Pham, S. W. Ng, J. Grundy, and A. Ghose, "Automatic feature learning for predicting vulnerable software components," in *IEEE Transactions on Software Engineering*, vol. 47, no. 1, pp. 67–85, 2021.

Retraction

Retracted: Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpener of Remote Sensing Images

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] A. I. Iskanderani, I. M. Mehedi, A. J. Aljohani et al., "Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpener of Remote Sensing Images," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 7615106, 11 pages, 2021.

Research Article

Artificial Intelligence-Based Deep Fusion Model for Pan-Sharpening of Remote Sensing Images

Ahmed I. Iskanderani ¹, Ibrahim M. Mehedi ^{1,2}, Abdulah Jeza Aljohani ^{1,2},
Mohammad Shorfuzzaman ³, Farzana Akhter,⁴ Thangam Palaniswamy ¹,
Shaikh Abdul Latif ⁵, Abdul Latif ⁶, and Rahtul Jannat ⁷

¹Department of Electrical and Computer Engineering (ECE), King Abdulaziz University, Jeddah, Saudi Arabia

²Center of Excellence in Intelligence Engineering Systems (CEIES), King Abdulaziz University, Jeddah, Saudi Arabia

³Department of Computer Science, CIT Taif University, Taif, Saudi Arabia

⁴Aarhus BSS, Aarhus University, Aarhus, Denmark

⁵Department of Nuclear Engineering, King Abdulaziz University, Jeddah, Saudi Arabia

⁶Department of Mathematics, King Abdulaziz University, Jeddah, Saudi Arabia

⁷Department of Electrical and Electronic Engineering (EEE), BRAC University, Dhaka, Bangladesh

Correspondence should be addressed to Rahtul Jannat; mrrahatuljannat@gmail.com

Received 7 November 2021; Revised 6 December 2021; Accepted 10 December 2021; Published 23 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Ahmed I. Iskanderani 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.

During the past two decades, many remote sensing image fusion techniques have been designed to improve the spatial resolution of the low-spatial-resolution multispectral bands. The main objective is fuse the low-resolution multispectral (MS) image and the high-spatial-resolution panchromatic (PAN) image to obtain a fused image having high spatial and spectral information. Recently, many artificial intelligence-based deep learning models have been designed to fuse the remote sensing images. But these models do not consider the inherent image distribution difference between MS and PAN images. Therefore, the obtained fused images may suffer from gradient and color distortion problems. To overcome these problems, in this paper, an efficient artificial intelligence-based deep transfer learning model is proposed. Inception-ResNet-v2 model is improved by using a color-aware perceptual loss (CPL). The obtained fused images are further improved by using gradient channel prior as a postprocessing step. Gradient channel prior is used to preserve the color and gradient information. Extensive experiments are carried out by considering the benchmark datasets. Performance analysis shows that the proposed model can efficiently preserve color and gradient information in the fused remote sensing images than the existing models.

1. Introduction

Fusion of multispectral (MS) and panchromatic (PAN) images has attracted researchers' interest, since it results in a fused image with better spatial resolution and spectral information [1]. The spatial resolution of a MS image is significantly better as compared to a PAN image. But a MS image only has a single band. Thus, to obtain an image with significant spectral information and better spatial resolution, efficient pan-sharpening approaches are required [2].

Many pan-sharpening techniques have been implemented so far. The traditional methods suffer from blurring effect and color distortion [1, 3]. Sparse representation theory-based fusion methods can easily overcome the problems of color distortion by enhancing the spatial resolution of MS images [4]. The intensity-hue-saturation (IHS) method was also used to fuse the images. These models were quite simple and efficient and can produce high-spatial-quality images [5, 6]. However, they experience spectral distortion. The spectral fidelity can be enforced using an edge-adaptive IHS method [7]. Compressed sensing (CS)

theory is also used for pan-sharpening of multispectral images. It can recover the sparse signal from a small number of linear measurements [8]. Optimized pan-sharpening techniques were also developed to preserve the spectral and geometry constraints [9, 10]. The Bayesian theory-based fusion model solved the problems of linear model and attained superior spatial and spectral fusion [11].

Recently, various deep learning models have been used to implement pan-sharpening techniques to produce HR PAN images. These techniques can effectively model complex relationships between variables via the composition of several levels of non-linearity [12]. In the deep pan-sharpening model, the correlation between the LR/HR MS image patches is the same as the LR/HR PAN image patches. Thereafter, this assumption is used to learn the mapping using convolutional neural network (CNN) [13]. Different types of CNN were used to fuse the images. CNN contains three convolutional layers such input, hidden, and output layers. Activation functions are contained in each layer. Input and hidden layers contain non-linear activation layer while output layer comprises linear activation function. For every layer, there are I input bands, J output bands, filters, parameters needed to be learned, tensors, weights, and biases. In case of fusion, PAN bands are given as input to the CNN. The components of MS are upsampled and then radiometric indices are extracted. Lastly, non-linear combinations of MS bands are made to improve the performance [14]. However, most of these methods suffer from inadequate spatial texture improvement and spectral distortion issues. To overcome these issues, many techniques were developed. A dual-path fusion network (DPFN) enhanced spatial texture and spectral distortion [15]. Shallow-deep convolutional network (SDCN) can produce fused images with minimal spectral distortion [16]. Dynamic deep learning models were proposed to build sensitive models towards input images [15]. Coupled multiscale convolutional neural network considered the PAN and MS images at different resolutions for better feature extraction [17]. A four-layer CNN and a loss function were designed which can extract spatial and spectral characteristics efficiently from original image. It did not require any reference fused image and hence did not need simulation data for training [18]. Generative adversarial learning (GAN) was also utilized to implement the fusion of PAN and MS images. It has an ability to produce high-fidelity fused images [19].

From the existing literature, it has been found that the deep learning and deep transfer learning models can efficiently fuse the remote sensing images. However, these models do not consider the inherent image distribution difference between MS and PAN images. Therefore, the obtained fused images may suffer from gradient and color distortion problems. To overcome these problems, in this paper, an efficient deep transfer learning model is proposed. The main contributions of this paper are as follows:

- (1) An efficient Inception-ResNet-v2 model is improved by using a CPL.

- (2) The obtained fused images are further improved by using gradient channel prior as a postprocessing step.
- (3) Extensive experiments are carried out by considering the benchmark datasets.

This paper is organized as follows. Section 2 discusses the literature review. Section 3 presents the proposed model. Comparative analysis is discussed in Section 4. Section 5 concludes the paper.

2. Literature Review

Wang et al. [20] proposed a pan-sharpening technique based on the channel-spatial attention model (CSA). In this, residual attention module was designed to produce high-resolution images. Xu et al. [21] implemented the soil prediction model using the pan-sharpened remote sensing indices. In this, images of Landsat 8, GeoEye-1, and WorldView-2 were fused. A prediction model was designed using random forest. Ma et al. [22] used the generative adversarial network to implement the pan-sharpening technique. For network training, it did not require ground truth. Akula et al. [23] implemented a pan-sharpening technique using adaptive principal component analysis and local variation contourlet transform. Wang et al. [24] utilized area-to-point regression kriging (ATPRK) for pan-sharpening. Wang et al. [25] presented a pan-sharpening technique based on compressed sensing. The joint sparsity model was used to recover the high-resolution multispectral images.

Wu et al. [26] utilized multiobjective decision for improving the fused multiband images. The information injection model was used to improve texture and gradient details of MS image. Spectral fidelity fusion was designed using injected information and spectral modulation to fused image. Zhuang et al. [27] designed a probabilistic model to fuse MS and PAN images. Gradient domain-guided image filtering was also used to refine the results. The maximum a posteriori model was also implemented on the difference between PAN and MS images and in respective gradient domains too. Sibiyi et al. [28] combined image texture obtained from a fused image with partial least squares discriminant analysis to monitor and map commercial forest species. This model proved that the image texture can discriminate commercial forest species.

Fang et al. [29] designed a framelet-based fusion model by using a variational model. The split Bregman iteration was also used to obtain better results. The Bregman method solves the convex optimization problems using regularization. It is best suited for those optimization problems where constraints are well specified. Due to error cancellation effect, it converges very fast. Wang et al. [30] proposed sparse tensor neighbor embedding for fusion of PAN and MS images using N-way block pursuit. A sparse tensor was concatenated with neighbor embedding to obtain a new high-dimensional sparse tensor embedding for fusion of PAN and MS images in an efficient manner. Saedi and Faez [31] utilized shiftable contourlet transform and

multiobjective particle swarm optimization (MPSO) to fuse PAN and MS images. PAN and MS images were histogram matched prior to fusion process. Fang et al. [32] designed a pan-sharpening technique using variational approach. In this, three assumptions were made to construct the energy function. The minimized solution was obtained using the Bregman algorithm. Zhang et al. [33] implemented a variational energy function to preserve the spectrum, geometry, and correlation information of the original images while pan-sharpening.

Ye et al. [34] proposed gradient-based deep network prior to fuse PAN and MS images. Convolutional neural network (CNN) was trained in gradient domain using the problem-specific recursive block. Xing et al. [35] implemented the pan-sharpening technique using deep metric learning (DML). The deep metric learning was used to train refined geometric multimanifold neighbor embedding. The hierarchical characteristics of masks were used by considering various non-linear deep learning models. Gogineni and Chaturvedi [36] used multiscale learned dictionary (MSLD) to design a pan-sharpening technique. It could obtain the underlying features of images, in which the characteristics of both learned dictionaries and multiscale were possessed. Huang et al. [37] developed a fusion model using multiple deep learning models (MDLMs). The non-subsampled contourlet transform (NSCT) used to decompose the PAN images into frequency bands. The characteristics of high-frequency bands were learned by the deep learning model.

From the literature, it is found that the deep learning model should be improved by using a better loss function and some preprocessing techniques [38–40].

3. Proposed Model

An efficient artificial intelligence-based deep transfer learning model is proposed. Inception-ResNet-v2 model is improved by using a CPL. The obtained fused images are further improved by using gradient channel prior as a postprocessing step.

3.1. Inception-ResNet-v2. Inception-ResNet-v2 is a well-known model which is improved InceptionNet with residual connections. It is achieved by replacing the filter concatenation stage of the InceptionNet (see [41]). Figure 1 shows the architecture of Inception-ResNet-v2.

3.2. Color-Aware Perceptual Loss. CPL [42] is utilized to assign small coefficients to feature channels which are more sensitive to colors for every L Inception-ResNet-v2 layer during the computation of perceptual loss.

The difference among the respective color and a gray-scale-inverted MS image (M_δ^{-1}) is utilized to compute coefficients of features. Higher difference indicates that the features are more sensitive to colors. CPL is also more sensitive to gradient information. The average feature difference is then assigned to the exponential function with a variable γ (see [42]). It can be represented as

$$M_\delta^{-1} = 1 - \frac{(M_\gamma^R + M_\gamma^G + M_\gamma^B)}{3}, \quad (1)$$

where M_γ^R , M_γ^G , and M_γ^B represent color channels of MS image. The CPL coefficients for color channels of every layer l , \mathbf{W}_{cpl}^l , can be computed as

$$\mathbf{W}_{cpl}^l = e^{-\gamma(1/N) \sum_n |\phi^l(M_\delta^{-1}) - \phi^l(M_\gamma)|}, \quad (2)$$

where γ is used to neglect features which are sensitive to colors. For PAN image (M_{pc}) and a CNN-based fused image (M_{ps}), CPL can be computed as

$$l_{cpl} = \sum_{l=1}^L \left\| \mathbf{W}_{cpl}^l \odot (\phi_{m_l}^l(M_{pc}) - \phi_{m_l}^l(M_{ps})) \right\|_1, \quad (3)$$

where $m_l = [7, 5, 3]$ shows the size of max-pool implemented on l^{th} layer feature which is used to achieve average of shift invariance to CPL. It can efficiently manage the misalignment problem.

Although CPL assigns high-frequency information to M_{ps} , additional loss is also required for color fidelity. Therefore, the perceptual and l_1 losses are used. The fidelity loss can be defined as

$$l_f = \alpha_{cpl} l_{cpl}(M_{pc}, M_{ps}) + \alpha_\gamma l_p(M_\gamma, M_{ps\downarrow}) + \alpha_{l_1} l_1(M_\gamma, M_{ps\downarrow}), \quad (4)$$

where $M_{ps\downarrow}$ shows a downscaled PS image to the MS resolution, l_1 loss is an average absolute difference, i.e., $1/N \sum_n |M_\gamma - M_{ps\downarrow}|$, and α_{cpl} , α_γ , and α_{l_1} are set to 0.85, 0.02, and 0.95, respectively.

3.3. Gradient Channel Prior. GCP is utilized to restore any kind of degradation from the images. It has an ability to preserve the gradient and texture information of restored images [43]. GCP can be defined as

$$\nabla I(m, n) = \begin{pmatrix} \psi_m \\ \psi_n \end{pmatrix} = \begin{pmatrix} \partial I / \partial m \\ \partial I / \partial n \end{pmatrix}. \quad (5)$$

An amplitude of I can be computed as

$$\text{mag}(I) = \sqrt{\psi_m^2 + \psi_n^2}. \quad (6)$$

An orientation angle of ∇I can be computed as

$$\nabla_O(m, n) = \arctan\left(\frac{\psi_n}{\psi_m}\right). \quad (7)$$

For $I(m, n)$, ψ_m and ψ_n can be computed using various masks (see [44]).

4. Performance Analysis

The proposed model is trained on using 100 epochs with a mini-batch size of 10 using Adam optimizer [45]. The learning rate is used as 5×10^{-5} . All experiments are performed on MATLAB2021a software. Experiments are performed on Pleiades, QuickBird, IKONOS, and WorldView-2

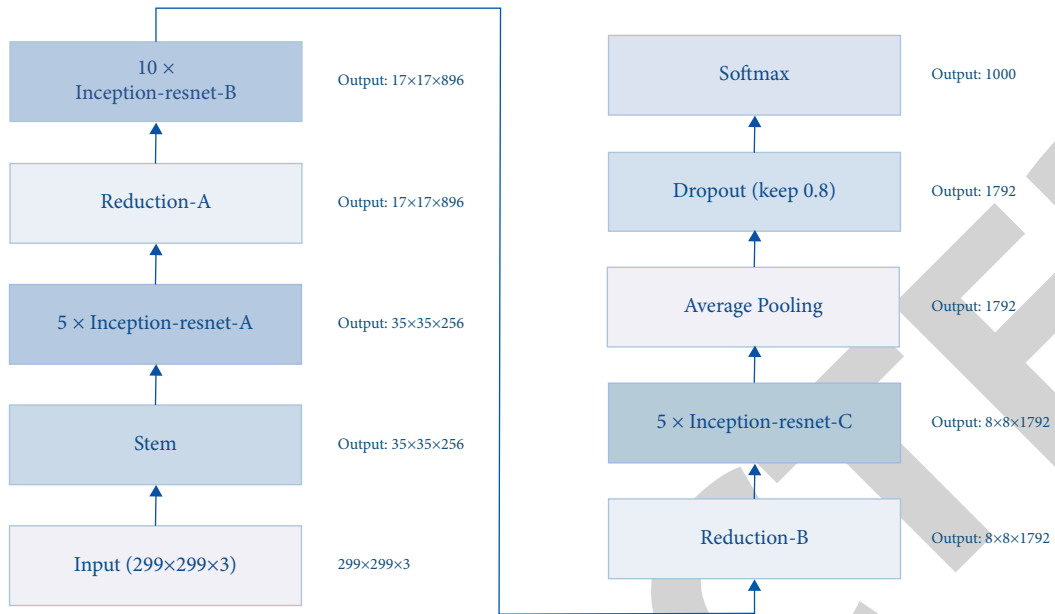


FIGURE 1: Architecture of Inception-ResNet-v2.



FIGURE 2: Continued.

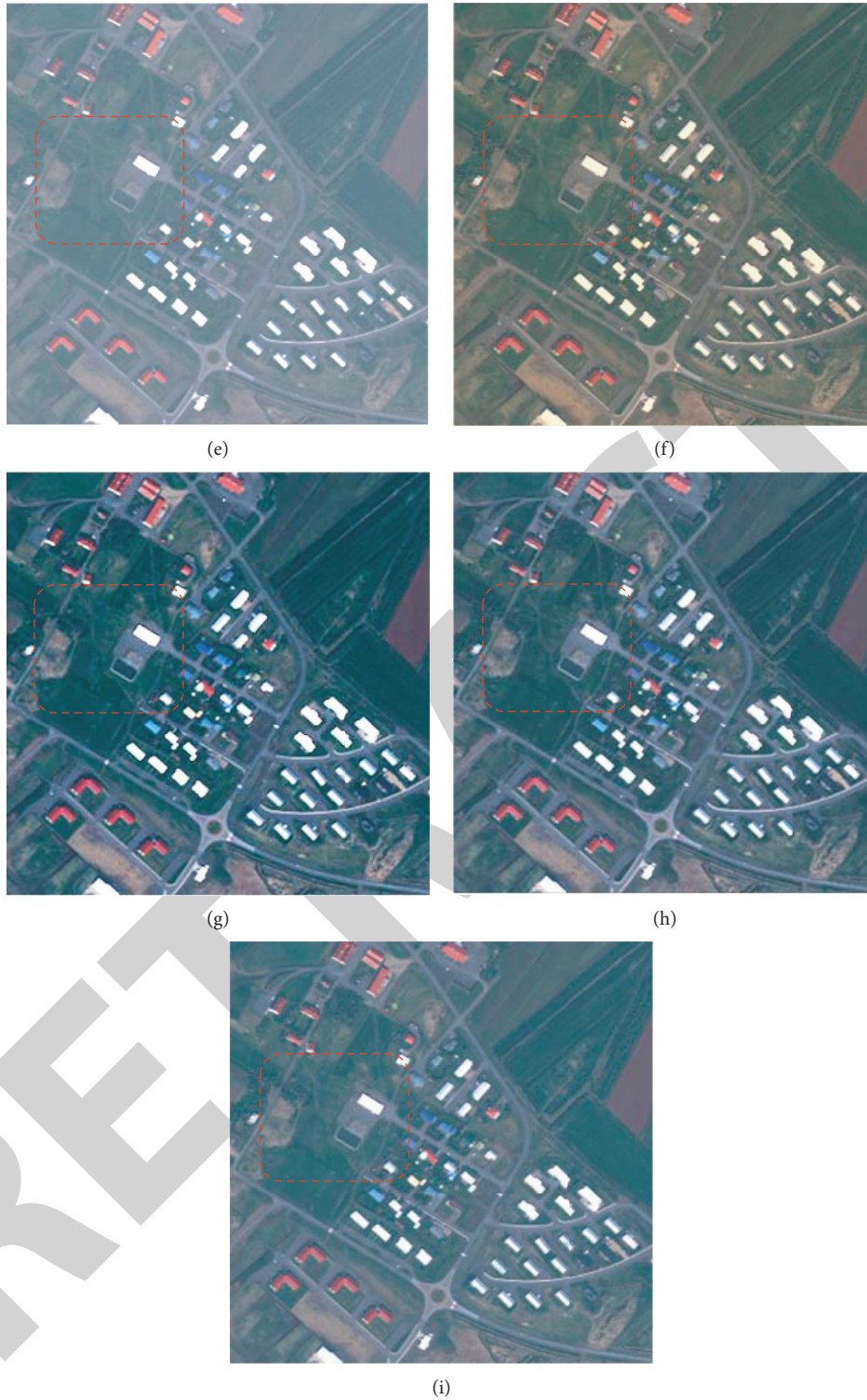


FIGURE 2: Visual analysis of pan-sharpening techniques. (a) Low-resolution multispectral (MS) image. (b) High-spatial-resolution pan-chromatic (PAN) image.

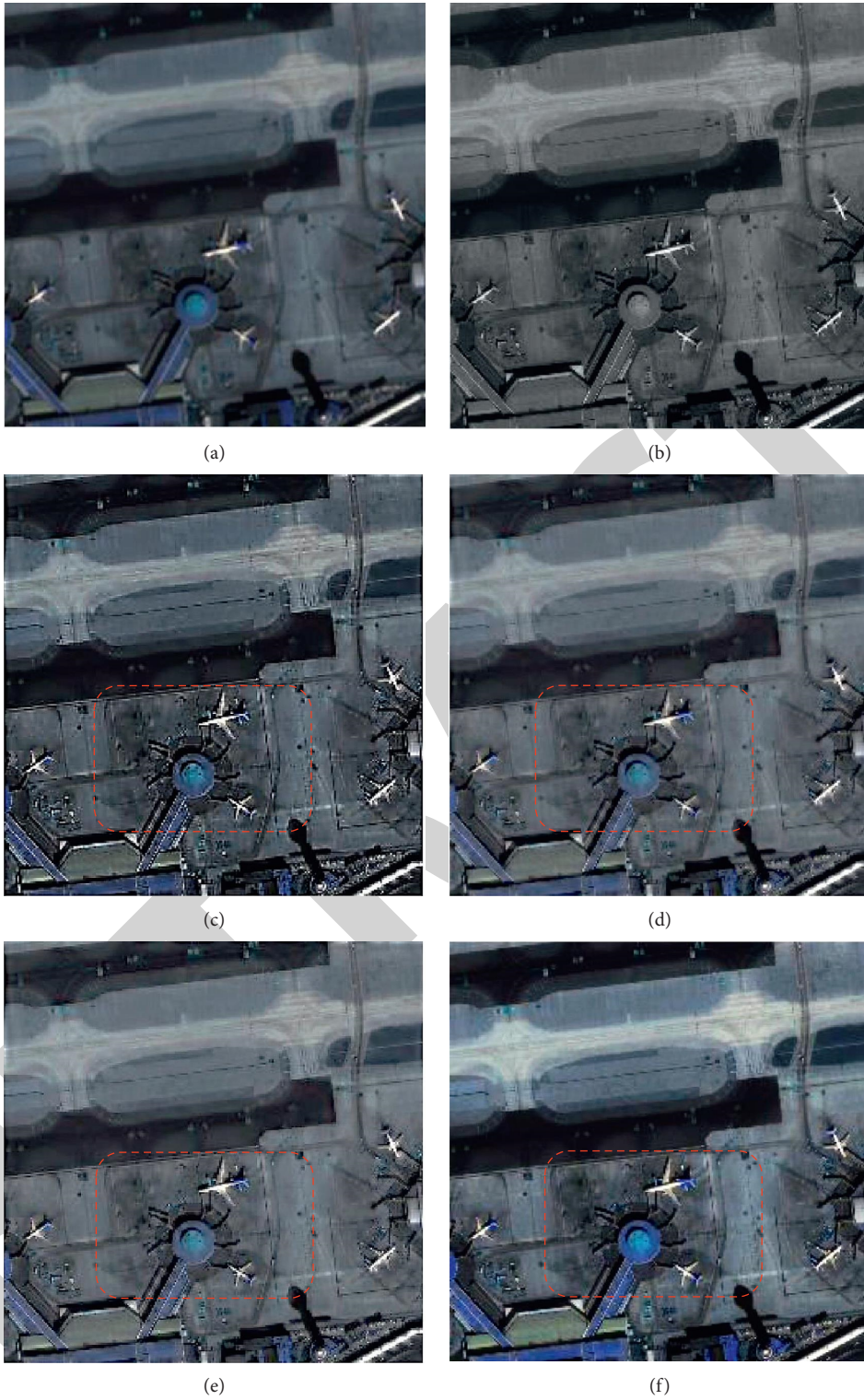


FIGURE 3: Continued.

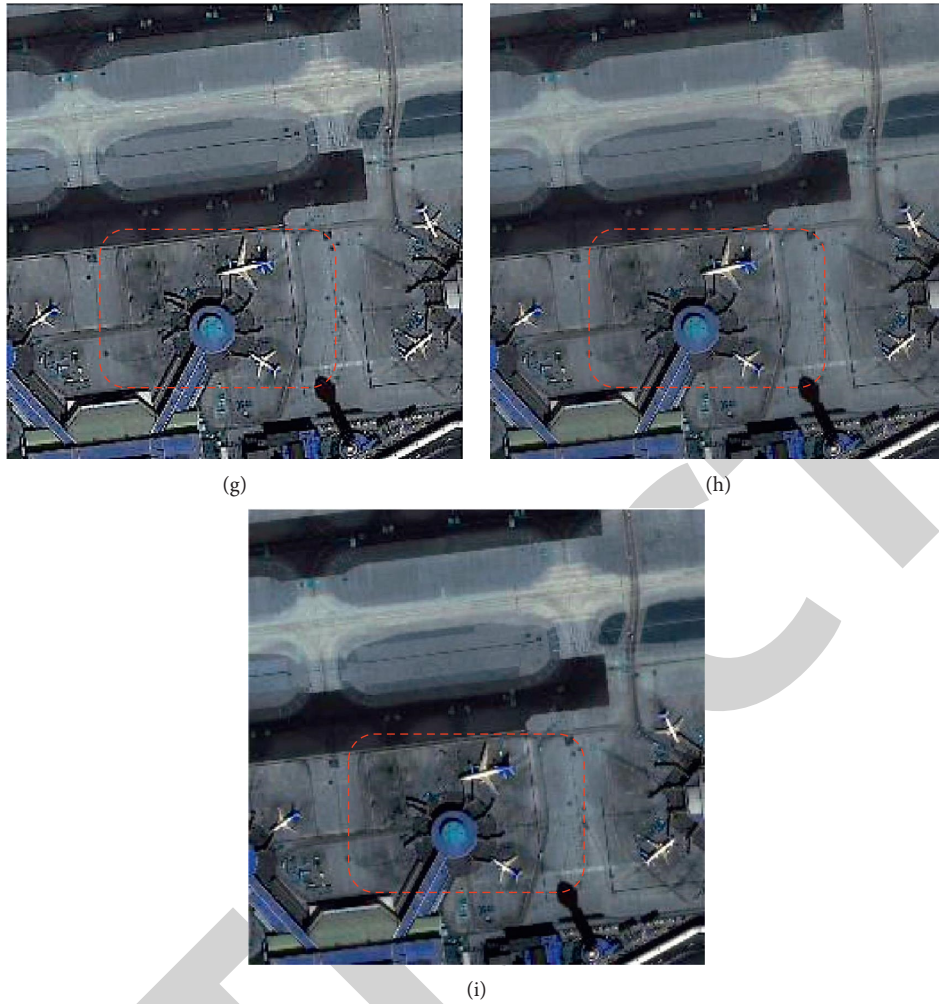


FIGURE 3: Visual analysis of pan-sharpening techniques. (a) Low-resolution multispectral (MS) image. (b) High-spatial-resolution pan-chromatic (PAN) image.

images. Comparisons are performed with six well-known competitive techniques.

4.1. Visual Analysis. Figures 2 and 3 show the visual analysis of the proposed model. It is found that the proposed model has better visibility as compared to the existing techniques. Red rectangles show the specific region in the obtained fused images. The selected region reflects the spatial and spectral information along with any kind of artifacts which are present in the obtained fused images. Also, the proposed model shows better gradient and color preservation as compared to the existing techniques. The results obtained from the proposed model show better spatial and spectral information. It clearly shows that the existing models are able to fuse the images by improving the spatial and spectral information of fused images. But whenever there is redundant information in both PAN and MS images, then the existing method fails to fuse the content efficiently. Also, the texture and gradient preservation is significantly more in the fused image obtained using the proposed model.

4.2. Quantitative Analysis. Five well-known quality metrics, i.e., root mean square error (RMSE) [46], universal image quality index (UIQI) [47], correlation coefficient (CC) [46], spectral angle mapper (SAM) [46], and Erreur relative globale adimensionnelle de synthese (ERGAS) [48], are used for comparative analysis.

Table 1 shows CC analysis of the proposed deep pan-sharpening model. CC is desirable to be maximum. It is found that the proposed model outperforms the competitive pan-sharpening models by 1.7824%.

Table 2 depicts UIQI analysis of the proposed deep pan-sharpening model. UIQI is desirable to be maximum. It is found that the proposed model outperforms the competitive pan-sharpening models by 1.2498%.

Table 3 shows SAM analysis of the proposed deep pan-sharpening model. SAM is desirable to be minimum. It is found that the proposed model outperforms the competitive pan-sharpening models by showing an average reduction of 1.3457%.

The quality of pan-sharpened images can be assessed using ERGAS. It determines the transition between spectral

TABLE 1: Analysis of correlation coefficient (maximum is desirable).

Images	CSA	CNN	MSLD	MDLM	DML	MPSO	Proposed
Pleiades 1	0.9482	0.9585	0.9525	0.9356	0.9495	0.9531	0.9612
Pleiades 2	0.9384	0.9545	0.9588	0.9636	0.9595	0.9594	0.9668
QuickBird 1	0.9543	0.9525	0.9517	0.9594	0.9353	0.9406	0.9626
QuickBird 2	0.9388	0.9387	0.9627	0.9435	0.9451	0.9388	0.9659
QuickBird 3	0.9422	0.9451	0.9386	0.9457	0.9481	0.9416	0.9513
IKONOS 1	0.9601	0.9493	0.9589	0.9366	0.9547	0.9645	0.9677
IKONOS 2	0.9391	0.9617	0.9442	0.9452	0.9405	0.9578	0.9649
IKONOS 3	0.9422	0.9562	0.9385	0.9563	0.9602	0.9529	0.9634
WorldView-2 1	0.9399	0.9572	0.9609	0.9533	0.9401	0.9549	0.9641
WorldView-2 2	0.9446	0.9418	0.9393	0.9386	0.9528	0.9537	0.9569

TABLE 2: Analysis of UIQI (maximum is desirable).

Images	CSA	CNN	MSLD	MDLM	DML	MPSO	Proposed
Pleiades 1	0.8227	0.8253	0.8252	0.8256	0.8207	0.8163	0.8288
Pleiades 2	0.8286	0.8222	0.8168	0.8277	0.8333	0.8303	0.8365
QuickBird 1	0.8276	0.8292	0.8161	0.8247	0.8219	0.8238	0.8324
QuickBird 2	0.8324	0.8235	0.8182	0.8235	0.8188	0.8276	0.8356
QuickBird 3	0.8176	0.8302	0.8216	0.8310	0.8322	0.8157	0.8354
IKONOS 1	0.8253	0.8205	0.8152	0.8266	0.8155	0.8253	0.8298
IKONOS 2	0.8324	0.8335	0.8325	0.8337	0.8305	0.8254	0.8369
IKONOS 3	0.8209	0.8338	0.8292	0.8154	0.8249	0.8237	0.8375
WorldView-2 1	0.8236	0.8347	0.8165	0.8294	0.8304	0.8253	0.8379
WorldView-2 2	0.8185	0.8183	0.8275	0.8236	0.8216	0.8323	0.8355

TABLE 3: Analysis of SAM (minimum is desirable).

Images	CSA	CNN	MSLD	MDLM	DML	MPSO	Proposed
Pleiades 1	6.1233	5.2497	4.8504	5.2314	5.6140	5.7876	4.8472
Pleiades 2	5.0322	5.0038	5.2320	5.6089	5.0142	5.027	5.0006
QuickBird 1	5.5875	5.7493	5.6528	5.5909	5.5925	5.1086	5.1054
QuickBird 2	5.3167	4.9801	5.2275	5.4522	5.0770	4.8495	4.8463
QuickBird 3	4.8664	5.4697	5.5565	5.4524	5.6393	4.9815	4.8632
IKONOS 1	5.6374	4.9882	4.8493	5.3591	4.9381	5.4698	4.8461
IKONOS 2	5.5923	5.4901	5.5478	4.8991	5.7502	5.7276	4.8959
IKONOS 3	5.1743	4.8863	5.6782	5.0255	5.1709	5.0138	4.8831
WorldView-2 1	5.4105	5.1662	5.6827	5.6985	5.8012	5.3316	5.1635
WorldView-2 2	5.0545	4.8830	5.3821	5.3335	5.0876	5.5423	4.8798

TABLE 4: Analysis of ERGAS (minimum is desirable).

Images	CSA	CNN	MSLD	MDLM	DML	MPSO	Proposed
Pleiades 1	6.4651	5.7022	6.0366	5.5356	7.0193	6.5008	3.0324
Pleiades 2	7.0733	6.7986	5.5258	5.8420	7.5852	7.1921	3.7944
QuickBird 1	5.1640	6.2983	7.1174	5.6054	7.0157	5.8708	3.2941
QuickBird 2	6.2149	6.3734	7.7259	7.4269	6.0897	6.6930	3.0855
QuickBird 3	5.3291	6.4481	6.4994	7.0956	6.7466	6.5206	3.4439
IKONOS 1	5.0443	7.5506	5.3478	5.6993	7.6662	7.9703	2.9661
IKONOS 2	5.4928	7.2282	6.8215	6.2475	5.4354	5.0493	3.2433
IKONOS 3	7.3333	5.2985	5.9716	7.3313	6.4497	6.4498	3.4455
WorldView-2 1	7.3629	5.9718	7.6287	6.0475	7.1961	5.4670	4.0428
WorldView-2 2	5.9306	6.9576	6.2451	6.6974	7.1292	5.8911	3.2409

TABLE 5: Analysis of RMSE (minimum is desirable).

Images	CSA	CNN	MSLD	MDLM	DML	MPSO	Proposed
Pleiades 1	17.3301	17.5812	17.0435	18.6407	17.5075	18.8094	16.9898
Pleiades 2	17.4840	17.4399	17.883	18.6324	17.2132	17.7383	17.1600
QuickBird 1	16.9437	17.2044	17.3443	18.1278	18.4979	18.7183	16.8905
QuickBird 2	18.7044	17.9968	17.2075	18.698	18.3412	18.3956	17.1543
QuickBird 3	18.2260	18.7599	17.5858	16.9439	16.9921	18.1098	16.8907
IKONOS 1	18.7411	17.6879	17.9462	17.4335	17.9966	17.1842	17.1314
IKONOS 2	17.9322	18.275	18.2602	18.1607	17.2863	16.9537	16.9005
IKONOS 3	17.0796	17.8841	18.5387	17.9606	17.7461	17.8408	17.0264
WorldView-2 1	17.9519	17.5074	17.3929	18.5105	18.4653	17.1286	17.0754
WorldView-2 2	17.0827	18.5631	17.5701	18.6942	17.1731	18.1685	17.0295

and spatial information [49]. Table 4 demonstrates ERGAS analysis of the proposed deep pan-sharpening model. ERGAS is desirable to be minimum. It is found that the proposed model outperforms the competitive pan-sharpening models by showing an average reduction of 1.0985%.

Table 5 demonstrates RMSE analysis of the proposed deep pan-sharpening model. RMSE is desirable to be minimum. It is found that the proposed model outperforms the competitive pan-sharpening models by showing an average reduction of 1.5486%.

5. Conclusion

To obtain a remote sensing image with better spatial and spectral information, efficient image fusion techniques are desirable. However, it has been found that the existing models do not consider the inherent image distribution difference between MS and PAN images. Therefore, the obtained fused images suffer from gradient and color distortion problems. To overcome these problems, in this paper, an efficient deep transfer learning model has been proposed. Inception-ResNet-v2 model was improved by using a color-aware perceptual loss (CPL). The obtained fused images were further improved by using gradient channel prior as a postprocessing step. Gradient channel prior was utilized to preserve the color and gradient information. Extensive experiments were carried out by considering the benchmark datasets. Performance analysis has shown that the proposed model can efficiently preserve color and gradient information in the fused remote sensing images than the existing models. The proposed model outperformed the competitive pan-sharpening models in terms of CC and UIQI by 1.7824% and 1.2498%, respectively. Also, compared to the existing models, the proposed model has achieved an average reduction in SAM, ERGAS, and RMSE by 1.3457%, 1.2847%, and 1.5486%, respectively.

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

The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia, for funding this research work through the project number (IFPRC-027-135-2020) and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

References

- [1] S. Ayas, E. T. Gormus, and M. Ekinici, "An efficient pan sharpening via texture based dictionary learning and sparse representation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 7, pp. 2448–2460, 2018.
- [2] C. Shi, F. Liu, L. Li, L. Jiao, Y. Duan, and S. Wang, "Learning interpolation via regional map for pan-sharpening," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3417–3431, 2015.
- [3] H. S. Basavegowda and G. Dagnev, "Deep learning approach for microarray cancer data classification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 22–33, 2020.
- [4] C. Jiang, H. Zhang, H. Shen, and L. Zhang, "Two-step sparse coding for the pan-sharpening of remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 5, pp. 1792–1805, 2014.
- [5] M. Ghahremani and H. Ghassemian, "Nonlinear IHS: a promising method for pan-sharpening," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 11, pp. 1606–1610, 2016.
- [6] S. Ghosh, P. Shivakumara, P. Roy, U. Pal, and T. Lu, "Graphology based handwritten character analysis for human behaviour identification," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 1, pp. 55–65, 2020.
- [7] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, and T. Wittman, "An adaptive IHS pan-sharpening method," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 4, pp. 746–750, 2010.
- [8] M. Ghahremani and H. Ghassemian, "A compressed-sensing-based pan-sharpening method for spectral distortion reduction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2194–2206, 2016.
- [9] P. Liu, L. Xiao, and T. Li, "A variational pan-sharpening method based on spatial fractional-order geometry and spectral-spatial low-rank priors," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1788–1802, 2018.
- [10] B. Gupta, M. Tiwari, and S. Lamba, "Visibility improvement and mass segmentation of mammogram images using quantile separated histogram equalisation with local contrast

- enhancement,” *CAAI Transactions on Intelligence Technology*, vol. 4, no. 2, pp. 73–79, 2019.
- [11] P. Guo, P. Zhuang, and Y. Guo, “Bayesian pan-sharpening with multiorder gradient-based deep network constraints,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 950–962, 2020.
- [12] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, “A new pan-sharpening method with deep neural networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 5, pp. 1037–1041, 2015.
- [13] J. Yang, X. Fu, Y. Hu, Y. Huang, X. Ding, and J. Paisley, “Pannet: a deep network architecture for pan-sharpening,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 5449–5457, Venice, Italy, October 2017.
- [14] G. Scarpa, S. Vitale, and D. Cozzolino, “Target-adaptive CNN-based pansharpening,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 9, pp. 5443–5457, 2018.
- [15] J. Wang, Z. Shao, X. Huang, T. Lu, and R. Zhang, “A dual-path fusion network for pan-sharpening,” *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–14, 2021.
- [16] L. Liu, J. Wang, E. Zhang et al., “Shallow-deep convolutional network and spectral-discrimination-based detail injection for multispectral imagery pan-sharpening,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 1772–1783, 2020.
- [17] J. Wei, Y. Xu, W. Cai, Z. Wu, J. Chanussot, and Z. Wei, “A two-stream multiscale deep learning architecture for pan-sharpening,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5455–5465, 2020.
- [18] Z. Xiong, Q. Guo, M. Liu, and A. Li, “Pan-sharpening based on convolutional neural network by using the loss function with no-reference,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 897–906, 2021.
- [19] Q. Liu, H. Zhou, Q. Xu, X. Liu, and Y. Wang, “PSGAN: a generative adversarial network for remote sensing image pan-sharpening,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 12, pp. 1–16, 2020.
- [20] P. Wang and E. Sertel, “Channel-spatial attention-based pansharpening of very high-resolution satellite images,” *Knowledge-Based Systems*, vol. 229, Article ID 107324, 2021.
- [21] Y. Xu, S. E. Smith, S. Grunwald, A. Abd-Elrahman, and S. P. Wani, “Incorporation of satellite remote sensing pan-sharpened imagery into digital soil prediction and mapping models to characterize soil property variability in small agricultural fields,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 123, pp. 1–19, 2017.
- [22] J. Ma, W. Yu, C. Chen, P. Liang, X. Guo, and J. Jiang, “PanGAN: an unsupervised pan-sharpening method for remote sensing image fusion,” *Information Fusion*, vol. 62, pp. 110–120, 2020.
- [23] R. Akula, R. Gupta, and M. R. V. Devi, “An efficient pan sharpening technique by merging two hybrid approaches,” *Procedia Engineering*, vol. 30, pp. 535–541, 2012.
- [24] Q. Wang, W. Shi, and P. M. Atkinson, “Area-to-point regression kriging for pan-sharpening,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 114, pp. 151–165, 2016.
- [25] W. Wang, L. Jiao, S. Yang, and K. Rong, “Distributed compressed sensing-based pan-sharpening with hybrid dictionary,” *Neurocomputing*, vol. 155, pp. 320–333, 2015.
- [26] L. Wu, Y. Yin, X. Jiang, and T. C. E. Cheng, “Pan-sharpening based on multi-objective decision for multi-band remote sensing images,” *Pattern Recognition*, vol. 118, Article ID 108022, 2021.
- [27] P. Zhuang, Q. Liu, and X. Ding, “Pan-GGF: a probabilistic method for pan-sharpening with gradient domain guided image filtering,” *Signal Processing*, vol. 156, pp. 177–190, 2019.
- [28] B. Sibiya, R. Lottering, and J. Odindi, “Discriminating commercial forest species using image texture computed from a worldview-2 pan-sharpened image and partial least squares discriminant analysis,” *Remote Sensing Applications: Society and Environment*, vol. 23, Article ID 100605, 2021.
- [29] F. Fang, G. Zhang, F. Li, and C. Shen, “Framelet based pan-sharpening via a variational method,” *Neurocomputing*, vol. 129, pp. 362–377, 2014.
- [30] M. Wang, K. Zhang, X. Pan, and S. Yang, “Sparse tensor neighbor embedding based pan-sharpening via n-way block pursuit,” *Knowledge-Based Systems*, vol. 149, pp. 18–33, 2018.
- [31] J. Saeedi and K. Faez, “A new pan-sharpening method using multiobjective particle swarm optimization and the shiftable contourlet transform,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 365–381, 2011.
- [32] F. Fang, F. Li, C. Shen, and G. Zhang, “A variational approach for pan-sharpening,” *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2822–2834, 2013.
- [33] G. Zhang, F. Fang, A. Zhou, and F. Li, “Pan-sharpening of multi-spectral images using a new variational model,” *International Journal of Remote Sensing*, vol. 36, no. 5, pp. 1484–1508, 2015.
- [34] F. Ye, Y. Guo, and P. Zhuang, “Pan-sharpening via a gradient-based deep network prior,” *Signal Processing: Image Communication*, vol. 74, pp. 322–331, 2019.
- [35] Y. Xing, M. Wang, S. Yang, and L. Jiao, “Pan-sharpening via deep metric learning,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 145, pp. 165–183, 2018.
- [36] R. Gogineni and A. Chaturvedi, “Sparsity inspired pansharpening technique using multi-scale learned dictionary,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 360–372, 2018.
- [37] W. Huang, X. Fei, J. Feng, H. Wang, Y. Liu, and Y. Huang, “Pan-sharpening via multi-scale and multiple deep neural networks,” *Signal Processing: Image Communication*, vol. 85, Article ID 115850, 2020.
- [38] D. Jiang, G. Hu, G. Qi, and N. Mazur, “A fully convolutional neural network-based regression approach for effective chemical composition analysis using near-infrared spectroscopy in cloud,” *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 74–82, 2021.
- [39] Y. Xu and T. T. Qiu, “Human activity recognition and embedded application based on convolutional neural network,” *Journal of Artificial Intelligence and Technology*, vol. 1, no. 1, pp. 51–60, 2021.
- [40] G. Hu, S.-H. Kay Chen, and N. Mazur, “Deep neural network-based speaker-aware information logging for augmentative and alternative communication,” *Journal of Artificial Intelligence and Technology*, vol. 1, no. 2, pp. 138–143, 2021.
- [41] C. Szegedy, S. Ioffe, V. Vincent, 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*, San Francisco, CA, USA, February 2017.
- [42] J. L. Gonzalez Bello, S. Seo, and M. Kim, “Pan-sharpening with color-aware perceptual loss and guided re-colorization,” in *Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP)*, pp. 908–912, IEEE, Abu Dhabi, UAE, October 2020.

Retraction

Retracted: Secure Complex Systems: A Dynamic Model in the Synchronization

Computational Intelligence and Neuroscience

Received 15 August 2023; Accepted 15 August 2023; Published 16 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] A. A. Hamad, M. L. Thivagar, J. Alshudukhi et al., "Secure Complex Systems: A Dynamic Model in the Synchronization," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 9719413, 6 pages, 2021.

Research Article

Secure Complex Systems: A Dynamic Model in the Synchronization

Abdulsattar Abdullah Hamad ¹, M. Lellis Thivagar ², Jalawi Alshudukhi,³
Talal Saad Alharbi ³, Saud Aljaloud ³, Khalid Twarish Alhamazani ³,
and Zelalem Meraf ⁴

¹College of Sciences, Tikrit University, Tikrit, Iraq

²School of Mathematics, Madurai Kamaraj University, Madurai, Tamilnadu, India

³University of Ha'il, College of Computer Science and Engineering Department of Computer Science, Ha'il, Saudi Arabia

⁴Department of Statistics, Injibara University, Injibara, Ethiopia

Correspondence should be addressed to Zelalem Meraf; zelalemmeraf@inu.edu.et

Received 14 November 2021; Accepted 1 December 2021; Published 23 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Abdulsattar Abdullah Hamad 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.

Chaotic systems are one of the most significant systems of the technological period because their qualities must be updated on a regular basis in order for the speed of security and information transfer to rise, as well as the system's stability. The purpose of this research is to look at the special features of the nine-dimensional, difficult, and highly nonlinear hyperchaotic model, with a particular focus on synchronization. Furthermore, several criteria for such models have been examined; Hamiltonian, synchronizing, Lyapunov expansions, and stability are some of the terms used. The geometrical requirements, which play an important part in the analysis of dynamic systems, are also included in this research due to their importance. The synchronization and control of complicated networks' most nonlinear control is important to use and is based on two major techniques. The linearization approach and the Lyapunov stability theory are the foundation for attaining system synchronization in these two ways.

1. Introduction

The Lyapunov method, which is referred to as the Lyapunov stability criterion, uses a Lyapunov $V(x)$ function that is similar to the potential function of classical dynamics. It is given as follows for a system ($\dot{x} = f(x)$) that has an equilibrium point at $x=0$. In the recent past, the study of nonlinear continuous dynamical systems has been considerable. It is one of the first trials in the Lu model [1]. It presents the new structure of high dimension (9D), novel king of quaternion complete, and has some unusual properties [2–4]. There is another study, which introduces another chaotic and hyperchaotic complex nonlinear, and this type has and its phase-space behavior holds a great amount of weight [6–9]. It was previously structured, for example, a

three-dimensional auto system that is not differ-isomorphic with the Lorenz attractor. Lü [10] presented another 3D attractor that is chaotic in different ways and is not diffeomorphic with Lorenz [3, 4, 11, 12] in terms of the arrangement of values for a parameter k . Lorenz [5], an extension of the Lorenz system, proposed the first chaotic nonlinear system. The mechanics of liquid flows' calorific convection are simulated using the Lorenz system's messy structure [13]. Many chaotic and super-chaotic complex systems with nonlinear quadratic conditions have been proposed in [13–15]. The Lorenz equation is used to combine these systems with nonlinear quadratic components. In the suggested model, the variables x and y are taken to be functions for one real and three complex parameters. It was previously structured; for example, a three-dimensional auto

system that is not differ-isomorphic with the Lorenz attractor. Lü has suggested another 3D attractor that exhibits chaotic behavior in different ways and is not diffeomorphic with Lorenz in terms of the arrangement of values for a parameter k .

The z variable, on the contrary, can only be predicted using real variables. A set of nine equations illustrates the hypercomplex chaotic system after extensive mathematical modifications. The collected system dynamics were examined, including phase spaces, eigenvalue computations, and Lyapunov exponent calculations, as well as all other studies. When compared to the findings obtained with 6D models, the suggested approach reveals an acceptable level of accuracy. Several research publications have attempted to look into the geometry of nature, complex-dynamic network synchronization, and regulation of specific topologies. It can be assumed that the three-following 1st O.D.E. characterizes a model proposed by Alyami and Mahmoud [5]. Lots of systems have been studied such as 2D, 3D, 4D, and 8D.

In this paper, we will present a study on a proposed model 9D and try to verify its validity, as the mathematical equations will be mentioned in sequence.

1.1. Contribution

- (i) We have proposed the system which clearly depicts that while increasing the parameter, the corresponding values of are decreased
- (ii) From a detailed analysis and results, we have achieved featured for the 9-dimensional, complex, and highly nonlinear hyperchaotic model which has been executed to make a model more dynamic
- (iii) For various models, several criteria have been examined, such as Hamiltonian, Synchrony, Lyapunov expansion, and stability

2. Structure

In the present section, a short review of the model developed by Lellis, and Hamad is presented to ensure continuity of the concept of the present study.

The system depends entirely on the mathematical equations that enable us to reach numerical results, as the nine-year system is considered one of the most recent and most complex systems. Circuit simulation is a process in which a model of an electronic circuit is created and analyzed using various software algorithms, which predict and verify the behavior and performance of the circuit. Here, we can explain and give the mathematical equations that the system adopts to prove its testability of its validity:

$$\dot{x}(t) = \alpha(y - x), \quad (1)$$

$$y'(t) = (\gamma x - y - xz), \quad (2)$$

$$\dot{z}(t) = -\beta z + \frac{1}{2}(y\bar{x} + x\bar{y}), \quad (3)$$

where α, β , and γ are real parameters, and the variables x, y , and z are defined as

$$x = u_1 + iu_2 + ju_3 + ku_4, \quad (4)$$

$$y = u_5 + iu_6 + ju_7 + ku_8, \quad (5)$$

$$z = u_9. \quad (6)$$

The complex variables were calculated using the traditional Lu Model. It may be deduced from the given model that the model was created by replacing actual variables. This is seen in equations (4) and (5). The system has been proposed and achieved at large sizes. Two alternative strategies are intended to be used to build a higher-dimensional model, which may be accomplished by adding additional variables to the first system's original system. It takes into account the second by integrating two existing models in order to produce a stable system. This procedure requires extra caution. It is the initial way that has been picked to create the present process in this text.

Equations (1) through (2) are treated mathematically; equations (4) and (5) are substituted into both sides of equation (1) as follows:

$$\dot{u}_1 + i\dot{u}_2 + j\dot{u}_3 + k\dot{u}_4 = \alpha((u_5 + iu_6 + ju_7 + ku_8) - (u_1 + iu_2 + ju_3 + ku_4)). \quad (7)$$

Both sides have been the manipulation of equation (7), and after a long mathematical, the first-order differential equation will get it from the following system:

$$\begin{aligned} \dot{u}_1 &= \alpha(u_5 - u_1), \\ \dot{u}_2 &= \alpha(u_6 - u_2), \\ \dot{u}_3 &= \alpha(u_7 - u_3), \\ \dot{u}_4 &= \alpha(u_8 - u_4), \end{aligned} \quad (8)$$

$$\begin{aligned} \dot{u}_5 &= \gamma u_1 - u_5 - u_1 u_9, \\ \dot{u}_6 &= \gamma u_2 - u_6 - u_2 u_9, \\ \dot{u}_7 &= \gamma u_3 - u_7 - u_3 u_9, \\ \dot{u}_8 &= \gamma u_4 - u_8 - u_4 u_9, \\ \dot{u}_9 &= -\beta u_9 + (u_1 u_5 + u_2 u_6 + u_3 u_7 + u_4 u_8). \end{aligned} \quad (9)$$

2.1. Hamiltonian Dynamics. In this section, in particular, the study considers the system smooth and nonlinear; then, it is generalized, and the Hamiltonian canonical form takes the following way:

$$\dot{\chi}(x) = \tau(x) \frac{\partial H}{\partial x} + S(x) \frac{\partial H}{\partial x}, \quad x \in R^n. \quad (10)$$

In equation (9), H is a smooth energy function.

The vector-matrix on the right-hand side is the primary system itself; therefore, the Hamiltonian takes the following form:

$$\dot{\chi}(x) = \frac{1}{2\alpha} [U_1^2 + U_2^2 + U_3^2 + U_4^2 + U_5^2 + U_6^2 + U_7^2 + U_8^2 + U_9^2]. \quad (11)$$

Equation (10) is the new Hamiltonian.

2.2. Symmetry and Invariance. Both are invariant because of the invariance:

$$u_1, u_2, u_3, u_4, u_5, u_6, u_7, u_8, u_9, \quad (12)$$

$$-u_1, -u_2, -u_3, -u_4, -u_5, -u_6, -u_7, -u_8, -u_9. \quad (13)$$

It can be proven that there are two solutions by the following substitution by introducing

$$\begin{cases} u_K^r = u_K^r \cos \Theta - u_K^i \sin \Theta, \\ u_K^i = u_K^i \sin \Theta + u_K^r \cos \Theta, \quad k = 1, 2, 3, \dots, 9. \end{cases} \quad (14)$$

r : real

i : imaginary

$k = 1, 2, 3, \dots, 9$

2.3. Equilibria. In this section, equilibrium is needed to find the homogenous solution given by the new system by equation (8) as follows:

$$\begin{aligned} \alpha(u_5 - u_1) &= 0 \implies u_1 = u_5, \\ \alpha(u_6 - u_2) &= 0 \implies u_2 = u_6, \\ \alpha(u_7 - u_3) &= 0 \implies u_3 = u_7, \\ \alpha(u_8 - u_4) &= 0 \implies u_4 = u_8, \\ \gamma u_1 - u_5 - u_1 u_9 &= 0 \implies u_1(\gamma - u_9 - 1) = 0, \\ \gamma u_2 - u_6 - u_2 u_9 &= 0 \implies u_2(\gamma - u_9 - 1) = 0, \\ \gamma u_3 - u_7 - u_3 u_9 &= 0 \implies u_3(\gamma - u_9 - 1) = 0, \\ \gamma u_4 - u_8 - u_4 u_9 &= 0 \implies u_4(\gamma - u_9 - 1) = 0, \\ -\beta u_9 + (u_1^2 + u_2^2 + u_3^2 + u_4^2) &= 0. \end{aligned} \quad (15)$$

From the 5th row in equation (14) till the 8th row, one can obtain the only possible solution $u_9 = \gamma - 1$; while all remaining variables equal zero, i.e., $u_1 = u_2 = u_3 = u_4 = u_5 = u_6 = u_7 = u_8 = 0$.

2.4. Stability. The authors have referred back to Jacobin, and the study has proposed a system by equation (14) to find characteristic eques as follows to check the stability:

$$\|J - \lambda\| = \begin{pmatrix} -\alpha - \lambda & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & -\alpha - \lambda & 0 & 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & -\alpha - \lambda & 0 & 0 & 0 & \alpha & \alpha & 0 \\ 0 & 0 & 0 & -\alpha - \lambda & 0 & 0 & 0 & 0 & 0 \\ \gamma & 0 & 0 & 0 & -1 - \lambda & 0 & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 & 0 & -1 - \lambda & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 & 0 & -1 - \lambda & 0 & 0 \\ 0 & 0 & 0 & \gamma & 0 & 0 & 0 & -1 - \lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\beta - \lambda \end{pmatrix} = 0. \quad (16)$$

2.5. Dissipation. From the divergence condition and the dissipation of the system, it will be examined follows:

$$\nabla \bullet V = \sum_{i=1}^{N=9} \frac{\partial \dot{u}_i}{\partial u_i} < 0. \quad (17)$$

By computing after simplifying partial derivatives, the dissipation condition is as follows:

$$\nabla \bullet V = -4\alpha - \beta - 3 = -(4\alpha + \beta + 3). \quad (18)$$

It is possible to state that a circumstance is at the root of the occurrence of chaotic behavior.

2.6. Lyapunov Exponents. Referring back to the system derived by equation (8) and recasting, it is in a general matrix form, and it will take the following form:

$$\dot{U}(t) = \xi(U(t); \zeta), \quad (19)$$

where

$$\zeta = (\zeta_1, \zeta_2, \zeta_3, \dots, \zeta_9)^T. \quad (20)$$

Now, it defines the deviation in solving as follows:

$$\delta \dot{U}(t) = L_{ij}(U(t); \zeta) \delta U(t), \quad (21)$$

$$L_{ij}(U(t); \zeta) = \frac{\partial U(t)_i}{\partial u_j}, \quad (22)$$

$$J_{ij} = \begin{pmatrix} -\alpha & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\ 0 & -\alpha & 0 & 0 & 0 & \alpha & 0 & 0 & 0 \\ 0 & 0 & -\alpha & 0 & 0 & 0 & \alpha & 0 & 0 \\ 0 & 0 & 0 & -\alpha & 0 & 0 & 0 & \alpha & 0 \\ \gamma & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & \gamma & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & \gamma & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & \gamma & 0 & 0 & 0 & 0 & -\beta \end{pmatrix}, \quad (23)$$

$$\|J_{ij}\| = \alpha^4 \beta (-\gamma^4 + 4\gamma^3 - 6\gamma^2 + 4\gamma - 6). \quad (24)$$

The choice of the parameters α, β , and γ determines the hyperchaotic system, and it is essential to remain as it is by

increasing the number of positive Lyapunov exponents and the occurrence of the hyperchaotic.

2.7. Lyapunov Attractors. As reported by Yorke–Kaplan, the Lyapunov dimension of the dynamic system's attractors is

$$D = \alpha^* + \frac{\sum_{K=1}^{\alpha^*} L_{ij}}{|L_{\alpha^*+1}|}, \quad (25)$$

In equation (24), α^* is the highest integer for which $\sum_{K=1}^{\alpha^*} L_{ij}$ should be negative.

2.8. Model of a Weighted-Complex Network. In this study, the authors will apply a weighted-complex-dynamic model provided by [6, 9], to the dynamic model they built. The following is what the model suggested:

$$\begin{cases} \frac{\partial y_j}{\partial t} = g(y_j) + \sum_{j=1}^M C_{ij} \Delta y_j, & 1 \leq i \leq M, \\ y_j = (y_{i1}, y_{i2}, y_{i3}, \dots, y_{im})^T \in R^n. \end{cases} \quad (26)$$

3. The Method of Adaptive Control

The identity in this section is regulated by the following:

$$\begin{cases} \frac{\partial \hat{y}_j}{\partial t} = g(\hat{y}_j) + \sum_{j=1}^M \hat{C}_{ij} \Delta \hat{y}_j + \zeta_i, & 1 \leq i \leq M, \\ \hat{y}_j = (\hat{y}_{i1}, \hat{y}_{i2}, \hat{y}_{i3}, \dots, \hat{y}_{im})^T \in R^n. \end{cases} \quad (27)$$

By making use of the following assumption,

$$\begin{cases} \dot{\bar{y}} = \hat{y}_j - y_j, \\ \bar{C} = \hat{C}_j - C_j, \end{cases} \quad (28)$$

the error system formula will take the following form:

$$\begin{cases} \dot{\bar{y}} = g(\hat{y}_j) - g(y_j) + \sum_{j=1}^M \bar{C}_{ij} \Delta y_j + \sum_{j=1}^M C_{ij} \bar{y}_g + \zeta_i, \\ \bar{y} = \hat{y}_j - y_j, \\ \bar{C} = \hat{C}_j - C_j. \end{cases} \quad (29)$$

4. Numerical Results and Discussion

In this work, four sets of the parameters α, β , and γ are tested, and the corresponding Lyapunov exponents are computed. These sets are shown in Table 1, and Table 2 shows numerical values.

Initial conditions:

$$\begin{aligned} u_1(0) &= 0.999, \\ u_2(0) &= 1.999, \\ u_3(0) &= 2.999, \\ u_4(0) &= 3.999, \\ u_5(0) &= 4.999, \\ u_6(0) &= 5.999, \\ u_7(0) &= 6.999, \\ u_8(0) &= 7.999, \\ u_9(0) &= 8.999. \end{aligned} \quad (30)$$

By MATLAB software, we will include and analyze the numerical values.

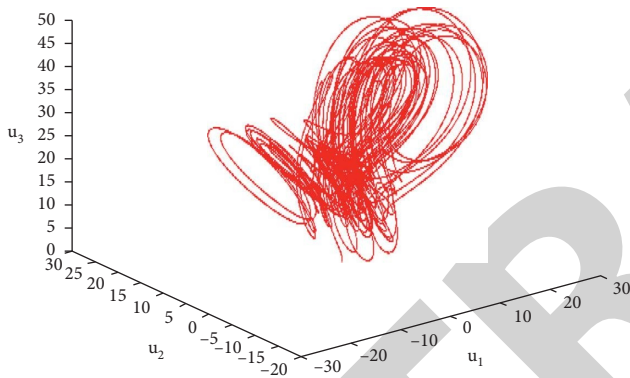
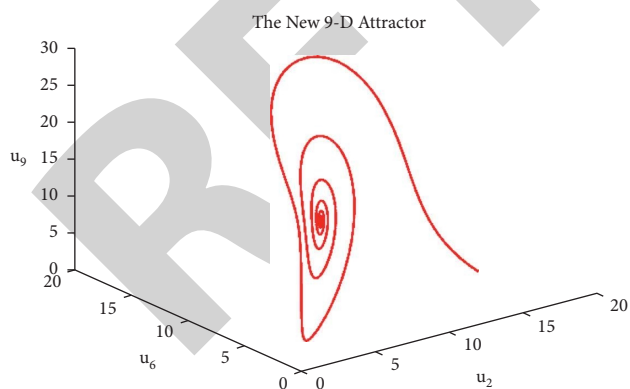
4.1. Attractors of a Proposed System. Figures 1 and 2 show studies of the control and synchronization behavior of a particular model are of practical interest. In light of this, the current work focuses on

TABLE 1: Lyapunov exponents.

Parameters (set)	Set (1)	Set (2)	Set (3)	Set (4)
α	10	12	14	16
β	3.7	4.7	5.7	6.7
γ	55	56	57	58

TABLE 2: Numerical values.

	Set (1)	Set (2)	Set (3)	Set (4)
L_1	1.01	1.009	1.005	1.002500
L_2	0.741	0.74050	0.74020	0.740050
L_3	0.41	0.40050	0.40025	0.4000100
L_4	0.141	0.14050	0.14004	0.140025
L_5	0	0.00000	0.00000	0.000000
L_6	-17.0051	-17.0071	-17.0081	-17.0095
L_7	-23.9108	-23.9118	-23.9127	-23.9132
L_8	-70.2753	-70.2761	-70.2753	-70.2772
L_9	-70.3127	-70.3138	-70.3142	-70.3145

FIGURE 1: Attractor of the system in the (u_1, u_2, u_3) space.FIGURE 2: Attractor of the system in the (u_2, u_6, u_9) space.

- (i) Design an adaptive control strategy for the synchronization phenomena of the chaotic system with known and unknown parameters
- (ii) Design a nonlinear control function capable of controlling the hyperturbulent system to stabilize any situation and follow any path while being smooth functionality

- (iii) Design a control function capable of synchronizing two identical or different chaotic systems evolving under different conditions

5. Conclusions

In the present research work, the authors have proposed the system by four different sets for the parameters as shown Tables 1 and 2. The results, due to these sets, are shown in Figures 1-2. As it is clear that while increasing the parameter, the corresponding values decreased on 9D. Dynamic systems are one of the important systems of the era of technology, as the properties of these systems always need to be updated, for the speed of security and information transfer increases as well as the stability of the system. The present study aims to analyze the detailed features of the nine-dimensional, complex, and highly nonlinear hyperchaotic model. Furthermore, various criteria, such as Hamiltonian, synchronization, Lyapunov expansion, and stability, have been investigated for such models.

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 regarding the publication of this paper.

Acknowledgments

This work was performed as a part of the Employment of Institutions.

References

- [1] M. L. Thivagar, A. A. Hamad, and S. G. Ahmed, "Conforming dynamics in the metric spaces," *Journal of Information Science and Engineering*, vol. 36, no. 2, pp. 279–291, 2020.
- [2] M. A. Alyami and E. E. Mahmoud, "Quaternion nonlinear Lü model and its novel quaternion complete synchronization," *Alexandria Engineering Journal*, vol. 59, no. 3, pp. 1391–1403, 2020.
- [3] B. Bao, J. Xu, Z. Liu, and Z. Ma, "Hyperchaos from an augmented Lü system," *International Journal of Bifurcation and Chaos*, vol. 20, no. 11, pp. 3689–3698, 2010.
- [4] N. Smaoui, M. Zribi, and T. Elmokadem, "A novel secure communication scheme based on the Karhunen-Loève decomposition and the synchronization of hyperchaotic Lü systems," *Nonlinear Dynamics*, vol. 90, no. 1, pp. 271–285, 2017.
- [5] S. Sengan, O. I. Khalaf, P. D. K. Vidya Sagar, D. K. Sharma, L. Arokia Jesu Prabhu, and A. A. Hamad, "Secured and privacy-based IDS for healthcare systems on E-medical data using machine learning approach," *International Journal of Reliable and Quality E-Healthcare*, vol. 11, no. 3, pp. 1–11, 2022, p.
- [6] G. Zhang, Z. Guo, Q. Cheng, and I. Sanz, "Multi-level integrated health management model for empty nest elderly

Retraction

Retracted: Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements

Computational Intelligence and Neuroscience

Received 25 July 2023; Accepted 25 July 2023; Published 26 July 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] M. Bader Alazzam, H. Mansour, M. M. Hammam et al., "Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 1094054, 6 pages, 2021.

Research Article

Machine Learning of Medical Applications Involving Complicated Proteins and Genetic Measurements

Malik Bader Alazzam ¹, Hoda Mansour,² Mohamed M. Hammam,³ Said Alsheikh,⁴ Ali Bakir,⁴ Saeed Alghamdi,⁵ and Ahmed S. AlGhamdi ⁶

¹Faculty of Computer Science and Informatics, Amman Arab University, Amman, Jordan

²College of Business Administration, University of Business and Technology, Jeddah, Saudi Arabia

³Theodor Bilharz Research Institute TBRI, Giza, Egypt

⁴University of Business and Technology, Jeddah, Saudi Arabia

⁵Taibah University, Taibah, Saudi Arabia

⁶Department of Computer Engineering, College of Computers and Information Technology, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

Correspondence should be addressed to Malik Bader Alazzam; m.alazzam@aau.edu.jo

Received 8 November 2021; Revised 16 November 2021; Accepted 30 November 2021; Published 21 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Malik Bader Alazzam 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.

Motivations. Breast cancer is the second greatest cause of cancer mortality among women, according to the World Health Organization (WHO), and one of the most frequent illnesses among all women today. The influence is not confined to industrialized nations but also includes emerging countries since the authors believe that increased urbanization and adoption of Western lifestyles will lead to a rise in illness prevalence. **Problem Statement.** The breast cancer has become one of the deadliest diseases that women are presently facing. However, the causes of this disease are numerous and cannot be properly established. However, there is a huge difficulty in not accurately recognizing breast cancer in its early stages or prolonging the detection process. **Methodology.** In this research, machine learning is a field of artificial intelligence that employs a variety of probabilistic, optimization, and statistical approaches to enable computers to learn from past data and find and recognize patterns from large or complicated groups. The advantage is particularly well suited to medical applications, particularly those involving complicated proteins and genetic measurements. **Result and Implications.** However, when using the PCA method to reduce the features, the detection accuracy dropped to 89.9%. IG-ANFIS gave us detection accuracy (98.24%) by reducing the number of variables using the “information gain” method. While the ANFIS algorithm had a detection accuracy of 59.9% without utilizing features, J48, which is one of the decision tree approaches, had a detection accuracy of 92.86% without using features extraction methods. When applying PCA techniques to minimize features, the detection accuracy was lowered to the same way (91.1%) as the Naive Bayes detection algorithm (96.4%).

1. Introduction

According to WHO data, millions of people died from cancer throughout the world, accounting for 70% of all fatalities and a nearly 50% rise in mortality in emerging countries compared with the preceding era [1, 2]. According to several physicians' studies, underdeveloped nations only have 5% of the worldwide budget to battle cancer. Furthermore, these countries have little material and human

resources. Breast cancer arises from breast cells, and there are two forms of cancer: benign and malignant. Breast cancer, on the other hand, is a deadly disease (a group of cancer cells). Breast cancer is most commonly associated with women, but it may also strike men. Breast cancer is a problem that has the potential to affect every region of the body [3–6]. It is possible to infect women with little bits of cancer that are not huge and may be felt or identified by changes that arise in the breast although clear symptoms do

not usually appear directly as a result of the disease. The most typical symptoms include a significant rise in breast size, as well as other associated symptoms (2):

- 1 Redness or emergence of the nipple
- 2 Changes in the skin, such as wrinkling and aggregation
- 3 Part of the breast swells

In statistics and computer learning, classification is one of the forms of supervised learning, which involves introducing data provided by a computer program and then doing a classification to discover new findings (2). These data can be of two types: numerous outcomes showing varying percentages or results with only two numbers (such as determining that the condition is acceptable or unacceptable, [7–10] that the person was male or female, or that the disease is benign or malignant). Handwriting recognition, document categorization, speech recognition, and biometric determination are all examples of classification issues that may be worked on (2).

2. Wisconsin Breast Cancer (WBC)

Dataset Description

This work used data on breast cancer patients provided by the Wisconsin University Hospitals, Madison. There are 699 specimens or samples with 10 + 1 qualities in this set of data (1 for class). Table 1 shows the results (2). These samples were split into two categories: benign data (458 instances) and malignant data (241 cases), and there are 16 cases with missing data (3).

After the data have been collected and scanned, it is separated into two groups: training and testing. The training data will be used to train algorithms, while the rest will be utilized to test them. This study's algorithms predict the diagnosis of breast cancer for each sample in the test group. Last but not least, these algorithms are used to do performance analysis, and the optimum analysis for breast cancer is established (3). (4).

3. Performance Measurement Measures

A method known as confusion matrix can be used to improve the classification algorithm's performance. By comparing the number of positively/mistakenly categorized instances and the number of properly/incorrectly classified negative cases, it may be deemed the most effective technique to organize performance and simplify the taxonomy pattern (5). The confusion matrix's columns indicate the anticipated classification, while the rows describe the case's actual categorization, as seen in Table 1 (6).

- 1 True positive (TP): this representation refers to the classification of the patient that is benign
- 2 True negative (TN): this representation indicates the classification of the patient that is malignant disease
- 3 Positive false (FP): this representation indicates the classification of the patient that is benign disease but was malignant disease

- 4 False negative (FN): this representation indicates the classification of the patient that is malignant disease but was benign

Equations used in performance measures of the most widely used accuracy, sensitivity, and specificity of medicine and biology are as follows (6), (7):

$$\text{FP rate} = \frac{\text{FP}}{(\text{TN} + \text{FP})} \%, \quad (1)$$

$$\text{accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \%, \quad (2)$$

$$\text{TP} = \text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \%, \quad (3)$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \%, \quad (4)$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \%, \quad (5)$$

$$\text{F1 score} = \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} * 2\%. \quad (6)$$

- (1) *Specificity*. It is the percentage of true negative outcomes that are correctly identified by the model and in a more precise manner the quality of the model designed to identify correctly the women who did not die from breast cancer
- (2) *Recall*. It is the measurement of the proportion of patients who are expected to have complications and who are already suffering from complications
- (3) *Precision*. It is to measure the proportion of patients who have complications due to the disease such as those that are complications based on the model
- (4) *F1 Score*. It is a weighted average of precision and recall
- (5) *Matthew Correlation Coefficient (MCC)*. It is a performance parameter of a binary classifier

4. Choice Plants (J48)

In data extraction implementations, DT is frequently employed. Because it is simple to grasp, it aids the end-user in data mining. An efficient connection between the data set's properties is provided in an easy-to-understand format. In comparison to other classification methods, this one requires a few computations.

The agreement and the rules are split into two sections in the DT (tests). When building the arbor, a feature test is represented by each node. A flowchart depicts the main idea of this algorithm, which consists of a root node that serves as a starting point for all (nonleaf) nodes but may be considered the algorithm's basic concept. Contracts are viewed as trial runs on the path to the paper node (final result). When using DT to identify breast cancer, the nodes are categorized into two categories: benign and malignant. The

TABLE 1: Detailed accuracy (SVM) without PCA.

	TP rate	FP rate	TP rate	FP rate	PR	SP	RE	FI score	MCC
Benign	100		26		95	70.2	88.1	98.2	58.100
Malignant	77		15		90.3	99.1	65	25	88
WGH	82.10		18		50	70	45	66	41.9

rules will be built based on the properties of the provided dataset to assess whether the tumor is malignant or benign. Figure 1 shows how to use the DT approach to identify breast cancer. cancer (1).

J48 is the enforcement of the DT algorithm ID3 which produces a binary tree (7). The tree is fitted to every line in the database next it is created. J48 was used because it had a relatively high speed compared with other DT algorithms. Moreover, simplicity is one of its unique features, and the results of algorithm can easily be sensed by the end user and accept the performance metric. Based on data from the UCI Machine Learning Repository, the commonly used ratio divides a data set into 80% training group and 20% test group which was applied on the J48. And results are given in Table 2. Information we got in the J48 method by using all futures in dataset with one exception removes instances that are have missing values.

5. The Information Gain and Adaptive Neural Fuzzy Inference System (IG-ANFIS)

Researchers have been working on artificial intelligence (AI) solutions to be utilized in medical and health-related sectors for several years. The following are the most often utilized AI strategies used by researchers to construct extremely efficient automated diagnostic systems:

- (1) Networks of neurons
- (2) Support vector machines
- (3) Fuzzy logic
- (4) Genetic programming algorithms

Because medical diagnosis requires ambiguous and higher-dimensional clinical data, a pressing demand for AI solutions to cope with the different nature of data sets has arisen, which will assist medical practitioners make more effective and precise decisions.

The adaptive neural fuzzy inference system (ANFIS) is a machine learning technique that combines two machine learning approaches: neural networks (NNs) and fuzzy inference systems (FISs). The k-nearest neighbor's technique was employed in this study to create a neural network (NN). ANFIS is developing input and output mapping by combining humanitarian expertise with machine learning capabilities (7).

Information gain (IG) is the simplest method for selecting the best features and is commonly used in text categorization. By assessing the difference between the before and postattributes, the IG method was utilized to evaluate the quality of each feature utilizing attributes (8).

Diseases are diagnosed using the IG-ANFIS technique (in our case, breast cancer). This approach or algorithm is a

hybrid of IG and ANFIS. The goal of IG is to reduce the number of input features to ANFIS (7) (8) by selecting the quality of characteristics for the input data. The outcome of IG is a group of features with high ranking values of input. The features group will be used that has a higher degree as input for ANFIS. The features selected having higher degree will be applied for training and testing on the ANFIS method. The general structure of IG-ANFIS is illustrated in Figure 2 where $Z = (z1, z2, \dots, xn)$ are the original features in the UCI dataset, $V = \{v1, v2, \dots, vm\}$ are the features obtained after information gain, and Q indicates the final output after applying V to ANFIS (diagnostic) (8).

The database having 699 records was divided to (341, 342) records for training and testing sequentially. And there are 16 records which were removed because they contain missing values. The class attributes have been normalized to 0 = benign and 1 = malignant. Table 2 shows the ranking of attributes after applying IG; its selects the quality of attributes (8).

The output for ANFIS after applying the features selected by IG used at WBC dataset gave us 98.24% accuracy ANFIS, while the accuracy of the ANFIS algorithm in detection without extracting characteristics was 59.9% (8).

6. SVM (Support Vector Machine)

Support vector machine (SVM) is a machine learning algorithm that supervises and works on classification and regression problems. In this method, we plot every element of data as a point in the space of n dimensions when n is quantity of features you have and the value of every feature is the value of specific coordinates (7). After that, we make the classification through finding the very high level that characterizes the two classes very well as shown in Figure 3.

Characteristics of the support vector machine are as follows (7) (8):

- (1) Flexibility in the function selection process is given as it is not specified by a particular type
- (2) It has the ability to handle a large number of features in the search space

Machine learning entails predicting and classifying data, and we use a variety of machine learning methods to accomplish this depending on the dataset.

The support vector machine, or SVM, is a linear model that can be used to solve classification and regression issues. It can solve both linear and nonlinear problems and is useful for a wide range of applications. SVM is a basic concept: the method divides the data into classes by drawing a line or hyperplane.

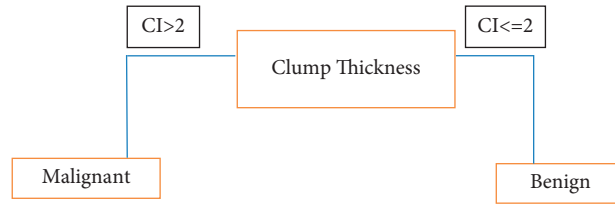


FIGURE 1: Decision tree (J48).

TABLE 2: Details of the comparisons.

Method	Accuracy	Precision	Recall	Specificity	F1-score	MCC
J48	92.86	97.6	91.1	96	94.3	85.2
ANFIS	59.9	*	*	*	*	*
IG-ANFIS	98.24	*	*	*	*	*
SVM	91.7	90	99.1	74.5	94.4	80.25
SVM with PCA	89.9	86.9	100	69.6	93	77.8
Naïve Bayes	96.4	96.9	98.4	90.2	97.7	90.22
Naïve Bayes with PCA	91.1	90	98.3	74	94.15	78.34

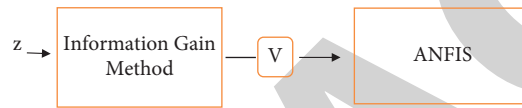


FIGURE 2: General structure for IG-ANFIS.

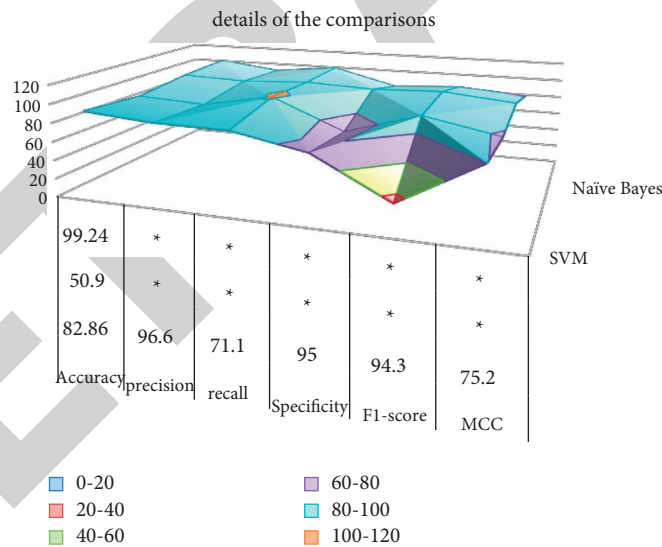


FIGURE 3: Details of the comparisons.

The support vector machine (SVM) is a supervised machine learning technique that can solve classification and regression problems. However, SVM is a borderline that separates two groups. Running on SVM, we also got a method for the process of separating the categories (benign and malignant) in a hyperplane. Here, there are three hyperplanes: A, B, and C. The correct hyperplane is recognized to classify star (benign) and circle (malignant) (6). You must memorize a rule to specify the correct hyperplane. Choose the hyper-plane that separates the two categories best. In this study, hyperplane “B” did an excellent job in this work. Determining the correct hyperplane here, we have three

hyperplanes A, B, and C and all are separating the classes well. Here, maximizing distances between the nearest data points (any category) and the hyperplane will help us to determine the correct hyperplane (9). This is called as margin.

This research studies 569 instances, and there were 357 instances of benign breast cancer and 212 instances of malignant breast cancer. Dataset will be divided as 70% for training and 30% for testing. We have slotted 70% of the dataset to training. Out of the 70% dataset for training, we are using 63% and the rest 7% for validation test (5) which was applied on the SVM. And accuracy results are obtained

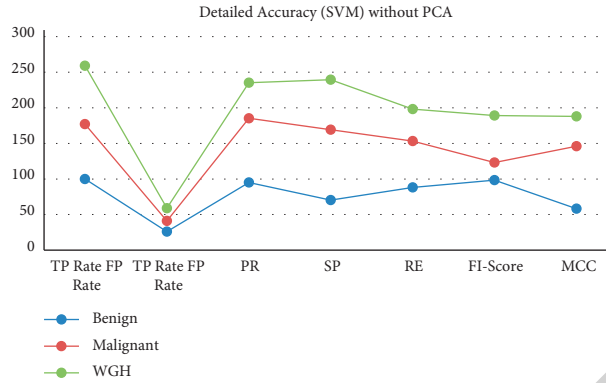


FIGURE 4: Detailed accuracy (SVM) without PCA.

on the SVM method. Table 1 gives us detailed information for SVM on the confusion matrix with all features as shown Figure 4.

And shown in Table 1, accuracy obtained by the SVM method gives us detailed information for SVM on the confusion matrix but with removing some features using the PCA method to get most important features (5).

7. Naïve Bayes

The Bayes theory supports a collection of classification methods known as Naive Bayes. There is not a single mathematical rule involved. Every attempt of the possibilities, however, is a family of algorithms, each of which has a common premise. Being categorized, it is self-employed in a variety of ways. Bayes’ theory employs contingency probability, which calculates the likelihood of a future event based on prior data. The classifier in Naive Bayes is that the input variables are expected to be independent of each other, with each scan choice contributing to the target variable’s probability individually (10). As a result, having a variable for one feature has no influence on the feature variables that relate to it. This might be the cause behind the Naive label. However, in real learning sets, the feature variables are interdependent, which might be one of the Naive Bayes classifier’s drawbacks. In any case, the Naive Bayes classifier is effective for large knowledge groups. Overall, the easy classifiers outperformed the tough classifiers for each form. The hypothesis of Naive Bayes is as follows:

$$P(CK|x) = \frac{(P(CK) * P(x|CK))}{P(x)} \tag{7}$$

For doing so, we need to estimate $P(x|CK)$ and assume that any particular value of vector x conditional on C_K is statistically independent of each dimension. $P(CK|x)$ is the probability guide (7) (10). When premise is true Naive Bayes algorithmic program,

- (1) works for multcategory and binary classification
- (2) It may be trained on a small set of little information and can be a great advantage
- (3) It is the fastest and climbable

- (4) It immigrated the case growing from the damn of locative monarchy to some degree

However, as previously stated, this results in a misleading assumption that the input variables are self-employed from another. This cannot be the case in real-world data sets because there are several high-level correlations among the feature variables. Measurement of prediction (11) is as follows:

- 1 Step 1. Create a frequency table from the data collection
- 2 Step 2. Using likelihoods, create a table of probabilities
- 3 Step 3. Calculate the back probabilities using the Naive Bayes algorithm

Prognosis determines which class has the highest posterior probability.

There were 357 cases of benign cancer and 212 cases of malignant cancer among the 569 cases studied by WBC. 70% of the dataset will be used for training and 30% in testing. We have dedicated 70% of the dataset to training (5). We used 63% of the 70% dataset for training and the remaining 7% for validation testing. On the Naive Bayes, this was used (5).

Another statistic for evaluating the success of a classification algorithm is the confusion matrix. The confusion matrix’s language, true to its name, might be perplexing, but the matrix itself is straightforward to comprehend. I first learned about the confusion matrix, accuracy, precision, recall, F1-score, ROC curve, true positives, false positives, true negatives.

8. Results and Discussion

Figures 4 and 5 show that for classification and decision-making, the J48 group of classifiers is commonly employed. As assessed in this article, three prominent J48 group classifiers, namely, J48, J48Consolidated, and J48Graft, are unique in their field, employing both single- and multidatasets over thirteen performance matrices for suitable rank allocation, whereas ANFIS exhibited less detection performance when using all the features. Then, by using the knowledge gain (IG) method to give the best characteristics and applying them on ANFIS, we got the highest detection performance compared with other methods,

Retraction

Retracted: Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy

Computational Intelligence and Neuroscience

Received 3 October 2023; Accepted 3 October 2023; Published 4 October 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.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

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] J. Alshorman, Y. Wang, G. Huang, T. B. Serebour, and X. Guo, "Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8560668, 6 pages, 2021.

Research Article

Detection and Analysis of Perfusion Pressure through Measuring Oxygen Saturation and Requirement of Dural Incision Decompression after Laminectomy

Jamal Alshorman , Yulong Wang, Guixiong Huang, Tracy Boakye Serebour, and Xiaodong Guo 

Department of Orthopedics, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

Correspondence should be addressed to Xiaodong Guo; xiaodongguo@hust.edu.cn

Received 2 November 2021; Revised 16 November 2021; Accepted 19 November 2021; Published 6 December 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Jamal Alshorman 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. Traumatic spinal cord injury (SCI) can continue and transform long after the time of initial injury. Preventing secondary injury after SCI is one of the most significant challenges, and early intervention to return the blood flow at the injury site can minimize the likelihood of secondary injury. **Objective.** The purpose of this study is to investigate whether laminectomy can achieve the spinal cord blood flow by measuring the spinal blood oxygen saturation intraoperatively without the presence of light. **Methods.** Between June and August 2021, eight patients were admitted after traumatic spinal cord injury for surgical treatment. We explored the effectiveness of laminectomy and whether the patients required further procedures or not. We used a brain oxygen saturation monitor at the spine injury site under dark conditions. **Results.** Eight cervical trauma patients, six males and two females, underwent laminectomy decompression. Three patients' ASIA grade improved by one level, and one patient showed slight motor-sensory improvement. Oxygen saturation was in the normal range. **Conclusion.** Performing bony decompression can show good results. Therefore, finding an examination method to confirm the improvement of blood perfusion by measuring oxygen saturation at the injury site after laminectomy is essential to avoid other complications.

1. Introduction

SCI is a life-threatening condition, and deciding the proper treatment method has a significant role in its prognosis. However, some patients showed substantially positive outcomes after laminectomy, while some required an additional durotomy. A primary spine injury can lead to consequential injury. This secondary injury is often catastrophic and is accompanied by a variety of complications [1]. At least 25 well-established secondary injury mechanisms can occur within minutes, weeks, and months following SCI [2]. If the probability of the occurrence of secondary injury is reduced, the loss of nerve function after SCI will be decreased and it will significantly improve the chances of regaining some lost neurological functions. Decompression surgery opens the

bony canals through which the spinal cord and nerve pass, creating more space for them to move freely. The laminectomy opens a window allowing the dural sac (containing nerves) to re-expand in the space. After extensive bony decompression, cord compression against the dura may only occur in a few patients [3]. Performing durotomy with duroplasty can reduce the intraspinal pressure (ISP) and decrease the spinal cord compartment syndrome (SCCS), improving spinal cord perfusion pressure (SCPP) and blood perfusion at the injury site [4, 5]. However, to perform durotomy and duroplasty, an experienced surgeon is needed because it causes a higher risk of injury such as spinal cord swelling and cerebrospinal fluid leakage [6]. Confirming blood perfusion after performing laminectomy is essential to avoid durotomy complications and their associated risk

factors. However, if decompression during laminectomy cannot release dural pressure, an additional dural incision and decompression can be done, and the cerebrospinal fluid circulation is highly likely to return.

Spinal dural decompression can reduce intradural pressure, thereby inhibiting spinal cord swelling. The occurrence of SCCS after SCI is common and leads to poor prognosis. Dural incision and duroplasty are designed to alleviate the symptoms of this SCCS [5, 7]. With the control of SCCS and the reduction of secondary damage, patients may have better nerve recovery. When decompression is performed at the level of the dura, nerve function is significantly restored [8, 9]. The prognosis will depend on the blood flow to the damaged area. This study shows the effectiveness of measuring oxygen saturation after laminectomy and assesses whether the patients require further procedures (dural decompression). Also, this study aims to confirm the benefits and effectiveness of bony decompression on SCPP in SCI patients.

1.1. Laminectomy Alone versus Durotomy with Duroplasty. Durotomy helps to decompress the spinal cord thoroughly and improves CSF circulation in severe SCI; also, it decreases ISP to less than 20 mmHg. While laminectomy alone is insufficient in some cases, durotomy decompression will be needed [4, 5]. The hematoma between the dura and spinal cord can increase the pressure that cannot be relieved by laminectomy. However, durotomy with duroplasty effectively reduces ISP, improves SCPP, and microcirculation of the spinal cord; CSF pulsation will be restored [10]. Restoring normal gradients of ISP across the spinal cord enhances blood perfusion, blocking possible secondary ischemia [11]. Because of the risks associated with durotomy and its complications, examining the laminectomy procedure's effectiveness (case-by-case) before performing durotomy is essential. Finding a method to ensure SCPP improvement after laminectomy is necessary to produce a better prognosis. If the spinal cord oxygen saturation is normal after sufficient bony decompression is completed, then SCPP has improved. A dural incision decompression is not required.

2. Method and Surgical Procedure

2.1. Inclusion and Exclusion Criteria. We have chosen the trauma patients older than 18 years old, with radiological abnormalities in the cervical region.

Between June and August 2021, eight patients underwent bony decompression and internal fixation. The operation was performed as early as possible (between 24 hours and 16 days).

After obtaining general anaesthesia, the patients were placed in the prone position, with the neck gently bent for a better view, and a head holder was used. After making a midline incision, the fascia and muscles were dissected to reveal the posterior arch. The midline incision should be more than 12 cm. We adopted the same surgical procedure: posterior laminectomy followed by oxygen saturation

measurement in all patients. Before extensive laminectomy, spinal blood oxygen saturation was measured first and recorded. After extensive laminectomy, spinal blood oxygen saturation was measured and recorded again. If spinal blood oxygen saturation was not improved or remained the same as the first measurement, more extensive laminectomy or durotomy with duroplasty was needed. After an extensive laminectomy (3–4 segments), the spinal cord pulsation (CSF circulation) was observed to check the improvement of blood perfusion. The injury site was checked for the existence of hematoma by direct visualization. The probe was inserted to measure the spinal oxygen saturation about 1 cm away from the dural sac (see Figures 1 and 2). Spinal blood oxygen saturation was measured in a dark environment. After that, depending on the oxygen saturation being in normal range (58%–68%), the decision to perform durotomy or not was made. Posterior fusion was performed in all cases.

2.2. Working Principle. The brain oxygen monitor uses near-infrared light with a wavelength between 700 and 900 nm. Since near-infrared light has good penetrability to the skin, fat, and skull, it can penetrate the aforementioned outer tissues and detect the blood oxygen parameters of deeper brain and muscle tissues. The near-infrared light enters the tissue from the light source of the probe, passes through the tissue to be tested, and is received by the receivers at two different positions of the probe. The probe of oxygen saturation measurement will not cause any trauma or pain to the patient; there is no lag in data collection and continuous and uninterrupted data collection; there is no need to infer blood oxygen information of local tissues indirectly through the blood oxygen information of the whole body. We used this equipment to monitor spinal blood oxygen saturation directly at the injury site. This was done to investigate whether there was a change in the blood oxygen saturation of the dura mater in patients with SCI and the effect on the prognosis of the patient.

3. Results

Eight patients with traumatic cervical SCI came to the Orthopedic Department of Union Hospital of Tongji Medical College for treatment. The youngest patient was 35 years old, and the oldest one was 66 (six males and two females). Physical examination and American Spinal Injury Association impairment scale were used to grade the severity of SCI. However, ASIA grades before the operation were (A in five patients, D in two patients, and C in one patient). Surgical intervention was between 24 hours and 16 days after injury. Laminectomy and vertebral body fusion were performed, and bone graft was applied.

Before extensive laminectomy, spinal blood oxygen saturation was measured first and recorded. After extensive laminectomy, spinal blood oxygen saturation was measured and recorded again. If spinal blood oxygen saturation was not improved or remained the same as the first measurement, more extensive laminectomy or durotomy with

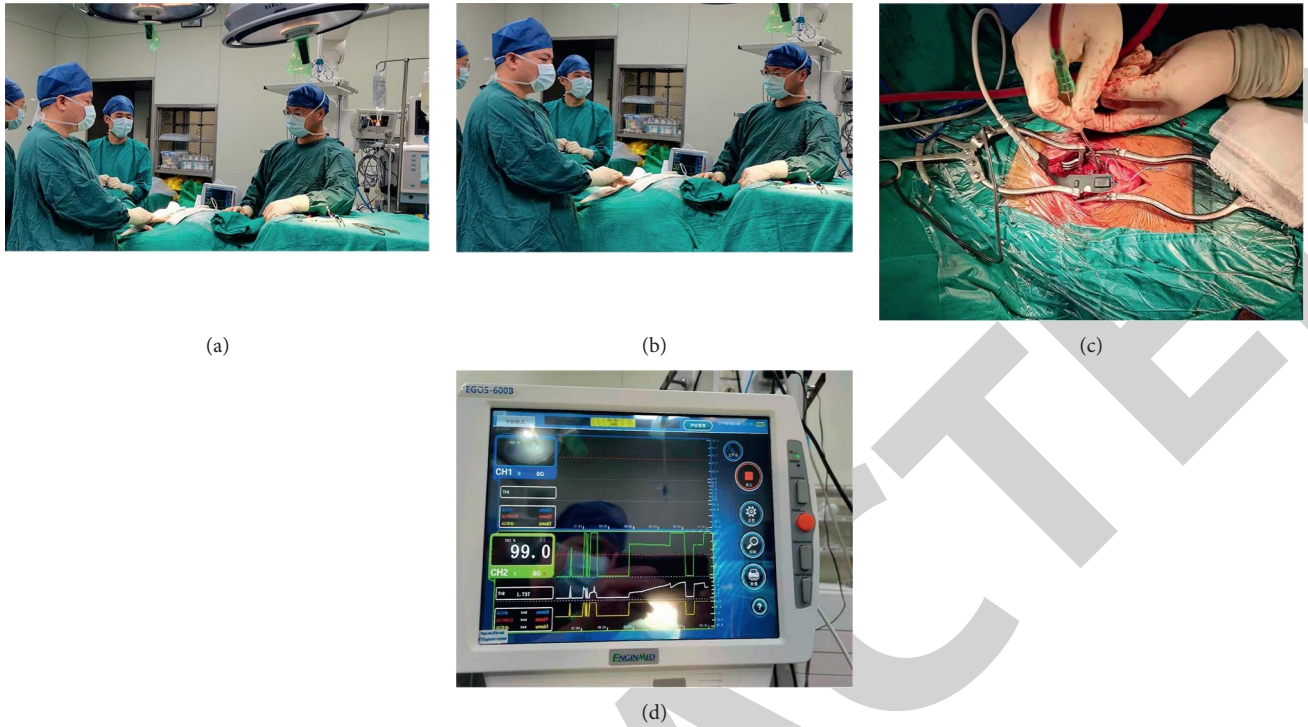


FIGURE 1: (a, b) Laminectomy procedure and spinal blood oxygen saturation measurement under dark conditions during the operation. (c) The probe. (d) The intraoperative display indicating that the spinal blood oxygen saturation is 68.3%.

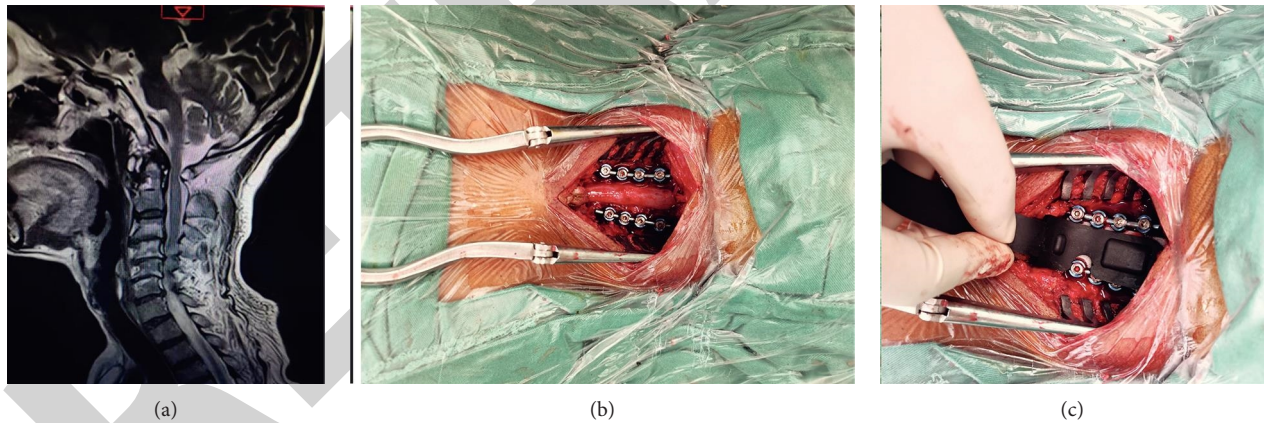


FIGURE 2: 66-year-old male with ASIA grade D underwent laminectomy and spinal blood oxygen saturation measurement. (a) Preoperative MRI. (b) C3–C6 bony decompression. (c) Insertion of the probe for oxygen saturation measurement. Oxygen saturation measurement failed in this patient.

duroplasty was needed. We failed to measure oxygen saturation of one patient. Two patients were transferred to the intensive care unit after the operation and underwent a tracheotomy. On the last follow-up, the ASIA grade improved in three patients, 2 patients (A to B) and one (A to D), two remained the same (one D and one A), and one patient died; the details of other two patients are not stated here. Sensory function improved in all patients except two (the one whose grade remained the same and the one who died) (see Table 1).

4. Discussion

High ISP and low SCPP can cause low oxygen saturation of the spinal cord tissue, which requires surgical decompression. Finding a method to examine blood flow at the injury site is essential in every patient to predict the prognosis, ensure no other treatment method is required, and avoid other procedure complications [6]. The pressure around the spinal cord can lead to a decrease in the blood flow at the injury site that can cause ischemia of the spinal tissue [12].

TABLE 1: A clinical summary of patients (laminectomy followed by spinal blood oxygen saturation measurement).

Cases	Age	Gender	Injury mechanism	Operation level	ASIA (preoperative)	ASIA (postoperative)	Time of operation	Oxygen saturation
1	35 y	M	Fall	C2–C6	D	D	102 hours	Normal
2	55 y	M	Fall	C2–C6	A	B	24 hours	Normal
3	64 y	F	Fall	C3–C5	A	A	30 hours	Normal
4	58 y	M	Fall	C3–C6	A	Died	24 hours	Normal
5	52 y	F	Fall	C3–C6	A	B	64 hours	Normal
6	57 y	M	Head injury by heavy object	C2–C6	A	D	29 hours	Normal
7	56 y	M	RTA	C3–C6	C	—	3.18 hours	Normal
8	66 Y	M	NS	C3–C6	D	—	16 days	Failed to measure

ASIA impairment scale: American Spinal Injury Association impairment scale; M: male; F: female; RTA: road traffic accident.

Measuring the oxygen saturation at the injury site can reveal whether laminectomy is enough or not (see Figure 3).

The significance of monitoring the blood oxygen saturation of the spinal cord during the operation is as follows. (1) During the operation, the blood oxygen saturation of the spinal cord tissue at the injury site can be directly measured, which can indirectly reflect the blood perfusion pressure of the spinal cord. (2) Whether the spinal cord decompression is complete can be determined. The tissue oxygen saturation is restored or improved, indicating that the decompression is effective. (3) Guidance on whether to do more extensive lamina decompression or more thorough dural incision decompression and other operations can be given.

Secondary injuries of SCI occur through multi-variant pathological mechanisms [2, 13]. Inflammatory process and increase in the pressure lead to tissue damage [14, 15]. Preventing secondary injury after SCI is one of the significant challenges [10, 16, 17]. However, the dura mater compressed the swollen spinal cord, leading to raised ISP and a progressive reduction of blood flow, leading to chronic hypoxia [8, 18, 19]. An appropriate treatment method is essential to decrease ISP and improve SCPP. However, performing bony decompression is not enough to reduce the pressure around the injury site. Even if bony decompression is performed early, gradually, or instantly, swelling can create subarachnoid occlusion [20]. Ensuring the blood flow circulation returns to normal is the primary purpose of predicting the prognosis in surfer SCI [21].

In brain trauma, maintenance of adequate cerebral oxygenation is critical in managing patients with brain injury; otherwise, it leads to hypoxia [22]. Maintaining the supply of oxygen and other nutrients at the injury site will end the inflammatory process from progression. Measuring the adequate oxygen delivery to a tissue depends on sufficient oxygen content in arterial blood and blood flow to the tissue utilizing pulse oximetry [23]. Also, acceptable oxygen delivery can assess susceptibility to injury and tissue damage [24].

There is a growing concern about the synergistic regeneration of bone and correlated vasculature, nerves, cartilage, tendon, ligaments, and muscles. Therefore, designing examination methods intraoperatively would open new doors to repairing and regenerating tissues after SCI [25]. Monitoring tissue oxygenation can evaluate both tissue oxygenation and microcirculation beyond macro-hemodynamic measurements [26]. Low blood flow at the injury site will cause hypoxemia (abnormally low oxygen level in tissue

(<60 mmHg)) [27]. Low arterial blood oxygenation indicates a reduction in oxygen delivery due to vasospasm or hypotension or inadequate blood perfusion [28]. Measuring the spinal blood oxygen saturation can provide information about blood perfusion [29]. Measuring the spinal oxygen saturation in dark conditions is done because the light can affect the examination results. However, this method is based on light transmission in the near-infrared spectrum across tissue and its absorption [30]. Measuring spinal blood oxygen saturation can help the clinician justify the change in the treatment method to influence potential deficiencies and the requirement of other procedures in every patient [6, 31].

Before decompression, we found that the oxygen saturation of spinal cord tissue was low, which may be caused by high ISP and low SCPP. We emphasized the need for surgical decompression and other interventions. There are many ways to determine decompression during operation, the level of decompression, cerebrospinal fluid pulsation, or blood circulation recovery. However, laminectomy decompression is empirical and cannot accurately determine the CSF circulation recovery. Simultaneously, these methods cannot directly reflect the situation of spinal cord tissue oxygen saturation. The significance of intraoperative monitoring of spinal cord oxygen saturation is as follows: direct measurement of the spinal cord blood oxygen saturation can indirectly reflect the blood perfusion pressure; whether the decompression is complete or not and the oxygen saturation of spinal cord tissue is restored or improved can be determined, indicating that decompression is effective; and guidance on whether to do a wider range of lamina decompression or more thorough dural incision decompression and other operations can be given.

Moreover, performing further procedure in case of abnormal oxygen saturation can affect the specific physiological deficit and hasten the healing of the injury. This can also make medical care decisions more personal and tailored to each case, minimizing unnecessary interventions and making patient follow-up easier. Also, this strategy can increase the chances of improving the current medical practice. Further monitoring tissue oxygenation can evaluate blood perfusion and confirm whether laminectomy is enough or the patient will need further durotomy. If the oxygen saturation is in the normal range, laminectomy will be enough to achieve significant outcomes, while if it is abnormal, then bony decompression is not enough to reduce ISP and improve the SCPP, requiring durotomy.

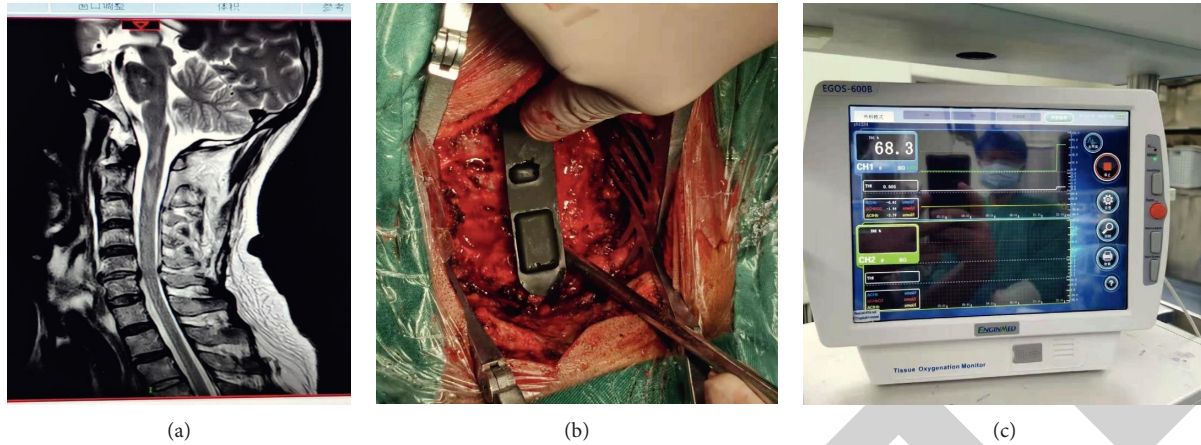


FIGURE 3: (a) An old male admitted to hospital for “head injured by a heavy object” (ASIA grade A) (diagnosis: spinal cord injury and cervical spine fracture). (b) Oxygen saturation probe was inserted after C2–C6 lamina decompression and internal fixation. (c) Spinal blood oxygen saturation was in a normal range (68.3%). Postoperative ASIA grade improved to grade D.

Oxygen saturation measurement equipment can be used for

- (1) Monitoring the oxygen metabolism of the patient’s spine tissue.
- (2) Assisting the clinical realization of noninvasive and continuous monitoring of spinal oxygen saturation.
- (3) Early detection of hypoxia and ischemia in the patient’s spine tissue.
- (4) Judging short-term neurodevelopmental prognosis.
- (5) Evaluation of the safety and effectiveness of the clinical treatment plan.
- (6) Assisting in evaluating the degree of SCI.

In these cases, the oxygen saturation was normal. We did not perform dural decompression as the SCPP improved at the injury site. Many factors affect the prognosis of the SCI, including severity of the injury, time of surgery, hematoma, and edema. Measuring spinal blood oxygen saturation is essential to confirm the improvement of SCPP after laminectomy. We failed to measure the oxygen saturation in one patient because the patient was thin and short, and we could not make the incision longer than 12 cm, so the space for the probe was not enough.

5. Limitations

This study has some limitations. First, this study evaluated a limited and small number of cases. Eight patients were involved and underwent spinal blood oxygen saturation measurement, and one of them failed the oxygen measurement. Second, spinal blood oxygen saturation was within normal range in all patients after laminectomy. Third, we chose only the patients with cervical abnormalities. In addition, the follow-up time is short for confirming the improvement after laminectomy. This is to say, more oxygen saturation measurement cases are thus needed.

6. Conclusion

Treating SCI patients with bony decompression can achieve significant outcomes but not in all cases. Decreasing the pressure around the injured spinal cord and increasing the blood flow are the mainstay interventions to achieving better results. Detecting the improvement of SCPP at the injury site is essential to confirm the benefit of laminectomy. Spinal blood oxygen saturation measurement under dark conditions might help in deciding whether to perform further procedures (durectomy with duroplasty or extensive laminectomy) or not. By monitoring spinal cord oxygen saturation in real time throughout the operation, it is possible to monitor the spinal cord perfusion pressure.

Data Availability

The patients’ history and other 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

We would like to thank the medical team of the Orthopedic Department, Union Hospital, who contributed to this study by providing ideas and helping in editing and revising this manuscript.

References

- [1] M. G. Fehlings, C. H. Tator, and R. D. Linden, “The relationships among the severity of spinal cord injury, motor and somatosensory evoked potentials and spinal cord blood flow,” *Electroencephalography and Clinical Neurophysiology: Evoked Potentials Section*, vol. 74, no. 4, pp. 241–259, 1989.
- [2] C. A. Oyibo, “Secondary injury mechanisms in traumatic spinal cord injury: a nugget of this multiply cascade,” *Acta*

Retraction

Retracted: Human-Computer Interaction Using Manual Hand Gestures in Real Time

Computational Intelligence and Neuroscience

Received 19 September 2023; Accepted 19 September 2023; Published 20 September 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] M. Alsaffar, A. Alshammari, G. Alshammari et al., "Human-Computer Interaction Using Manual Hand Gestures in Real Time," *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 6972192, 5 pages, 2021.

Research Article

Human-Computer Interaction Using Manual Hand Gestures in Real Time

Mohammad Alsaffar ¹, Abdullah Alshammari ¹, Gharbi Alshammari ¹,
Tariq S Almurayziq ¹, Saud Aljaloud ¹, Dhahi Alshammari,¹ and Assaye Belay ²

¹University of Ha'il, College of Computer Science and Engineering, Department of Computer Science and Information, Ha'il, Saudi Arabia

²Department of Statistics, Mizan-Tepi University, Tepi, Ethiopia

Correspondence should be addressed to Assaye Belay; assaye@mtu.edu.et

Received 5 November 2021; Accepted 16 November 2021; Published 28 November 2021

Academic Editor: Deepika Koundal

Copyright © 2021 Mohammad Alsaffar 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 paper describes the construction of an electronic system that can recognise twelve manual motions made by an interlocutor with one of their hands in a situation with regulated lighting and background in real time. Hand rotations, translations, and scale changes in the camera plane are all supported by the implemented system. The system requires an Analog Devices ADSP BF-533 Ez-Kit Lite evaluation card. As a last stage in the development process, displaying a letter associated with a recognized gesture is advised. However, a visual representation of the suggested algorithm may be found in the visual toolbox of a personal computer. Individuals who are deaf or hard of hearing will communicate with the general population thanks to new technology that connects them to computers. This technology is being used to create new applications.

1. Introduction

The word gesture has its origin in the Latin “gestus,” which refers to a form of nonverbal communication based on body language. Gestures are facial expressions or movements of the hands or any part of the body through which thoughts, feelings, or moods are manifested. Its purpose is to efficiently exchange a message between the person making the gesture and the person interpreting it [1]. Additionally, the Latin cestus is related to Greer, which in turn means “to carry out”; hence, the relationship of the word gesture with others such as “manage,” “gestate,” or “management”. In human-computer interaction (HCI), the use of gestures as a means of communication with computing devices is investigated [2]. The body is the primary agent in contact. Additionally, on some occasions, it is used to evaluate the user experience when facing certain interactions, for example, to estimate the emotions generated by an exchange according to the gestures made by the user.

A set of compelling techniques is available to deal with the recognition problem; however, their computational cost is usually very high, making them impossible to implement in real time using an embedded processor.

The technique proposed by Viola and Jones is highlighted and, later, improved by Lienhart and Maydt, which is based on Haar's wavelets [3], which perform a multi-resolution analysis of the image. In addition, the OpenCV library [4] includes functions that allow finding the hand and face through AdaBoost-type parallel classifiers [5], which yield excellent results. Other authors use morphological techniques such as skeletonization to identify organs in the human body [6, 7].

In order to find certain characteristics that classify each of the gestures for different individuals, it is proposed as a solution to carry out a morphological analysis of the image. In this way, a new alphabet is established. Each finger of the hand represents a bit, establishing a set of highly differentiable gestures and a problem of binary nature, which is addressed through morphological analysis of the image.

Finally, the viability and efficiency of the developed algorithm is demonstrated, and quite good processing times are obtained in the Blackfin 533 processor (ADSP BF-533) [8], using the EZ-Kit Lite evaluation card from Analog Devices.

2. Review of Literature

The implemented system executes the processing of a video sequence, in which a person is wearing a dark long-sleeved shirt gesturing in the foreground, and a camera captures the image of him under controlled lighting and background conditions [5]. The block diagram of the system is shown in Figure 1. The image is captured from a video sequence in the first stage; after that, the region in which the gestural interlocutor's hand is located is determined in the segmentation stage (region of the image to be processed). A thinning process is performed in the region of interest to limit the amount of processed information and allow the recognition stage to be successful [9]. Once the image is thinned, points of interest are identified, whose position concerning the centre of mass of the hand allow its representation in a vector, whose dimension is equal to the number of fingers. Each of its components has information on the inclination of the finger concerning the inclination of the forearm. Finally, using the mean square error, it is decided if the vector is sufficiently similar to any of the base vectors established in a training stage before the system's operation.

2.1. Alphabet. In the framework of this work, a new alphabet is proposed, based on the number of fingers and their location on the hand; this gives the flexibility of obtaining a fairly broad set of gestures, with reasonably well-defined structural differences. Likewise, the objective is to identify the presence of the finger and its location to the center of mass of the hand, that is, this new alphabet is based on the modelling of the fingers of the hand as binary inputs to the recognition system and is considered thumb as the most significant bit. The symbols are generated in this work; the letter A is represented in binary terms by 10,000 because the gesture only presents the thumb. By designating a hand in a binary way, thirty-two gestures can be obtained, with the future possibility of expanding the alphabet, to approximately sixty-four gestures (using the two faces of the hand) and more than two thousand gestures with the use of the two hands [9].

2.2. System Training. The training phase is a stage in the system's design to establish the base vectors; these contain information on each of the gestures for which the system will respond, being a small database of vectorization of images corresponding to valid gestures. Similarly, the success of the recognition depends on the base vectors, which is why they are established through a series of tests and statistical analysis of the results obtained by applying the algorithm developed in different interlocutors. Suppose the data corresponding to several individuals are averaged. In that case, the base vectors can be defined, and when samples are taken

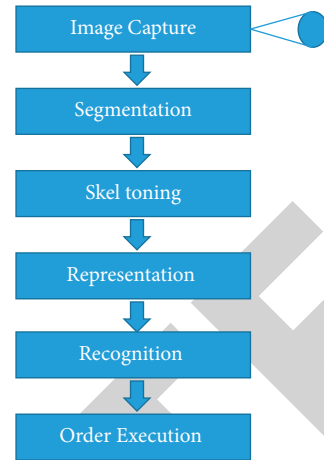


FIGURE 1: Block diagram of the system.

from a significant population, it is possible to develop a functional system for the population, in general [10].

2.3. Segmentation. Through background and controlled lighting, the region corresponding to the skin turns out to be the brightest. In this way, if luminance information is used, those pixels that exceed a set threshold are considered skin and must be taken into account for further analysis. If good lighting and a sufficiently opaque background are guaranteed, a threshold can be set at which good segmentation can be achieved [11].

2.4. Region of Interest (ROI). At this stage of the process, a region is established that must contain only the segmented hand to guarantee subsequent recognition, that is a region made by an interlocutor person with one of the hands in a situation with regulated lighting and background in real-time for the hand rotations, translations, and scale changes in the camera plane is supported by the implemented system, and other objects may appear; the system must be able to locate these objects, which turn out to be noise for the application, and filter them. With this in mind, the image must be processed to establish a region of interest (ROI) [12]. Setting the ROI reduces the area of the image over which to search for the target, thereby optimizing the process.

2.5. Skeletonization. To reduce the amount of information to be processed while preserving the topological distribution of the hands, a morphological operation can be carried out on the region of interest, such as thinning [13].

Thinning removes redundant information, which produces a more superficial image, reduces memory access time and space, and facilitates the extraction of topological features from the region of interest. The result of the thinning process of the segmented image must maintain certain properties to allow correct conservation of the topological characteristics of a determined gesture and allow a correct future recognition. The morphological operation in question must ensure that the resulting image is one pixel wide; this makes it much easier to find the branches that correspond to

pixels with more than two neighbours of interest and the terminal points that correspond to pixels with only one neighbour of interest. In this work, two thinning algorithms were evaluated to determine the appropriate method for real-time recognition compliance. Finally, the algorithm and the Medial Axis Transform (MAT) [14] were implemented.

2.6. Skeletonized Image Filtering. Before obtaining the final points, it is necessary to conduct a cleaning process of the resulting thinned image. The distances of the endpoints obtained from the thinned image relative to the centre of mass of the segmented image are determined, and a threshold is established. In effect, the endpoints corresponding to the distances smaller than the said threshold are discarded.

To maintain the robustness of the system at the camera-user distance [16–20], the thinning image cleaning process establishes a threshold, which is a proportion of the most significant distance between one of the endpoints and the center of mass of the hand. Therefore, an adaptive threshold is established.

2.7. Representation and Recognition. To recognise a gesture in an image, a morphological analysis of the image is carried out in search of an appropriate vectorization of the hand that allows later recognition.

It is attempted to justify the choice of control points, which will be preponderant in extracting the topological characteristics of the hand, from an image thinned out, after the image has been thinned—based on a formal theoretical support. From a thinned image, we want to find the most suitable control points for representing the curve resulting from the thinning process, bearing in mind that limiting their number is important for fulfilling the objective of operations in real-time processing to be carried out [21–29]. In addition to the centre of mass of the segmented hand, the endpoints of the thinned image are chosen to represent the idea of the geometry that it contains the points control to be analyzed because they provide key information of the topological structure of the hand. The recognition process is based on finding the angles of the fingers and forearm (endpoints) concerning a reference point. The center of mass of the interlocutor's segmented hand is used as a reference (Figure 2).

When the hand is fully vertical, the angle of the forearm to the origin is ideally 270 degrees, and the angle between the thumb and forearm is slightly greater than 90 degrees, as shown in Figure 3.

When the hand is rotated in the plane, the angle of the forearm concerning the origin changes, as expected. The angle difference between the thumb and the forearm, on the other hand, is still slightly greater than 90 degrees, and the angle difference between the forearm and each of the fingers remains constant.

The system generates a vector with the angle differences, which then compares to some base vectors established during the training phase, using the quadratic error criterion, to see if the calculated vector is sufficiently similar to

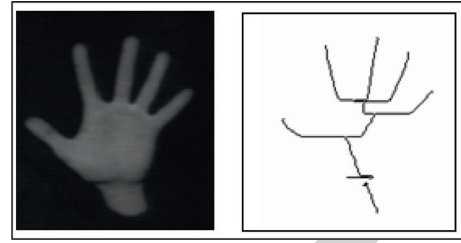


FIGURE 2: Endpoints of the skeleton. * At best, they correspond to the fingers or the endpoint of the forearm.

any of the vectors stored in memory, and uses this decision to recognise each of the gestures. These vectors will be arrangements with length equal to the number of fingers in each gesture, so the maximum length of one of these base vectors is five (corresponding to the five angles between the forearm and each of the five fingers). Each of the gestures for which the system was trained has at least one vector. We use angles to make the system resistant to translations and scale changes because angles are based on length relationships that will remain constant as long as the objective is in the camera's plane. To find each of the angles, the centre of mass is used as a reference point, which can vary as the interlocutor gestures. To achieve a more static center of mass, it is possible to discover more of the forearm of the interlocutor, which results in the angles between the fingers being more similar, with a greater probability of error. After a series of tests, it was established that the optimum point in which the sleeve should be left is approximately 3 cm below the interlocutor's hand.

3. Development on the Analog Devices ADSP BF-533 Processor

The implementation of the system is oriented in real time, using a dedicated processor, which allows portable applications. The development of the system mainly used the ADV7183 video encoder, the parallel peripheral interface (PPI), the DMA controller, and the asynchronous memory SDRAM. The PPI together with the DMA allows implementation of a subsampling of the image exclusively with hardware [15]. This subsampling does not significantly affect the application's performance, but it does optimize the processing speed since memory accesses are reduced. The DMA is configured to generate an interrupt once the entire image has been stored in memory and interrupts the data transfer. In this way, a black and white image corresponding to the captured scene is stored in memory. The DMA interrupt routine corresponds to image processing. The DMA is enabled again to transfer another image after the image has been processed, and the process is repeated.

4. Evaluation of Results

To evaluate the recognition algorithm, a total of nineteen thousand two hundred images corresponding to different individuals gesturing were analyzed. Four efficiency aspects of the system were evaluated, each with 4,800 independent images, and the results are as follows:

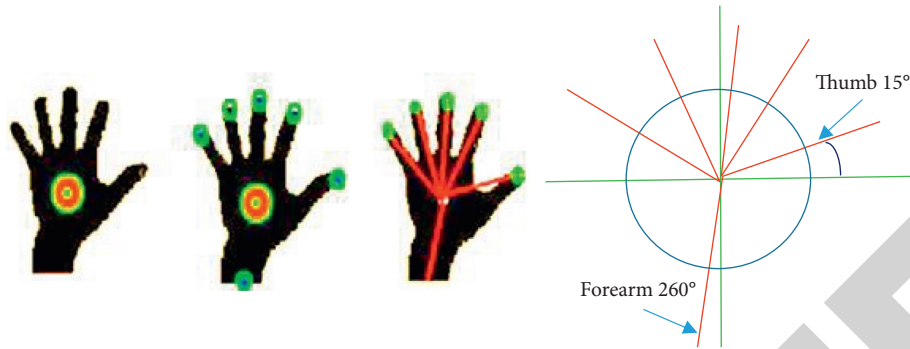


FIGURE 3: Angles relative to the forearm.

- (i) True hits 79.27%
- (ii) True rejections 99.50%
- (iii) False hits 0.27%
- (iv) False rejections 38.39%

It should be noted that the system recognizes 79% of the frames analyzed, which is very high if one takes into account that twenty-five images are processed in one second. The first stage of image processing involves locating the hand within the image, and the second stage involves recognising the object. The ROI fixation process, which determines the location of the hand in the image, is a much more computationally expensive process than the recognition process; in this process, a signal corresponding to the programmable flag of the Analog Devices processor is used to determine the processing time (half period of the signal corresponds to image processing).

In this way, the decision is made to implement an ROI location for every hundred processes. In effect, the region of interest is established and the subsequent one hundred surveys are carried out on this region. Finally, the ROI is refreshed, a greater number of recognitions is obtained in a given time. The thinning algorithm implemented in the DSP was the MAT because the processing time (35 ms) turns out to be between five and six times less than the time achieved with the Shang Zhang thinning algorithm (150 ms) [16]. A more robust system is developed in the Visual C++ programming environment than the one implemented in the development board. It is possible to constantly determine the region of interest recursively without affecting its operation in real time. The hand location function with recursive algorithms turns out to be optimal concerning its nonrecursive version in terms of time; however, in the evaluation card, due to the large number of iterations involved, the nonrecursive version of the location algorithm is implemented.

When implementing a recursive function, the processor must save the context in each iteration and considering that the number of iterations is proportional to the number of pixels analyzed in the image; it turns out to be a drawback due to space limitations, in fast memory, in which context can be stored.

5. Conclusion

An efficient tool was obtained that allows communication between a user and a machine, opening the possibility of

controlling it remotely and in real time. In the same way, it opens the possibility of managing ports and other peripherals of the personal computer, allowing future developments focused on enabling teleconferences guided by deaf-mute people. Allowing said population to limit their isolation is then envisaged to interact with people who are alien to the implemented language through a machine that synthesizes a sound or generates a text. In a future improvement of the system, it is proposed to work with colour spaces with which it could be possible to work with any background.

Data Availability

The data underlying the results presented in the study are available within the manuscript.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this study.

References

- [1] K. Cooperrider, N. Abner, and S. Goldin-Meadow, "The palm-up puzzle: meanings and origins of a widespread form in gesture and sign," *Frontiers in Communication*, vol. 3, 2018.
- [2] W. Wing Kwong Chung, X. Xinyu Wu, and Y. Yangsheng Xu, "A realtime hand gesture recognition based on Haar wavelet representation," in *Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics*, pp. 336–341, Bangkok, Thailand, February 2009.
- [3] S. J. Abdallah and G. Khalid Ouda, "Using integrated library management systems for the improvement of information services based on cloud computing," *Tikrit Journal of Pure Science*, vol. 25, no. 4, pp. 101–102, 2020, p.
- [4] V. Harini, V. Prahelika, I. Sneka, and P. Adlene Ebenezer, "Hand gesture recognition using OpenCv and Python," *New Trends in Computational Vision and Bio-inspired Computing*, vol. 174, pp. 1711–1719, 2020.
- [5] G.-W. Wang, C. Zhang, and J. Zhuang, "An application of classifier combination methods in hand gesture recognition," *Mathematical Problems in Engineering*, vol. 2012, Article ID 346951, 17 pages, 2012.
- [6] L. M. Thivagar, A. A. Hamad, and S. G. Ahmed, "Conforming dynamics in the metric spaces," *Journal of Information Science and Engineering*, vol. 36, no. 2, pp. 279–291, 2020, p.