Princess Sumaya University for Technology
جامعة الأميرة سميّة للتكنولوجيا

صندوق دعم البحث العلمي والابتكار
Scientific Research and Innovation Support Fund

Jordanian Journal of Computers and Information Technology

March 2025 VOLUME 11 NUMBER 01
ISSN 2415 - 1076 (Online)
ISSN 2413 - 9351 (Print)

JJCIT

PAGES	PAPERS
1 - 15	CURATING DATASETS TO ENHANCE SPYWARE CLASSIFICATION Mousumi Ahmed Mimi, Hu Ng and Timothy Tzen Vun Yap
16- 32	A NOVEL EVIDENTIAL COLLABORATIVE FILTERING FRAMEWORK BASED ON DISCOUNTING CONFLICTING PREFERENCES Khadidja Belmessous, Faouzi Sebbak and M'hamed Mataoui
33- 53	STATE-OF-THE-ART OF MACHINE LEARNING IN NEURO DEVELOPMENT DISORDER: A SYSTEMATIC REVIEW Lilian Lee Yen Wei, Ag Asri Ag Ibrahim and Rayner Alfred
54- 72	A NEW APPROACH COMBINING RSA AND ELGAMAL ALGORITHMS: ADVANCEMENTS IN ENCRYPTION AND DIGITAL SIGNATURES USING GAUSSIAN INTEGERS Yahia Awad, Douaa Jomaa, Yousuf Alkhezi and Ramiz Hindi
73- 84	OPTIMIZATION OF FALSE ALARM RATE AND MISDETECTION RATE FOR A DESIRED THRESHOLD VOLTAGE IN COOPERATIVE COMMUNICATION Satish Kumar Gannamaneni and Jibendu Sekhar Roy
85- 99	PRIVACY-AWARE MALARIA DETECTION: U-NET MODEL WITH K-ANONYMITY FOR CONFIDENTIAL IMAGE ANALYSIS Ghazala Hcini and Imen Jdey
100- 116	BLOCKCHAIN-BASED DEVICE AUTHENTICATION IN EDGE COMPUTING USING QUANTUM APPROACH Vinayak A. Telsang, Mahabaleshwar S. Kakkasageri and Anil D. Devangavi
117- 135	ENHANCING MICRO-EXPRESSION RECOGNITION: A NOVEL APPROACH WITH HYBRID ATTENTION-3DNET Budhi Irawan, Rinaldi Munir, Nugraha Priya Utama and Ayu Purwarianti

www.jjcit.org jjcit@psut.edu.jo

An International Peer-Reviewed Scientific Journal Financed by the Scientific Research and Innovation Support Fund

Jordanian Journal of Computers and Information Technology (JJCIT)

The Jordanian Journal of Computers and Information Technology (JJCIT) is an international journal that publishes original, high-quality and cutting edge research papers on all aspects and technologies in ICT fields.

JJCIT is hosted and published by Princess Sumaya University for Technology (PSUT) and supported by the Scientific Research Support Fund in Jordan. Researchers have the right to read, print, distribute, search, download, copy or link to the full text of articles. JJCIT permits reproduction as long as the source is acknowledged.

AIMS AND SCOPE

The JJCIT aims to publish the most current developments in the form of original articles as well as review articles in all areas of Telecommunications, Computer Engineering and Information Technology and make them available to researchers worldwide. The JJCIT focuses on topics including, but not limited to: Computer Engineering & Communication Networks, Computer Science & Information Systems and Information Technology and Applications.

INDEXING

JJCIT is indexed in:

EDITORIAL BOARD SUPPORT TEAM

JJCIT ADDRESS

WEBSITE: www.jjcit.org
EMAIL: jjcit@psut.edu.jo
ADDRESS: Princess Sumaya University for Technology, Khalil Saket Street, Al-Jubaiha
B.O. BOX: 1438 Amman 11941 Jordan
TELEPHONE: +962-6-5359949
FAX: +962-6-7295534

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

CURATING DATASETS TO ENHANCE SPYWARE CLASSIFICATION

Mousumi Ahmed Mimi[1], Hu Ng[1] and Timothy Tzen Vun Yap[2]

(Received: 22-Jun.-2024, Revised: 26-Aug.-2024, Accepted: 14-Sep.-2024)

ABSTRACT

Current methods for spyware classification lack effectiveness as well-structured datasets are typically absent, especially those with directionality properties in their set of features. In this particular research work, the efficacy of directionality properties for classification is explored, through engineered features from those on existing datasets. This study curates two datasets, Dataset A which includes features extracted from only single directional packet flows and Dataset B which includes those from bi-directional packet flows. Classification with these features is performed with selected classifiers, where SVM obtained the highest accuracy with 99.88% for Dataset A, while the highest accuracy went to RF, DT and XGBoost for Dataset B with 99.24%. Comparing these results with those from existing research work, the directional properties in these engineered features are able to provide improvements in terms of accuracy in classifying these spywares.

KEYWORDS

Datasets curation, Feature engineering, Packet analysis, Spyware classification.

1. INTRODUCTION

Cybercrimes are increasing due to careless use of online applications and technologies [1]-[2]. Users install various applications on their devices for different purposes, but many are not safe or secure as some disguise themselves as normal applications, such as spyware [3]. Spyware, a malicious software, is installed on the device, gathers sensitive information and transfers it to third parties without user consent [4]. It's very tricky and challenging to distinguish between spyware and legitimate applications, as it disguises itself as a legitimate application [5]. While significant research has been conducted on malware, the exploration of spyware has been overlooked. This creates a research gap for further investigation of the classification methods. Current spyware-classification methods have limitations in feature engineering based on directional properties, which hampers accurate classification. The accuracy of existing spyware classification is often hindered due to inadequate datasets and insufficient analysis for different classifiers, leading to overfit or underfit [6]. Additionally, users sometimes overlook security considerations when installing applications, creating opportunities for hackers. Cybercriminals create clones of popular software on untrustworthy sites with security vulnerabilities. Users carelessly install these cloned applications, allowing attackers to gain access to sensitive information [7]. Thus, this study makes several key contributions to the field of spyware classification.

1.1 Contribution

First, this study aims to enhance spyware classification by curating two new datasets. Dataset A involves annotation based on single-direction packet flow, while Dataset B involves bi-directional packet flow. Information about the packet flow is extracted from major static parameters, such as IP pairs, ports and protocols.

Secondly, feature engineering is applied to form dynamic features from static parameters like Total Forward (Fwd) Packet (Pkt), Total Backward (Bwd) Packet (Pkt), Flow Bytes per Second (Flow Bytes/s), Flow Packets per Second (Flow Pkt/s), … etc., derived from the annotated datasets. The goal is to identify significant features that can provide insights into the effectiveness of using packet-flow information in curating datasets and classifying spyware.

1. M. A. Mimi and H. Ng are with Faculty of Computing and Informatics, Multimedia University, 63100 Cyberjaya, Malaysia. Emails: 1221404218@student.mmu.edu.my and nghu@mmu.edu.my
2. T. T. V. Yap is with School of Mathematical and Computer Sciences, Heriot-Watt University Malaysia, 62200 Putrajaya, Malaysia. Email: timothy.yap@hw.ac.uk

Third, exploring machine-learning (ML) models for spyware classification from the two datasets, with the aim of improving accuracy. This involves applying Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), Naïve Bayes (NB) and Extreme Gradient Boosting (XGBoost) on curated datasets. Comparative analyses of the models trained on these datasets are then conducted to observe clustering patterns for the six different classes (five types of spyware with normal traffic) within each dataset, examining the extent of overlap between them.

The methodologies and the findings offer valuable contributions for enhancing the capabilities of Security Operation Center (SOC) system and Intrusion Detection System (IDS). The curated datasets, based on directional properties, benefit IDS by providing more accurate and contextually rich data for classifiers, that improves classification accuracy and reduces false positive rates for SOC environment. The engineered dynamic features, such as Flow Bytes/s, Flow Pkt/s, enable real-time threat analysis, allowing SOC and IDS to swiftly prioritize alerts for emerging threats. The validation of classification models, including SVM, NB, XGBoost, DT and RF, offer SOCs proven tools for more accurate spyware classification, which improves incident response reliability. Overall, the practical benefits of this research, such as enhanced detection accuracy and improved anti-spyware tools, strengthen the cybersecurity defences of SOCs, reducing the risk of unauthorized data access and privacy violations.

The remaining sections of the study are organized as follows: Section 2 provides a concise summary of prior research on dataset collection, feature extraction and engineering. Section 3 outlines the methodology, including details about dataset acquisition, spyware characteristics and data pre-processing. In Section 4, the proposed method is explained in detail with a focus on packet-flow direction and various approaches for curating Datasets A and B through feature engineering. This section also covers detailed methods for curating Datasets A and B while comparing them to the raw dataset. Section 5 discusses model construction, while Section 6 explores the results and their discussion. Lastly, concluding thoughts are presented in Section 7 to wrap up the document.

2. RESEARCH BACKGROUND

2.1 Spyware Dataset Collection

Cybersecurity is a widely discussed research topic. Researchers have criticized and discussed disciplines within cybersecurity, including spyware. Researchers have collected datasets to detect and characterize spyware.

Qabalin et al. [8] collected a dataset of five different spyware types - Flexispy, Mobilespy, uMobix, TheWispy and mSPY - by capturing packets using PCAPDroid [9]. Then, DT was applied to the dataset, achieving 79% accuracy for binary classification and 77% for multi-class classification.

Conti et al. [10] used the ASAINT (A Spy App Identification System based on Network Traffic) application to identify spyware apps and collected data. Packets were captured using Wireshark [11] and then network traffic was manually analyzed to distinguish between spyware apps and normal ones. The identified classes of spyware applications were cImg, cSms, mSpy and tSpy. Dropbox (DB) [12] and Google Foto (GF) [13] photo uploads were considered as normal applications. Four steps were applied: data collection, pre-processing, training and testing. The data distribution was adjusted using the Synthetic Minority Over-Sampling Technique (SMOTE) during pre-processing. The effectiveness of the datasets was assessed using RF, Linear Regression (LR) and K- Nearest Neighbour (KNN) algorithms, with RF ultimately achieving the best F1-score of 0.85.

M. Naser and Q. A. Al-Haija [14] utilized the Android Spyware-2022 dataset [8] to identify Android spyware, focusing on two out of five spyware classes: MobileSPY and FlexiSPY. After pre-processing the data by removing null and duplicate entries, they analyzed Source IP, Destination IP, Source Port, Destination Port, Duration and Protocol. The testing involved the application of a Fine Decision Tree (FDT), resulting in a 98% accuracy rate.

Noetzold et al. [15] implemented integrated spyware to monitor workplace computers. Integrated spyware represented the utilization of spyware techniques as a fundamental component within the design and functionality of a workplace computer-monitoring solution. The spyware initially sent harmful messages to a Twitter account developed using Python before being applied to the computer. Then, this spyware computer was connected to the workplace computer through an Application

Programming Interface (API) gateway. Subsequently, hate messages were sent from the spyware computer to the workplace one, triggering alerts generated by the API gateway which was connected to a relational database for storing information. Data pre-processing involved the use of normalization and classification techniques to differentiate between hateful and non-hateful speeches. Furthermore, LR, SVM and NB algorithms were utilized to assess prediction validity. NB demonstrated superior accuracy at 80%.

Pierazzi et al. [16] utilized the VirusTotal website [17] to collect spyware. Five types of spyware were identified: HeHe, UaPush, AceCard, Pincer and USBCleaver. Twenty-five features were extracted from each spyware, including file size, permission for sending short-message service, author information, permission for checking phone state, permission to write messages and permission to reboot the system, among others. The Ensemble Late Fusion (ELF) method identified these features as crucial in distinguishing spyware from normal applications. This involved comparing the characteristics of each feature with those of a normal application. Histograms were used to illustrate the variance between spyware and regular applications. Differentiating spyware from normal applications using RF with ELF resulted in an impressive F1-score of 0.96.

Mahesh et al. [18] utilized a Particle Swarm Optimization (PSO) algorithm with Artificial Neural Network (ANN) to improve the prediction of spyware detection. A benchmarked dataset of malware [19] was obtained for the study conducted by Kaggle [20]. Utilizing multi-objective PSO for data pre-processing, the features were then scaled using standard scaling. Additionally, a multi-layer perceptron was utilized along with the Jordan canonical form to remove less significant features and enhance accuracy. The ANN model was finally used to predict the accuracy of the proposed method, achieving an impressive 99% accuracy rate.

Zahan et al. [21] developed a benchmark dataset of malicious and benign software packages from NPM and PyPI to enhance malware-detection tools. The dataset was compiled from existing malicious databases and new malicious and neutral packages. They collected malicious packages from open-source datasets and an internal Socket benchmark and curated a set of neutral packages using manual annotation and automated scanning. The final MalwareBench dataset contained 20,792 samples, of which 6,659 were malicious.

A comprehensive analysis of the search results obtained through the adopted keyword-search approach has been conducted. The reviewed findings are summarized in Table 1.

Table 1. Summary of the reviewed literature in this area of study.

Author(s)	Dataset	Spyware Types	ML Model	Key Findings	Strengths	Weaknesses
Qabalin et al. [8]	Android Spyware-2022	Flexispy, Mobilespy, uMobix, TheWispy, mSPY	DT	Binary classification accuracy: 79.00%, Multi-class: 77.00%	Network-traffic analysis, dataset available	Limited to binary and multi-class classification, lower multi-class accuracy
Conti et al. [10]	ASAINT application, Wireshark	cImg, cSms, mSpy, tSpy	RF, LR, KNN	Best F1-score achieved by RF: 0.85	Effective use of ASAINT	Manual network-traffic analysis
M. Naser and Q. A. Al-Haija [14]	Android Spyware-2022	MobileSPY, FlexiSPY	FDT	Accuracy: 98.00%	High accuracy with FDT	Limited to two spyware classes
Noetzold et al. [15]	Integrated spyware for workplace monitoring	Not specified	LR, SVM, NB	NB demonstrated superior accuracy: 80.00%	Innovative use of integrated spyware	Focused on workplace computers, not mobile spyware
Pierazzi et al. [16]	VirusTotal website	HeHe, UaPush, AceCard, Pincer, USBCleaver	ELF, RF	F1-score: 0.96	High F1-score, Effective feature extraction	Complex ELF method

2.2 Feature Extraction and Engineering

Feature extraction transforms raw data into numerical features that retain the original information, enabling effective processing and improved ML-model performance over direct application of algorithms. Feature engineering, a crucial element in successful ML research, involves data

presentation, refinement and pre-processing tasks. Poorly engineered features can adversely impact model predictions.

Zhang et al. [22] developed a low-cost feature-extraction method for deep learning-based malware detection. The approach involved monitoring API call behaviour, encoding heterogeneous information into homogeneous features using feature hashing and applying gated convolutional neural networks and Bi-Directional Long Short-Term Memory (Bi-LSTM) to capture sequential API call correlations. This yielded a 98.80% area under the ROC curve.

Gibert et al. [23] described a feature-extraction process that combined hand-crafted features from hexadecimal and assembly-language source codes, as well as deep features extracted using deep learning architectures. The hand-crafted features included metadata, byte unigrams, entropy statistics, Haralick features and local binary pattern features. The assembly-language features covered metadata, opcode unigrams, register features, symbol frequency, pixel intensity, API function calls, data define features, section features and miscellaneous features. Deep features were extracted from raw data, including grayscale image-based features, entropy-based features, opcode N-gram features and byte N-gram features. These features were then fused using an early fusion mechanism to create a joint representation, which was used to train a Gradient Boosting (GB) model for malware classification, achieving an accuracy of 99.81%.

Masabo et al. [24] developed a feature-engineering method to classify polymorphic malware (can transform into various forms). The researchers collected a dataset of 5 malware classes (API, Crypto, Locker, Zeus and Shadow brokers.), pre-processed the data and performed feature engineering to identify 11 top features. These included static analysis of portable executable files, packing techniques, file access and registry reading. The developed feature-engineering approach outperformed traditional ML methods (GB), achieving a 94% accuracy.

Nawaz et al. [25] proposed a system to classify Android malware using the Drebin dataset [26]. Static analysis focused on Android intents and permissions, while dynamic analysis utilized network requests and API calls. Apktool [27] was used to decompile and decode the APK files. Feature selection with Info Gain reduced the dimensionality of permissions, intents, API calls and network features. These features were extracted from the APK components and used to train ML classifiers, with RF and GB performing best on the permission features, achieving an F1-score of 0.98.

Jung et al. [28] utilized an APK file from the AndroZoo dataset [29], extracted information on API calls and permissions and generated a feature vector for each application. They applied feature-selection methods to choose the top 20 features from API calls and permissions. The authors then employed RF and grid search to establish optimal hyperparameters and the best accuracy of 96.95% was obtained using Gini importance with the RF model.

Low et al. [30] explored two feature-engineering methods, label encoding and evidence counting, for malware detection. The study involved four main steps: data pre-processing, feature selection, model construction and evaluation. Five malware classes (Advanced Persistent Threats (APT), Crypto, Zeus, Locker and Shadow Brokers) were extracted from the dataset. During pre-processing, data integration, cleaning and transformation were applied. Boruta was used for feature selection and several ML models were constructed, with the optimal parameters identified through grid search. The dataset was balanced using SMOTE. The results showed that RF provided better accuracy for label encoding at 91.34%, while LSTM achieved higher accuracy of 94.64% for evidence counting.

A comprehensive analysis of the search results obtained through the adopted keyword-search approach has been conducted. The reviewed findings are summarized in Table 2.

3. METHODOLOGY

The dataset is initially prepared through pre-processing, organizing the data and converting it into a Comma-Separated Values (CSV) format. Feature engineering is then conducted to choose significant features for classification-model performance. Subsequently, classification models are constructed and their results are recorded for evaluation purposes. Figure 1 illustrates the workflow of the methodology.

Table 2. Summary of the reviewed literature in this area of study.

Author(s)	Dataset	Spyware/ Malware Types	Features and Techniques	ML Model	Performance	Strengths	Weaknesses
Zhang et al. [22]	AV-TEST 2017	Various PE malware	Feature extraction using Cuckoo, Feature hashing, Multiple gated CNNs, Bi-LSTM	Multiple gated CNNs, Bi-LSTM	AUC: 98.80%	Effective deep learning for malware detection	Focused on PE files, not mobile apps
Gibert et al. [23]	Not specified	Malware (not specified)	Hand-crafted and deep features, Fusion mechanism, Joint representation of features from multiple modalities	XGBoost	Accuracy: 99.81%	Comprehensive feature extraction from multiple sources	Complex and time-consuming
Masabo et al. [24]	Malware Training Sets	API, Crypto, Locker, Zeus, Shadow brokers	Compute feature importance for feature engineering	KNN, Linear Discriminant Analysis (LDA), GB	Accuracy: 94.00%	Focus on polymorphic malware	Limited dataset, feature-engineering complexity
Nawaz et al. [25]	Drebin dataset	Android malware	Permissions and intents extraction, Network requests, API calls,	RF, NB, GB, Ada Boosting	F1 score: 0.98	High F1-scores with RF and GB	Dynamic-analysis complexity
Jung et al. [28]	AndroZoo dataset, static extraction of API calls	Not specified (general malware)	Gini importance-based method	RF	Accuracy: 96.95%	Effective feature-selection methods	Complexity in feature extraction
Low et al. [30]	Dataset provided by Ramili-2016	Advanced Persistent Threats (APT), Crypto, Zeus, Locker, Shadow Brokers	Label encoding and evidence counting	RF, DT, KNN, SVM, LSTM	Label encoding: 91.34%, Evidence counting: 94.64%	High accuracy with LSTM	High computational requirements

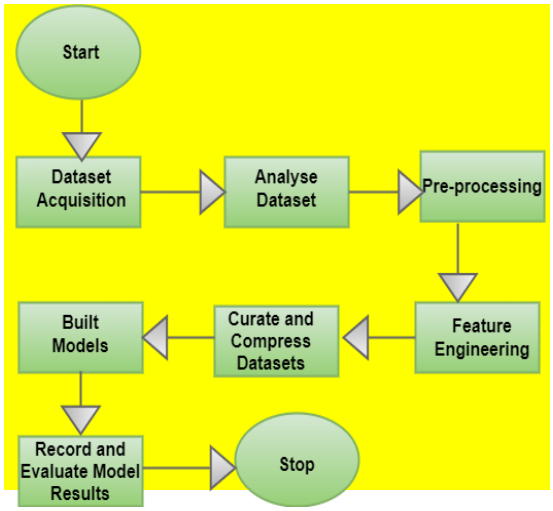


Figure 1. Methodology workflow.

3.1 Dataset Acquisition

The dataset used is called "Android Spyware-2022" [8]. The dataset was generated using PCAPDroid. It is a data-collection tool that can be installed on the Android operating system. The data consisted of five different spywares: FlexiSPY, MobileSPY, mSPY, TheWispy and uMobix; and one normal-traffic class which represents normal-smartphone traffic. Each row in the PCAP file represents a single packet. Analyzing the PCAP data involved extracting static information from each packet, such as Source IP, Destination IP, Source Port, Destination Port, Protocol type, Flags information, Acknowledgment Number and Message content. It also included recording Flow Duration (the time taken for a packet to transfer from source to destination), Packet Header Length and Packet Full Length. The information in the PCAP files is presented in Table 3.

Table 3. Available information in PCAP files.

File Name (.pcap)	Number of Packets	System Name	File Size (MB)	Data Tag
Normal_Traffic	1,04,914	Smart Phone Normal Traffic	78.81	Normal Traffic
FlexiSPY_Installation	19,793	FlexiSPY Inst	16.78	FlexiSPY Inst
FlexSPY_Traffic	35,433	FlexiSPY Traffic	22.32	FlexiSPY Traffic
Mspy Traffic- Part1	35,560	mSPY	25.94	mSPY Traffic
Mspy Traffic- Part2	20,537	mSPY	20.32	mSPY Traffic
mSPY Installation Process	12,976	mSPY	11.34	mSPY Inst
uMobix_Installation	17,312	uMobix	14.37	uMobix Inst
uMobix_Traffic	18,561	uMobix	16.28	uMobix Traffic
MobileSpy_Traffic	28,154	MobileSPY	12.76	MobileSPY Traffic
Mobilespy_Intallation_1	10,139	MobileSPY	8.41	MobileSPY Inst
TheWiSPY_Installation	58,223	TheWiSPY	53.24	TheWiSPY Inst
TheWISPY_Traffic	27,343	TheWiSPY	21.36	TheWiSPY Traffic

In this context, there are two packet types: "Installation" and "Traffic." The "Installation" type represents traffic data captured during the spyware-installation process, while the "Traffic" type represents spyware operation traffic data. Figure 2 illustrates the distribution of the six classes within the dataset. It shows that all five spyware classes overlap with the normal-traffic class. Utilizing the PCAP file information directly for classification may not be ideal, as features are not distinct for each class. Therefore, there is a need to curate new datasets with more distinct features for each class. To achieve this, prominent characteristics of each spyware must be identified and analyzed in the next sub-section.

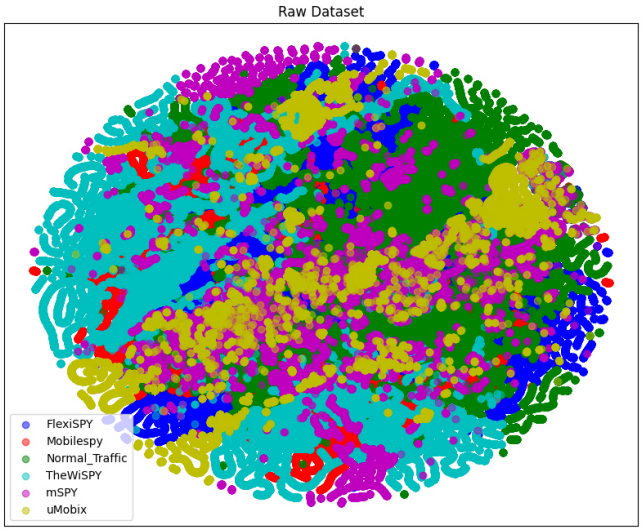


Figure 2. The distribution of the six classes.

3.2 Characteristics of Spyware Acquired from the Dataset

Table 4 presents a concise summary of each type of spyware. The Spying Scope represents different monitoring channels. The Platform highlights the language and framework used to develop spyware. The Upload represents how the data is transmitted to the Command and Control (C2C) server. Sniffing identifies sniffing strategies.

Based on the observations in Table 4, it is noted that each spyware shares similar characteristics in terms of spying scope, platform and sniffing features. This similarity will determine the next course of action for adapting the data pre-processing method.

3.3 Data Pre-processing

The process of converting every individual PCAP file into CSV format is explained in Algorithms 1 and 2. Each PCAP file represents a sample that contains information related to the corresponding spyware. Investigating the utilization of packet-flow direction in feature engineering is explored in the following section to minimize class overlap.

Table 4. Characteristics of spyware acquired from the dataset.

Spyware Classes	Spying Scope	Platform	Upload	Sniffing
FlexiSPY	a) Social-media applications b) Keylogger c) OS activity d) Update history e) Applications manifest f) Phone calls	Java	Periodic-based with fixed time interval	Event-based
MobileSPY	a) SIM Tracker b) Social-media applications c) Keylogger d) OS activity e) Update history. f) Applications manifest g) Phone calls	React Native, Java	Non-adjustable periodic	Event-based
TheWiSPY	a) SIM Tracker b) Social-media applications c)Keylogger d)OS activity e) Phone calls	React Native, Java	Adjustable periodic	Event-based
mSPY	a) Social-media applications b) Keylogger c) OS activity d) Update history. e) Applications manifest e) Phone calls	Java	Periodic-based with fixed time interval	Event-based
uMobix	a) Social-media applications b) Keylogger c) OS activity d) Update history e) Applications manifest f) Phone calls	Java	Adjustable periodic	Adjustable in terms of periodic or event-based

4. FEATURE ENGINEERING

Firstly, when observing the direction of the packet flow, two directional properties are apparent: single-direction and bi-directional. The direction of the packet flow is determined by the Source IP, Destination IP and Protocol. In this case, static features are extracted from the PCAP file including Source IP, Destination IP, Source Port, Destination Port, Protocol, Flow Duration and Packet Length.

The dynamic features include Total Forward Packets (Total Fwd Pkt), Total Backward Packets (Total Bwd Pkt), Total Length of Forward Packets, Total Length of Backward Packets, Flow Bytes per Second (bytes/s) and Flow Packets per Second (pkt/s), as well as the statistical values, such as minimum, maximum, average and standard deviation values.

Total forward and backward packets, along with the total length of forward and backward packets, are derived from the direction of packet flow and packet length. Additionally, flow bytes per second (bytes/s) and flow packets per second (pkt/s) are obtained from the direction of packet flow, flow duration and packet length.

After feature-engineering processes, the next step involves curating the two datasets: Dataset A and Dataset B.

4.1 Method for Developing Dataset A

A single-direction packet flow is utilized to curate Dataset A. It examines the IP pairs and Protocol for each row. When the Source IP, Destination IP and Protocol remain constant across two or more consecutive rows, statical measures are calculated to form new features. The features are presented in Table 5. These consecutive rows are combined into a single group, which represents a single-directional packet flow. However, if the Source IP and Destination IP remain unchanged for consecutive rows but the Protocol differs, they cannot be considered part of the same group. They will be considered as part of a different packet flow.

Table 5. Detail description of features after feature engineering.

Feature Name	Feature-engineering Process
Source (Src) IP	From PCAP file
Destination (Dst) IP	From PCAP file
Src Port	From PCAP file
Dst Port	From PCAP file
Protocol	From PCAP file
Flow Duration	Subtract the current flow start time from the last flow end time.
Total Forward (Fwd) Packets	Sum the forward packets.
Total Backward (Bwd) Packets	Sum the backward packets
Total Length of Fwd Packet	Sum the forward packet length
Total Length of Bwd Packet	Sum the backward packet length
Fwd Packet Length Min	Minimum forward packet length
Fwd Packet Length Max	Maximum forward packet length
Fwd Packet Length Mean	Average forward packet length per flow
Fwd Packet Length Std	Standard deviation of forward packet length
Bwd Packet Length Min	Minimum backward packet length
Bwd Packet Length Max	Maximum backward packet length
Bwd Packet Length Mean	Average backward packet length per flow
Bwd Packet Length Std	Standard deviation of backward packet length
Flow Bytes/s	Byte rate in a flow
Flow Pkt/s	Packet rate in a flow

Algorithm 1: Generating Dataset A. Here, p represents the previous row and n represents the next row.

Algorithm 1: Dataset A

```
1.  pcap = read (open ( pcap file))
2.  Require: ip, protocol
3.  for row number in row do
4.      if (p.IP pairs = = n.IP pairs && p.protocol = = n.protocol) then
5.          calculate feature value
6.      else
7.          move to the next row
8.      end if
9.  end for
10. function writeCsv(data, outputFile):
11. processedData = processPcap(pcapFile)
12. writeCsv(processedData, outputFile)
```

4.2 Method for Developing Dataset B

For Dataset B, the process is like Dataset A (Sub-section 4.1) with the exception of utilizing a bi-directional packet flow instead of a single-direction packet flow. Figure 3 illustrates this bi-directional flow. The Source IP in Row-1 and Row-2 subsequently becomes the Destination IP in Row-3 and Row-4. This process subsequently occurs also in Row-5. A similar process occurs for the Destination IPs as well.

Algorithm 2: Dataset B

```
1.  pcap = read (open ( pcap file))
2.  Require: ip, protocol
3.  for row number in row do
4.      if (p.source.ip ||p.destination.ip==n.source.ip||n.destination.ip && p.protocol== n.protocol) then
5.          calculate feature value
6.      else
7.          move to the next row
8.      end if
9.  end for
10. function writeCsv(data, outputFile):
11. processedData = processPcap(pcapFile)
12. writeCsv(processedData, outputFile)
```

4.3 Comparison of the Datasets

After curating Datasets A and B, Dataset A consists of 3,928 rows and Dataset B comprises 2,573 rows, while the raw dataset contains a total of 386,963 rows. Tables 6, 7 and 8 present the features and sample values for the raw dataset, Dataset A and Dataset B. It is important to note that while Table 8

includes features related to bi-directional packets, each feature is listed with respect to one source and destination, which may obscure the indication of data-flow direction. To address this, Figure 3 in sub-section 4.3 illustrates how the data-flow direction is represented in the datasets, enhancing the clarity of the bi-directional packet features.

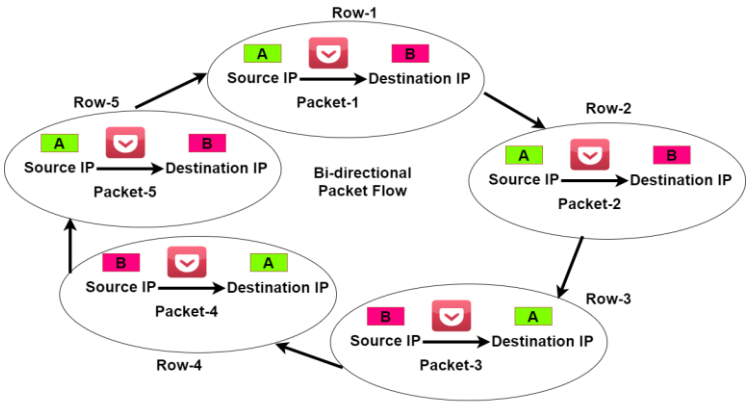


Figure 3. Bi-directional packet flow.

Table 6. Sample rows of the raw dataset.

Src IP	Dst IP	Protocol	Total Fwd. Packets	Total Bwd. Packets	Total Length of Fwd. Packet	Total Length of Bwd. Packet	Fwd. Packet Length Min.	Fwd. Packet Length Max.	Fwd. Packet Length Mean	Fwd. Packet Length Std.	Bwd. Packet Length Min.	Bwd. Packet Length Max.	Bwd. Packet Length Mean	Bwd. Packet Length Std.
10.215.173.1	161.117.185.166	TCP	1	0	60	0	60	60	60	0	0	0	0	0
10.215.173.1	157.240.195.54	TCP	1	0	60	0	60	60	60	0	0	0	0	0
8.8.8.8	10.215.173.1	DNS	0	1	0	48	0	0	0	0	48	48	48	0
8.8.8.8	10.215.173.1	DNS	0	1	0	40	0	0	0	0	40	40	40	0
10.215.173.1	157.240.195.54	TCP	0	1	0	44	0	0	0	0	44	44	44	0
157.240.195.54	10.215.173.1	TCP	1	0	40	0	40	40	40	0	0	0	0	0
10.215.173.1	104.21.81.103	TLSv1.2	1	0	43	0	43	43	43	0	0	0	0	0
157.240.195.54	10.215.173.1	UDP	0	1	0	40	0	0	0	0	40	40	40	0
142.250.200.243	10.215.173.1	TLSv1.3	0	1	0	44	0	0	0	0	44	44	44	0

Table 7. Sample rows of Dataset A.

Src IP	Dst IP	Protocol	Total Fwd. Packets	Total Bwd. Packets	Total Length of Fwd. Packet	Total Length of Bwd. Packet	Fwd. Packet Length Min.	Fwd. Packet Length Max.	Fwd. Packet Length Mean	Fwd. Packet Length Std.	Bwd. Packet Length Min.	Bwd. Packet Length Max.	Bwd. Packet Length Mean	Bwd. Packet Length Std.
10.215.173.1	161.117.185.166	TCP	23	0	6976	0	88	1472	303.30	390.05	0	0	0.00	0.00
10.215.173.1	157.240.195.54	TCP	20	0	1880	0	88	152	94.00	15.10	0	0	0.00	0.00
161.117.185.166	10.215.173.1	TLSv1.2	0	2	0	176	0	0	0.00	0.00	88	88	88.00	0.00
10.215.173.1	10.215.173.2	DNS	38	0	21492	0	88	1548	565.58	586.28	0	0	0.00	0.00
10.215.173.2	10.215.173.1	DNS	0	65	0	11408	0	0	0.00	0.00	116	436	175.51	75.25
10.215.173.1	157.240.196.60	TCP	72	0	82640	0	88	1548	1147.78	564.12	0	0	0.00	0.00
10.215.173.1	157.240.196.60	TLSv1.3	1	0	116	0	116	116	116.00	0.00	0	0	0.00	0.00
10.215.173.1	10.215.173.2	UDP	48	0	7592	0	100	1424	158.17	187.85	0	0	0.00	0.00
10.215.173.1	172.217.171.206	TLSv1.3	2	0	252	0	112	140	126.00	14.00	0	0	0.00	0.00

Figures 4 and 5 provide a comparison of the Datasets A and B and their respective values. As shown in Figure 2, there is an overlap in the values of all six classes within the raw dataset, making it challenging to differentiate between distinct clusters. Conversely, Datasets A and B (depicted in Figures 4 and 5) show minimal or no overlap among the classes, clearly distinguishing between them. These visual representations encompassed all features and depicted the distribution of the six classes.

Table 8. Sample rows of the Dataset B.

Src IP	Dst IP	Protocol	Total Fwd. Packets	Total Bwd. Packets	Total Length of Fwd. Packet	Total Length of Bwd. Packet	Fwd. Packet Length Min.	Fwd. Packet Length Max.	Fwd. Packet Length Mean	Fwd. Packet Length Std.	Bwd. Packet Length Min.	Bwd. Packet Length Max.	Bwd. Packet Length Mean	Bwd. Packet Length Std.
10.215.173.1	161.117.185.166	TCP	10	10	896	4484	88	96	89.60	3.20	88	1004	448.40	373.16
157.240.196.60	10.215.173.1	TCP	64	64	6100	79780	88	500	95.31	51.46	88	1548	1246.56	468.12
10.215.173.1	10.215.173.2	DNS	3	3	184	2800	56	64	61.33	3.77	44	1378	933.33	628.85
37.44.39.12	10.215.173.1	DNS	993	993	61816	1E+06	61	78	62.25	2.91	54	1378	1375.33	59.36
10.215.173.1	157.240.196.60	TCP	63	63	80404	7456	88	1548	1276.25	476.01	88	500	118.35	102.80
172.217.171.206	10.215.173.1	TLSv1.2	10	10	1580	8568	88	648	158.00	168.58	88	1548	856.80	628.06
74.125.206.188	10.215.173.1	TCP	14	14	2464	2748	88	488	176.00	135.51	88	736	196.29	170.65
10.215.173.1	74.125.206.188	TCP	13	13	1152	4080	88	96	88.62	2.13	108	488	313.85	144.58
10.215.173.1	10.215.173.2	UDP	22	22	3356	2520	124	320	152.55	40.96	104	124	114.55	4.76

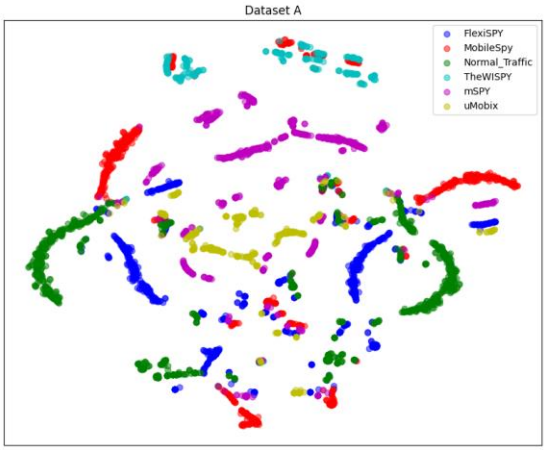


Figure 4. Dataset A.

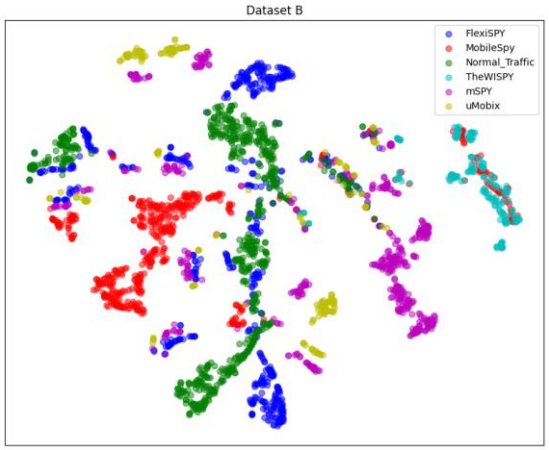


Figure 5. Dataset B.

5. Model Construction

To assess whether a dataset is appropriate for a detection model, thorough testing and analysis are crucial. The aim is to identify the most suitable ML model that aligns with the features of the curated datasets. Through analyzing the confusion-matrix values from different ML models, valuable information about the curated datasets' efficiency and performance will be obtained, enabling well-informed decisions regarding their use in detection tasks.

5.1 Classifiers

To determine the most suitable ML model, this study utilizes NB, XGBoost, RF, DT and SVM with the Radial Basis Function (RBF) kernel. DTs are recognized for their simplicity and interpretability, as they iteratively divide the data based on feature values to create a tree-like structure for classification. They can effectively handle both numerical and categorical data, making them well-suited for datasets with mixed-data types. RF improves upon DT by combining multiple trees, which enhances accuracy and reduces overfitting. SVM with RBF kernel and employing the One- *vs*-Rest (OVR) strategy excels in managing high-dimensional data by identifying optimal decision boundaries for classification. NB is particularly effective for text and categorical data due to its reliance on independence between features. Lastly, XGBoost combines gradient boosting with regularized learning to perform well in diverse datasets as an ensemble model. This research applied these traditional methods because it focused on feature engineering.

6. Results And Discussion

6.1 Results

The study evaluated classification performance using various metrics and different training-testing

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

dataset splits. The 70-30 split yielded the best results, which were consistent across different approaches and datasets, as illustrated in Tables 9 and 10. This consistency can be attributed to the rigorous curation of the datasets, as described in sub-sections 4.1 and 4.2. The similar results suggested the classification model's robustness across dataset configurations. A thorough review has been conducted to ensure the accuracy and reliability of the dataset-preparation and model-evaluation methods.

Table 9. Results for accuracy, precision, recall and F1-score for Dataset A.

ML Models	Accuracy (%)	Precision (%)	Recall (%)	F-1 Score (%)
DT	97.97	97.97	97.97	97.97
RF	97.97	97.97	97.97	97.97
XGBoost	96.38	96.43	96.38	96.39
SVM	99.88	99.88	99.88	99.88
NB	97.02	97.02	97.02	97.02

Table 10. Results for accuracy, precision, recall and F1-score for Dataset B.

ML Models	Accuracy (%)	Precision (%)	Recall (%)	F-1 Score (%)
DT	99.24	99.25	99.24	99.24
RF	99.24	99.25	99.24	99.24
XGBoost	99.24	99.25	99.24	99.24
SVM	99.12	99.12	99.12	99.12
NB	99.02	99.02	99.02	99.02

Figures from 6 to 9 show the difference in the evaluation metrics (including accuracy, precision, recall and F1-score) between Datasets A and B. For Dataset A, SVM demonstrated superior performance in accuracy, precision, recall and F1-score. Conversely, DT, RF and XGBoost exhibited better performance on these metrics for Dataset B.

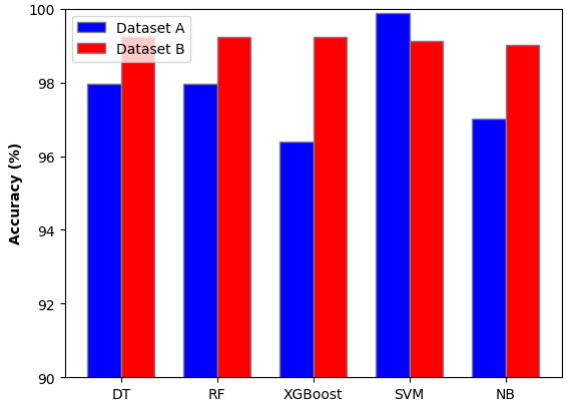


Figure 6. Evaluation of the accuracy of ML models in the context of both Dataset A and Dataset B.

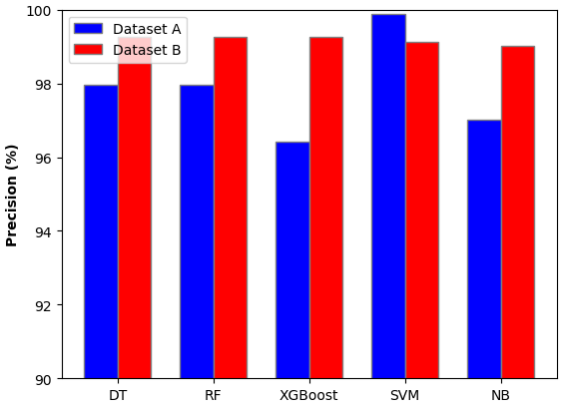


Figure 7. Evaluation of the precision of ML models in the context of both Dataset A and Dataset B.

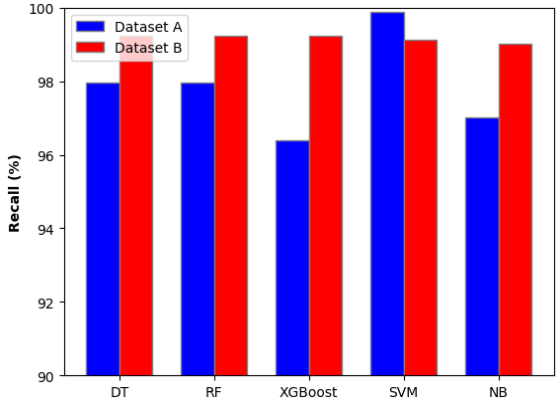


Figure 8. Evaluation of the recall of ML models in the context of both Dataset A and Dataset B.

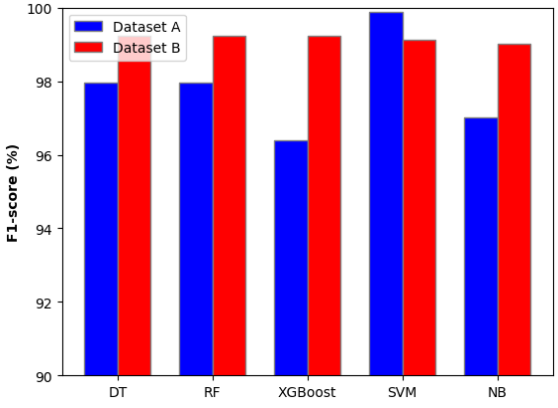


Figure 9. Evaluation of the F1-score of ML models in the context of both Dataset A and Dataset B.

6.2 Discussion

The comparative evaluation aimed to enhance the classification of spyware. Table 11 presents a comparison between the outcomes of earlier studies and those of this research, all utilized on the same dataset. M. Naser and Q. A. Al-Haija [14] categorized two types of spyware: FlexiSPY and MobileSPY, while [8] classified five varieties: FlexiSPY, MobileSPY, TheWiSPY mSPY and uMobix without extracting new features from the datasets used in their research. Feature engineering involved extracting new features based on packet-flow direction, IP pairs and Protocol information to improve spyware classification by identifying five distinct classes.

Table 11. Comparison of classification accuracy.

Research Work	ML Scheme	Dataset	Spyware Classes	Accuracy
Qabalin et al. [8]	DT	Android Spyware-2022 [8]	5	79.00%
M. Naser and Q. A. Al-Haija [14]	FDT	Android Spyware-2022 [8]	2	98.00%
T. N. AlMasri and M. A. N. AlDalaien. [31]	RF	Android Spyware-2022 [8]	5	92.00%
This Work	SVM	Dataset A	5	99.88%
	RF, DT, XGBoost	Dataset B	5	99.24%

After the curation process, both Datasets A and B showed improved performance compared to previous studies in identifying spyware. M. Naser and Q. A. Al-Haija [14] achieved strong results in spyware identification with only two spyware classes, but positive outcomes were achieved by performing well across five different spyware classes. By having more classes, the model had to comprehend more detailed patterns and characteristics associated with each type of spyware. This higher level of detail enabled the model to better discriminate, leading to improved accuracy. Other researchers utilized features directly from the raw dataset without considering directional properties, while this research focused on features utilizing packet flow direction. This approach led to the creation of more pertinent features related to different types of spyware behaviours.

The engineered features of curated Datasets A and B were highly effective, because they incorporated directional properties. These properties facilitated a deeper understanding of network behaviour, aiding in anomaly detection, optimization support and enhancing network analysis for modelling objectives and spyware classification [32]. The data-flow direction enhanced anomaly detection and network-performance optimization by mapping expected sequences, standard frequency, volume of data transmissions and common paths taken by data packets. However, some of these features, although characterized, are relatively trivial. The major contribution of this study is curating two new datasets based on directional properties. Though both datasets have the same features, the values of each feature for both datasets were different because of directional properties. Engineered features included directional properties within the network data for in-depth insight into behavioural baselines, protocol analysis, traffic volume, directionality and time-based patterns. Because of the complex nature of spyware, this method was crucial. By understanding the data flow direction, accurate behaviour modeling could distinguish between normal and suspicious activities with greater precision. The detailed analysis of packet-flow direction and its impact on spyware classification is the novelty of this study. Unlike previous studies, this study insights into data movement were provided by directional properties, which helped identify unusual patterns that signify potential threats or inefficiencies. Unusual data paths represented the potential threats, like exfiltration by spyware. The model also identified network inefficiencies, such as sub-optimal routing or congestion. This novel approach focuses on these directional properties, which provides the model clearer view of network dynamics for early performance optimization and threat detection [33].

The derived features: Total Fwd Pkt and Total Bwd Pkt from IP pairs, Ports and Protocol; provided valuable insights into the source and destination of network traffic with communication protocols (e.g. TCP, UDP). This improved pattern recognition, device-connection analysis and anomaly detection in both TCP and UDP traffics by identifying missing packets or abnormal activities (out-of-order sequences) by recognizing predictable packet sequences during normal activities, such as connection setup (SYN-ACK-ACK) during data transfer. Anomalies in UDP traffic included unexpected increases in packet rate or unusually large packet sizes. Additionally, unusual patterns in destination ports were

observed, which could indicate a DDoS attack. Normal UDP packets consisted of independent, sequenced packets typically used for real-time applications, such as DNS queries or streaming [34].

Metrics like packet-length statistics (minimum, maximum, mean and standard deviation) and flow duration (Flow Bytes/s and Flow Pkt/s) were analyzed to understand the transmission rates and detect anomalies (e.g. unusual packet sizes, irregular transmission rates and unexpected flow durations) in network traffic. An unusually large maximum packet size could indicate data aggregation before transmission, which might be normal for certain applications (e.g. video-streaming services, file-transfer protocols, cloud-storage services and backup solutions), but suspicious for others. The average packet size helped understand the typical packet load on the network. A sudden increase in the mean packet size indicated bulk data transfers. Standard deviation measured the variability in packet sizes. High variability suggested a mix of different types of traffic represented suspicious, while low variability indicated uniform traffic represented normal. A higher Flow Bytes/s rate indicated a high-volume data transfer, which could be legitimate (e.g. video-streaming) or suspicious (e.g. data exfiltration). Packet-length statistics were effective for detecting anomalies by observing baseline establishment, deviation detection and statistical methods [35]. These detail analyses helped identify those unusual patterns which were crucial for maintaining network security and efficiency.

Combining packet-length statistics with flow duration provides detailed analysis encompassing short-lived and long-lasting interactions. Flow metrics like Flow Bytes/s and Flow Pkt/s helped understand data-transmission rates, aiding in identifying abnormal patterns. Integrating packet length statistics with flow duration enables differentiation between short-lived spikes and sustained high-traffic periods, enhancing the understanding of network-flow dynamics.

Traditional ML models incorporated directional properties into their feature sets to more accurately distinguish between normal and anomalous behaviours. For instance, RF and DT benefited from the added granularity in their decision-making processes, while SVM could better separate data points in the feature space. NB, with its probabilistic approach, could more effectively categorize behaviours based on the directional data. This improved recall and precision of anomaly detection, because the algorithms were configured to recognize patterns specific to single-direction and bi-directional flows, yielding more reliable and accurate classification with better detection of anomalies.

Anomaly identification techniques focussed on analyzing unique behaviours within bi-directional and single-direction flows. Understanding the differences between these traffic patterns was important not only for anomaly detection, but also efficient resource allocation. It helped improve tactics by adapting feature sets specific to each type of flow. Targeted flow behaviours contributed to improving data-classification accuracy by reducing false-negative and false-positive results. The model could more precisely identify anomalies by focusing on each flow type's unique characteristics. Also, it reduced false alarms, which improved precision. A higher recall rate for anomaly detection could be achieved by analysing traffic patterns. A higher F1-score was achieved by the balanced improvement in both precision value and recall value, which indicated better overall performance in anomaly detection [36].

By observing dynamic load balancing, fault prediction, forecasting of traffic, protocol routing and the adaptive quality of service, these algorithms addressed bottlenecks and network inefficiencies. Those ensured that the network operated smoothly. As a result, the need for significant processing resources to solve problems after they occurred was reduced. For that reason, the overall processing time for detecting anomalies was shortened, which improved the response time [37].

7. CONCLUSIONS

Engineering features based on the directional properties that captured detailed characteristics of network-traffic behaviour. This enabled the model to identify specific patterns to bi-directional and single-direction traffic, indicating various types of network threats or activities. The high F1-score, recall, accuracy and precision achieved with these features demonstrated their effectiveness in accurately classifying network traffic. These metrics also highlighted their importance in detecting anomalies, which was important for ensuring the security and reliability of network infrastructure.

The investigation analyzed the impact of packet directionality on spyware classification. This was done through the curation of datasets focusing on directional properties. A specific emphasis was placed on IP pairs and Protocol. The analysis found that considering the directional properties

significantly improved spyware classification. RF, NB, SVM, DT and XGBoost were constructed and compared between the two curated datasets. The findings suggested that DT, RF and XGBoost performed better for Dataset B, while SVM showed better performance for Dataset A. These ML approaches demonstrated potential in spyware classification, but further improvements are needed to enhance the model, so that future work should integrate more spyware types with larger number of samples and explore advanced feature-selection and deep-learning techniques. The limited types of spyware and the small number of samples in the dataset represent limitations, so expanding them could improve detection mechanisms. Evaluating the model's performance in diverse real-world scenarios and incorporating realistic benign-traffic data could enhance its ability to distinguish between malicious and benign activities, providing a practical security solution. Integrating real-time data processing and adaptive learning could also be valuable directions for future research.

ACKNOWLEDGEMENTS

This research is supported by TM Research & Development Grant (TM R&D), MMUE/220028.

REFERENCES

[1] T. Munusamy and T. Khodadi, "Building Cyber Resilience: Key Factors for Enhancing Organizational Cyber Security," Journal of Informatics and Web Engineering, vol. 2, no. 2, pp. 59-71, 2023.

[2] M. Al-Hashedi, L.K. Soon, H. N. Goh, A. H. L. Lim and E. G. Siew, "Cyberbullying Detection Based on Emotion," IEEE Access, vol. 11, pp. 53907-53918, 2023.

[3] R. Thangaveloo et al., "Datdroid: Dynamic Analysis Technique in Android Malware Detection," Int. J. on Advanced Science, Engineering and Information Technology, vol. 10, no. 2, pp. 536-541, 2020.

[4] T.A.A. Abdullah, W. Ali, S. Malebary and A. A. Ahmed, "A Review of Cyber Security Challenges: Attacks and Solutions for Internet of Things-based Smart Home," Int. J. of Computer Science and Network Security, vol. 19, no. 9, pp. 139-146, 2019.

[5] A. S. Grillis, "What is Spyware?" [Online], Available: https://www.techtarget.com/ searchsecurity/definition/spyware, Dec. 12, 2023.

[6] S. S. Rawat and A. K. Mishra, "Review of Methods for Handling Class-imbalanced in Classification Problems," arXiv preprint, arXiv: 2211.05456, 2022.

[7] M. Botacin et al., "On the Security of Application Installers and Online Software Repositories," Proc. of the 17th Int. Conf. on Detection of Intrusions and Malware, and Vulnerability Assessment (DIMVA2020), pp. 192-214, Lisbon, Portugal, 2020.

[8] M. K. Qabalin, M. Naser and M. Alkasassbeh, "Android Spyware Detection Using Machine Learning: A Novel Dataset," Sensors, vol. 22, no. 15, pp. 5765-5790, 2022.

[9] Google Play, "PCAPdroid-Network Monitor Apps," [Online], Available: https://play.google.com/apps/, Jan. 08, 2024.

[10] M. Conti, G. Rigoni and F. Toffalini, "ASAINT: A Spy App Identification System Based on Network Traffic," Proc. of the 15th Int. Conference on Availability, Reliability and Security, Article no. 51, pp. 1-8, DOI:10.1145/3407023.3407076, August 2020.

[11] WireShark, Go Deep, [Online], Available: https://www.wireshark.org/, Dec. 12, 2023.

[12] Google Play, "DroidBox Mikrotik Config Tool-Apps," [Online], Available: https://play.Google.com/ store/ apps, Dec. 12, 2023.

[13] Google, "Google Photos," [Online], Available: https://www.google.com/photos/about/, Jan. 08, 2024.

[14] M. Naser and Q. A. Al-Haija, "Spyware Identification for Android Systems Using Fine Trees," Information, vol. 14, no. 2, pp. 1-10, 2023.

[15] D. Noetzold et al., "Spyware Integrated with Prediction Models for Monitoring Corporate Computers," Preprints.org, vol. 1, DOI: 10.20944/preprints .202301.0580.v1, 2023.

[16] F. Pierazzi, R. Emilia, R and I. V. S. Subrahmanian, "A Data-driven Characterization of Modern Android Spyware," ACM Transactions on Management Information Systems, vol. 11, pp. 1-38, 2020.

[17] VirusTotal-Home, [Online], Available: https://www.virustotal.com/gui/home/, Dec. 07, 2023.

[18] V. Mahesh and S. D. KA, "Detection and Prediction of Spyware for User Applications by Interdisciplinary Approach," Proc. of 2020 Int. Conf. on Computational Intelligence for Smart Power System and Sustainable Energy (CISPSSE), DOI: 10.1109/CISPSSE49931.2020.9212222, Keonjhar, India, July 1-6, 2020.

[19] O. F. Catak, "API Call Based Malware Dataset," [Online], Available: https://www.kaggle.com./datasets/ focatak/ malapi2019, Dec. 08, 2019.

[20] Kaggle, "Your Machine Learning and Data Science Community," [Online], Available: https://www.kaggle.com/, Nov. 01, 2024.

[21] N. Zahan, P. Burckhardt, M. Lysenko, F. Aboukhadijeh and L. Williams, "MalwareBench: Malware

Samples Are Not Enough," Proc. of 2024 IEEE/ACM 21st Int. Conf. on Mining Software Repositories (MSR), pp. 728-732, DOI: 10.1145/3643991.3644883, April 2024.
[22] Z. Zhang, P. Qi and W. Wang, "Dynamic Malware Analysis with Feature Engineering and Feature Learning," Proc. of 34th AAAI Conf. on Artificial Intelligence (AAAI-20), pp. 1210-1217, April 2020.
[23] D. Gibert et al., "Fusing Feature Engineering and Deep Learning: A Case Study for Malware Classification," Expert Systems with Applications, vol. 207, pp. 117957-117974, 2022.
[24] E. Masabo, K. S. Kaawaase, J. S. Otim, J. Ngubiri and D. Hanyurwimfura, "Improvement of Malware Classification Using Hybrid Feature Engineering," SN Computer Science, vol. 1, pp. 1-14, 2020.
[25] A. Nawaz, "Feature Engineering Based on Hybrid Features for Malware Detection over Android Framework," Turkish J. of Computer and Mathematics Education, vol. 12, no. 10, pp. 2856-2864, 2021.
[26] M. Humayun, N. Z. Jhanjhi and M. Z. Alamri, "Smart Secure and Energy Efficient Scheme for E-Health Applications Using IoT: A Review," Int. J. of Computer Science and Network Security, vol. 20, no. 4, pp. 55-74, 2020.
[27] Apktool, "Apktool," [Online], Available: https://apktool.org/, Dec. 01, 2024.
[28] J. Jung, J. Park, S. J. Cho, S. Han, M. Park and H. H. Cho, "Feature Engineering and Evaluation for Android Malware Detection Scheme," J. of Internet Technology, vol. 22, no. 2, pp. 423-440, 2021.
[29] K. Allix et al., "AndroZoo: Collecting Millions of Android Apps for the Research Community," Proc. of the 13th Int. Conf. on Mining Software Repositories (MSR), pp. 468-471, Austin, USA, May 2016.
[30] M. X. Low et al., "Comparison of Label Encoding and Evidence Counting for Malware Classification," Journal of System and Management Sciences, vol. 12, no. 6, pp. 17-30, 2022.
[31] T. N. AlMasri and M. A. N. AlDalaien, "Detecting Spyware in Android Devices Using Random Forest," Proc. of the 2023 Int. Conf. on Advances in Comput. Research (ACR'23), pp. 294-315, 2023.
[32] N. Ben-Asher, S. Hutchinson and A. Oltramari, "Characterizing Network Behavior Features Using a Cyber-security Ontology," Proc. of MILCOM 2016-2016 IEEE Military Communications Conf., pp. 758-763, Baltimore, USA, November 2016.
[33] S. Misra, M. Tan, M. Rezazad, M. R. Brust and N. M. Cheung, "Early Detection of Crossfire Attacks Using Deep Learning," arXiv preprint, arXiv: 1801.00235, 2017.
[34] L. Zhou et al., "DDOS Attack Detection Using Packet Size Interval," Proc. of the 11th Int. Conf. on Wireless Comm., Networking and Mobile Computing (WiCOM), pp. 1-7, Shanghai, China, 2015.
[35] A. Iorliam et al., "Flow Size Difference Can Make a Difference: Detecting Malicious TCP Network Flows Based on Benford's Law," arXiv preprint, arXiv: 1609.04214, 2016.
[36] N. Davis, G. Raina and K. Jagannathan, "A Framework for End-to-End Deep Learning-based Anomaly Detection in Transportation Networks," Transportation Research Interdisciplinary Perspectives, vol. 5, pp. 100-112, 2020.
[37] M. Kuchnik et al., "Plumber: Diagnosing and Removing Performance Bottlenecks in Machine Learning Data Pipelines," Proc. of Machine Learning and Systems, vol. 4, pp.33-51, 2022.

ملخص البحث:

تفتقر الطُّرق الراهنة لتصْنيف برامج التّجسُّس إلى الفعالية لغياب مجموعات البيانات جيدة التّنظيم، وبخاصّةٍ مجموعات البيانات ذات الخصائص الاتّجاهية. في هذه الورقة، يتمّ استكشاف فعالية الخصائص المرتبطة بالاتّجاهية في مجموعات البيانات من أجل التّصنيف، من خلال خصائص تجري هندستها من الخصائص الموجودة في مجموعات البيانات. وتعمل هذه الدّراسة على تنظيم اثنتين من مجموعات البيانات. المجموعة A تحتوي على خصائص يتمّ استخلاصُها من تدفّقات حُزَم أحادية الاتّجاه، بينما تحتوي المجموعة B على خصائص يتمّ استخلاصها من تدفّقات حُزَم ثنائية الاتّجاه. وتجدر الإشارة إلى أنّ عملية التّصنيف من خلال تلك الخصائص تتمّ باستخدام مجموعةٍ مختارة من خوارزميات التّصنيف. وقد حقّق مصنِّف (SVM) الدّقّة الأعلى بالنّسبة لمجموعة البيانات A 99.88%، في حين ذهبت الدّقّة الأعلى بالنّسبة لمجموعة البيانات B إلى مُصنِّفات (RF و DT و XGBoost) 99.24%.

وبمقارنة الطّريقة المقترحة في هذا البحث مع غيرها من الطّرق الواردة في أدبيات الموضوع، تبيَّن أنّ الخصائص الاتجاهية في مجموعات البيانات من شأنها أنْ تحقّق تحسيناتٍ من حيث الدّقّة في تصنيف برامج التّجسُّس.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

A NOVEL EVIDENTIAL COLLABORATIVE FILTERING FRAMEWORK BASED ON DISCOUNTING CONFLICTING PREFERENCES

Khadidja Belmessous, Faouzi Sebbak and M'hamed Mataoui

(Received: 20-Jun.-2024, Revised: 2-Sep.-2024 and 24-Sep.-2024, Accepted: 27-Sep.-2024)

ABSTRACT

This paper presents a novel framework to enhance Evidential Collaborative Filtering (ECF), a critical Recommender System (RS) designed for sensitive domains like healthcare and target tracking. The focus is on refining how user-rating imperfections are handled, particularly in managing conflicting preferences during neighborhood selection to boost recommendation quality. The newly proposed ECF architecture integrates a two-probabilities-focused approach with an advanced conflict-management technique, employing Deng relative entropy and the Best Worst Method. This allows for assigning more accurate reliability weights to each user, improving preference selection and rating prediction in ECF. Experimental evaluations on Movielens-100K and Flixster datasets show that our framework surpasses baselines in prediction error, precision, recall and F-score.

KEYWORDS

Recommender systems, Collaborative filtering, Dempster-Shafer theory, Conflict, Fusion.

1. INTRODUCTION

Recommender systems (RSs) have been categorized in the literature into three main approaches; namely, content-based filtering (CBF) [1], collaborative filtering (CF) [2], and hybrid filtering [3]. CBF provides recommendations based on user profiles, which are generally difficult to acquire. On the other hand, CF generates recommendations by using the preferences of the most similar users. Hybrid filtering is a combination of both CBF and CF. Compared to CBF, CF has made significant progress due to the ease with which real-world information about users' preferences on items may be obtained [4].

Collaborative filtering is a leading approach in RS, based on the idea that our purchase decisions are usually influenced by our similar neighbors. Sparsity is a key challenge in CF, representing the proportion of missing ratings to the overall rating-matrix size. In CF, the subjective nature of user ratings and their intrinsic sparsity not only increase the uncertainty, but also affect the trustworthiness of the recommendation outputs [5]-[6]. Evidential Collaborative Filtering (ECF) is a sub-class of CF that addresses the sparsity issue by handling the inherent uncertainty in RS under the framework of Dempster-Shafer Theory (DST), also called evidence theory [7]-[8]. ECF can be categorized into three main types [9]: ECF using evidential fusion to combine multi-source information, ECF offering soft ratings and ECF providing evidential predictions.

This paper primarily focuses on a specific type of ECF that utilizes soft rating systems. This ECF addresses the limitations of traditional hard-rating scores in capturing user preferences, which can sometimes be an inadequate representation [10]. For instance, consider a user who rates two items, i_1 and i_2, with scores of 3 and 4, respectively. If this user wants to rate a third item, i_3, as better than i_1, but not as good as i_2, standard rating scales might not accurately reflect this nuanced preference. The ECF framework discussed here allows for more flexible user ratings, like a range of {4,5}, to better capture these subtle preferences. Essentially, this branch of ECF is designed to account for the subjective and sometimes imprecise nature of user preferences [11].

Imperfections and conflicts in user preferences negatively affect the trustworthiness and effectiveness of ECF systems [5]. These imperfections can arise due to several factors, including uncertainty, ambiguity and contradictions in user feedback. Existing ECF frameworks rely mainly on the use of

Kh. Belmessous (Corresponding Author), F. Sebbak and M. Mataoui are with Department of Computer Science, Ecole Militaire Polytechnique, Bordj El Bahri, P.O.Box 17, Algiers, Algeria. Emails: kh.belmessous@gmail.com, faouzi.sebbak@emp.mdn.dz and mhamed.mataoui@emp.mdn.dz

Dempster's combination rule (DCR) in combining users' preferences [9]. Nguyen and Huynh explored the fusion of information in RS using DST, as detailed in [12]. They concluded that DCR is ineffective for combining user ratings due to its weakness to handle highly conflicting mass functions [13]. Recently, Belmessous et al. [9] highlighted shortcomings in the existing ECF framework, which often overlooks the importance of managing these conflicting preferences through advanced techniques. This paper addresses this gap by proposing a novel framework that integrates recent advancements in DST to better manage conflict, thereby enhancing the overall performance of ECF systems.

Research on ECF has provided limited solutions for managing conflicting user preferences [9], falling behind in the ongoing advancements within DST research. DST is continually evolving to tackle challenges related to dealing with highly conflicting information [13]. Many studies in ECF overlook the discounting factor, which is key in DST for determining the reliability of user ratings. Nguyen and Huynh [12] have explored the integration of information in RS, highlighting the difficulty in combining mass functions that are highly conflicting. In RS, it's quite common to encounter users giving completely opposite ratings to the same item, which leads to frequent conflicts in mass combination.

Although DST research is continually proposing new solutions for conflict management [14], ECF has not fully capitalized on these advancements. The ongoing challenge in DST of effectively combining highly conflicting evidence remains a significant issue. This gap in ECF [15]-[16], where advanced DST conflict-management proposals are underutilized, is a key focus of this paper. We intend to bridge this gap by incorporating recent solutions for conflict management [17] into ECF systems to enhance their performance.

This paper presents the following key contributions:

- Introduction of a novel framework designed to manage imperfections in users' preferences throughout the decision-making process in ECF;
- Proposition of a new neighbourhood-selection strategy in ECF, utilizing optimal discounting weights;
- Proposition of an efficient method for preference-prediction estimation in ECF.

The remainder of this article is structured as follows: In Section 2, we provide a summary of the theoretical concepts underlying our new approach. We then outline our proposed framework and its main components in Section 3. Section 4 describes our experimental design and presents the obtained results. Section 5 includes a discussion of our findings, strengths and limitations. The article concludes with Section 6, where we summarize our work and suggest areas for future research.

2. BACKGROUND AND RELATED WORK

In this section, we explore foundational theories and contributions important to our study, with a particular focus on the Dempster-Shafer Theory (DST) and conflict management. Initially, we introduce DST, known for its ability to manage uncertainty and make decisions based on evidence. This theory is valuable in decision-making, where the quality of information is crucial. Subsequently, we discuss the metrics used to evaluate conflict within this theoretical framework. Additionally, we explore the conflict-management methodology based on discounting optimal weights and present the related research on ECF offering soft ratings.

2.1 Dempster-Shafer Theory

The Dempster-Shafer theory is a flexible method for modeling uncertainty that does not require assigning a probability to every element in a set. The DST was introduced by Arthur P.Dempster in the context of statistical inference [18], and it was further developed by his student Shafer [19].

DST is founded on a number of concepts, including: the frame of discernment, the mass function also called basic probability assignment (BPA) and Dempster's combination rule. Concerning the frame of discernment Θ, it is a finite set representing the problem domain. All propositions of interest are defined by elements in 2^{Θ}. A BPA is defined as a mapping $m(.) \in [0,1]$ that meets the following properties:

$$m(\emptyset) = 0 \qquad \emptyset: the\ empty\ set$$

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

$$\sum_{H\in 2^{\Theta}} m(H) = 1 \qquad H: a\ subset\ of\ \Theta.$$

The quantity $m(H)$ can be interpreted as a measure of the belief that is committed exactly to H, given the available evidence. All subsets $H\in 2^{\Theta}$ having a positive mass are considered as focal elements of $m(.)$. Concerning the Dempster's combination rule, it is an operation that permits to combine evidence from multiple independent sources under the same frame of discernment. Let m_1 and m_2 be the two BPAs associated with two independent sources of evidence. H_1 and H_2 are the focal elements of m_1 and m_2, respectively. The resulting mass function m is the combination of m_1 and m_2 and is noted by $m = m_1 \oplus m_2$. The Dempster's combination rule (DCR) is defined by Equations 1 and 2, where m_{DS} is the result of Dempster's combination.

$$m_{DS}(H) = \frac{m_{12}(H)}{1-m_{12}(\emptyset)} \tag{1}$$

where,

$$m_{12}(H) = \sum_{\substack{H_1,H_2\in 2^{\Theta} \\ H_1\cap H_2 = H}} m_1(H_1)m_2(H_2) \tag{2}$$

The body of evidence in DST encompasses all BPAs from independent sources, serving as the aggregate evidence for decision-making. It forms the basis for applying Dempster's combination rule, enabling the synthesis of evidence across the problem domain.

In the DST framework, decision criteria can include: the maximum of the belief *Bel*(H), which indicates the comprehensive support that evidence lends to a hypothesis *H*; the maximum of the plausibility *Pl(H)*, reflecting the extent to which evidence does not contradict *H*; or the pignistic probability *BetP(H)* [20], which provides a practical way to make decisions under uncertainty by balancing the evidence supporting different hypotheses. The relationship between belief and plausibility is illustrated in Figure 1.

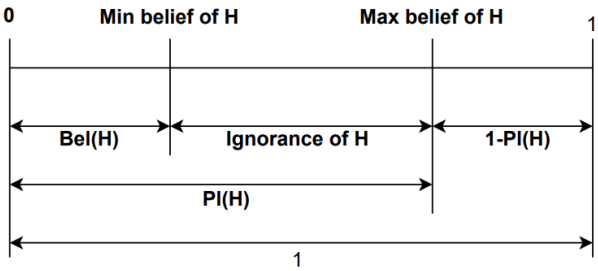


Figure 1. Relationship between belief and plausibility.

Another important tool in the DST framework is the discounting factor proposed by Shafer [19]. The factor α is considered as a discounting rate permitting to control the reliability of the BPA. When α is set to 1, the BPA is deemed fully reliable; conversely, an α value of 0 signifies that the BPA is entirely unreliable. The discounting of the BPA $m(\cdot)$ is defined as follow:

$$\begin{Bmatrix} m'(H) = \alpha.m(H), & \forall H \in 2^{\Theta}, H \neq \Theta \\ m'(\Theta) = (1-\alpha) + \alpha.\mathrm{m}(\Theta) & \end{Bmatrix} \tag{3}$$

with $m'(.)$ representing the unreliable source.

2.2 Conflict Metrics in Dempster-Shafer Theory

In scenarios where Dempster's combination rule is applied to fuse evidence from multiple sources, it's possible to reach counter-intuitive conclusions, especially when the evidence conflicts significantly [13], [21]. Consistently, there are novel propositions being introduced for conflict metrics to enhance the accuracy of assessing conflict levels of evidence. Consider two BPAs, m_1 and m_2, defined under the Frame of Discernment (FoD) $H = \{H_1, H_2, \dots, H_i, \dots, H_n\}$. Some representative metrics for evaluating conflict are summarized in Table 1.

In Jousselme et al.'s distance equation, $\widehat{m}_1$ and $\widehat{m}_2$ represent the vector forms of the basic probability assignments m_1 and m_2, respectively and $\overline{D}$ is the Jaccard matrix between all pairwise

propositions in m1 and m2. An increased distance in this measure indicates a higher level of conflict among the evidence. Similarly, in Song et al.'s correlation-coefficient equation, $\widehat{m}_1 = m_1.D$ and $\widehat{m}_2 = m_2.D$ are used, where D is the Jaccard matrix applicable to all propositions in m_1 and m_2. In Jiang's correlation-coefficient equation, H_i and H_j serve as focal elements within the power concentration of the frame and their relationship is quantified through a modulus calculation that involves the intersection and union of H_i and H_j.

Table 1. Summary of some representative conflict metrics for DST.

Metric Name	**Equation**
Jousselme et al.'s evidence distance *d* [22]	$d(m_1, m_2) = \sqrt{\frac{1}{2}(\widehat{m}_1 - \widehat{m}_2)^T \overline{D}(\widehat{m}_1 - \widehat{m}_2)}$
Song et al.'s correlation coefficient *cor* [23]	$K_{cor}(m_1, m_2) = 1 - cor(m_1, m_2)$, $cor(m_1, m_2) = \frac{\langle \widehat{m}_1, \widehat{m}_2 \rangle}{\|\widehat{m}_1 . \widehat{m}_2\|}$
Jiang's correlation coefficient k_r [24]	$k_r(m_1, m_2) = 1 - \sum_{i=1}^{2^{\|n\|}} \sum_{j=1}^{2^{\|n\|}} m_1(H_{i)}\, m_2(H_{j)} \frac{\|H_i \cap H_j\|}{\|H_i \cup H_j\|}$
Xiao et al.'s correlation coefficient *ECC* [25]	$k_{ECC}(m_1, m_2) = 1 - ECC(m_1, m_2)) = 1 - \left[\frac{\langle \widehat{m}_1, \widehat{m}_2 \rangle}{\|\widehat{m}_1 . \widehat{m}_2\|}\right]^2$

These methods provide different approaches to understand and quantify the level of agreement or conflict between various BPAs, each with its unique application and implications for decision-making.

2.3 Conflict Management by Considering the Optimal Discounting Weights Using the BWM Method

This sub-section introduces the conflict-management method by considering the optimal discounting weights based on the Best-Worst Method (BWM) [26] to manage evidential conflict in DST. This recent methodology involves selecting the best and worst BPAs to calculate discount weights effectively before the fusion process. The detailed steps of this method are outlined as follows:

1) Evidential distance-matrix establishment: an evidential distance matrix is calculated using Jousselme's distance measure to evaluate the distances between each pair of evidence, helping identify the relative degrees of conflict.

2) Determination of worst and best BPAs:

   - The worst BPA, represents the maximum contribution to overall system conflict.
   - The best BPA is determined based on its relative distance to the worst BPA.

3) Preference calculation for best and worst BPAs: Fei and Deng [27] introduced a new metric called Deng relative entropy to measure the discrepancy between BPAs. Deng relative entropy, as described by formula 4, is specifically designed for mass functions in the context of DST.

$$r(m_1 \| m_2) = \sum_i m_1(L_i) log \frac{m_1(L_i)}{m_2(L_i)} \quad (4)$$

   Deng relative entropy calculates the average logarithmic difference between two BPAs, m1 and m2, thus providing a measure of the informational divergence between them.

   Establishing preference vectors for the best and worst BPAs: by utilizing Deng relative entropy, the preference vector for the best BPA, denoted by m_B, relative to other BPAs is calculated as follows:

$$M_B = (\sigma(m_B \| m_1), \sigma(m_B \| m_2), \dots, \sigma(m_B \| m_n)) \quad (5)$$

   where $\sigma(m_B \| m_j)$ quantifies the relative preference of the best BPA m_B over other BPA j. It is defined such that $\sigma(m_B \| m_B) = 1$, indicating the highest self-preference. Similarly, the preference vector for the worst BPA, m_W in relation to other BPAs is given by:

$$M_W = (\sigma(m_1 \| m_W), \sigma(m_2 \| m_W), \dots, \sigma(m_n \| m_W))^T \quad (6)$$

In this vector, $\sigma(m_{Bi} \| m_W)$ measures the preference of each BPA mi over the worst BPA m_W. This measurement also adheres to the condition $\sigma(m_W \| m_W) = 1$, reflecting maximum self-preference and its role as the most conflict-contributing BPA.

4) Finding the optimal weights for BPAs:

In this phase, optimal weights $(\widetilde{w}_1, \widetilde{w}_2, \dots, \widetilde{w}_n)$ are determined to refine the evidential contributions more effectively. The Consistency Ratio (CR) ζ^* plays an essential role in this process by measuring the consistency of these weights, which is pivotal for evaluating the pairwise comparison's efficacy. Divergences in the expected proportional relationships, such as when $\frac{w_B}{w_j} \neq \sigma(m_B \| m_j)$ or $\frac{w_j}{w_W} \neq \sigma(m_{Bj} \| m_W)$, necessitate a re-evaluation of CR to ensure the reliability of the weight assignments.

$$min\, max_j \left\{ \left| \frac{w_B}{w_j} - m_{Bj} \right|, \left| \frac{w_j}{w_W} - m_{jW} \right| \right\}$$

$$s.t. \sum_{\substack{w_j \geq o \\ j=\{1,2,\dots,n\}}}^{min\, \zeta} w_j = 1 \quad \Rightarrow s.t. \begin{cases} \left| \frac{w_B}{w_j} - \sigma(m_B \| m_j \right| \leq \xi \\ \left| \frac{w_j}{w_W} - \sigma(m_j \| m_W \right| \leq \xi \\ \sum_j w_j = 1 \\ w_j \geq 0 \\ j = \{1,2,\dots,n\} \end{cases} \quad (7)$$

Therefore, the optimal weights $(\widetilde{w}_1, \widetilde{w}_2, \dots, \widetilde{w}_n)$ and CR (ζ_*) of BPAs could be calculated. Meanwhile, the CR can be defined as follows:

$$\eta_{CR} = \frac{\zeta_*}{max\{\zeta\}} \quad (8)$$

5) Discounting and fusion: optimal weights obtained from the previous steps are used to discount the BPAs before fusion using the DCR.

This conflict-management methodology [17] ensures the reliability of the optimal weights by employing a consistency ratio for reference comparisons, guaranteeing that each piece of evidence contributes appropriately to the final fused result.

2.4 Related Work on Evidential Collaborative Filtering Offering Soft Ratings

The pioneering effort in the area of ECF that introduces soft preferences in RS was initiated by Wickramarathne et al. [28]. This approach leverages the DST to effectively handle uncertainties in user preferences for a CF system. Emphasizing prediction accuracy, this evidential RS design accepts higher computational demands. Its sophisticated nature makes it suitable for critical and advanced applications, including those in medical and healthcare services and security-threat evaluations. Subsequently, Nguyen expanded on this foundation by developing an evidential RS that incorporates soft ratings, drawing inspiration from Wickramarathne et al. [28], and tackling the issue of data sparsity by leveraging community context under the DST framework [11]. Nguyen and Huynh further enhanced this system by integrating the reliability of predicted ratings, acknowledging their inherent imprecision compared to real ratings, to refine the recommendations [29]. Later, Nguyen aimed at reducing computational load by proposing an optimization that prioritizes the combination of focal elements with the top two probabilities within their sets [10].

Furthermore, Nguyen et al. extended the application of their ECF to incorporate social-media platforms [30]. In this context, user ratings and community preferences gathered from social networks are represented as mass functions. These are then combined according to Dempster's rule of combination. Moreover, Nguyen and Huynh introduced an innovative approach for combining evidence in their system as described in [29] through [31]. Their technique focuses on discarding focal elements with negligible probabilities, considered as noise in the fusion of information, thereby enhancing the efficiency of computations without sacrificing data integrity. In addition, their research

in [12] delves into optimizing evidence combination for DST-based RS. This study establishes the essential parameters for crafting a combination operator that aligns with the requisites of DST-based RS. Within this framework, Nguyen and Huynh unveiled new strategies for executing mixed combinations, showcasing their commitment to refining the DST-based RS.

Nguyen's 2017 study [15] introduced an innovative BPA combination approach named the "Two-probabilities focused-combination method". This method permits to combine belief masses with significant conflicts and offers the advantage of decreased computational time. Although the proposed method is not stable due to the fact that it is non-associative, indicating that the order of inputs can influence the results, the sequence in which inputs are combined has an impact on the outcome. Further, Nguyen and Huynh tackled the challenges of data sparsity and the cold-start problem in [32] through an ECF that incorporates soft ratings alongside community preferences. They also proposed a novel approach for assessing user-user similarity, prioritizing provided over predicted ratings, within a similar system [33]. Dong et al. followed up with a different strategy in [34], introducing the modified rigid coarsening method based on hierarchical decomposition to simplify the frame of discernment in the combination process. Lastly, Bahri et al. presented ECFAR in [16], a rule-based CF system that leverages the DST, marking another contribution to the field.

3. METHODOLOGY

This study presents a novel framework, Conflict-Aware Evidential Collaborative Filtering (CA-ECF), which integrates an advanced conflict-management methodology from recent research [17] into a classical ECF framework [15]. This methodology, aimed to managing evidential conflict within DST, optimizes weights for BPAs using the BWM, as elaborated in sub-section 2.3.

To ensure clarity and consistency of mathematical notations throughout our proposition, we have defined all the used variables and notations in Table 2.

Table 2. Notations' table.

Symbol	Description
R_{MN}	Rating matrix with M and N representing the total number of users and items, respectively. Here, M corresponds to the set of users $U = \{U_1, U_2, ..., U_M\}$ and N corresponds to the set of items $I = \{I_1, I_2, ..., I_N\}$.
$\hat{R}_{MN}$	Dense User-Item rating matrix.
Θ	Set of preference levels, denoted by $\Theta = \{\theta_1, \theta_2, ..., \theta_L\}$, where L is the number of the available preferences.
$r_{i,k}$	Rating of user U_i on item I_k.
C	Set of concepts within the contextual data, denoted by $C = \{C_1, C_2, ..., C_P\}$, where P is the total number of concepts. Each concept C_p, with $1 \leq p \leq P$, can consist of at most Q_p groups, indicating that $C_p = \{G_{p,1}, G_{p,2}, ..., G_{p,Qp}\}$.
$g_p(U_i)$	Groups within concept C_p that user U_i is interested in.
$g_q(I_k)$	Groups within concept C_q that item I_k is associated with.
$G_{p,q}$	The intersection of user and item interest groups associated with concept C_p.
$m_{i,k}$	BPA corresponding to a rating $r_{i,k}$.
$\uplus$	Two-probabilities focused combination.
$d(U_i, U_j)$	Jousselme's distance measure between users U_i and U_j.
$s(U_i, U_j)$	Similarity score between users U_i and U_j.
m_B, m_W	Best and worst BPAs.
$\sigma(m_1 \| m_2)$	Deng relative entropy measuring the conflict between two BPAs m_1 and m_2.
M_B, M_W	Vectors representing the preference of the best BPA and worst BPA over other BPAs using the Deng relative entropy.
N_{U_i}	Set of k closest neighbors for user U_i.
$knn_{i,k}$	Set of neighbors of user U_i that have rated the target item I_k.
ξ	Optimization variable used to minimize the optimal weights.
w_i	Weight assigned to the i^{th} BPA, used in the discounting and fusion processes.
$\hat{r}_{ik}$	Predicted rating.

The proposed CA-ECF framework represents an advanced version of the classical ECF, with its main characteristics detailed in [10][15][28]. CA-ECF innovates by using conflict in user preferences to identify the most similar neighbors. It then adjusts their influence in making predictions based on their optimal weights. The architecture of the proposed framework is depicted in Figure 2.

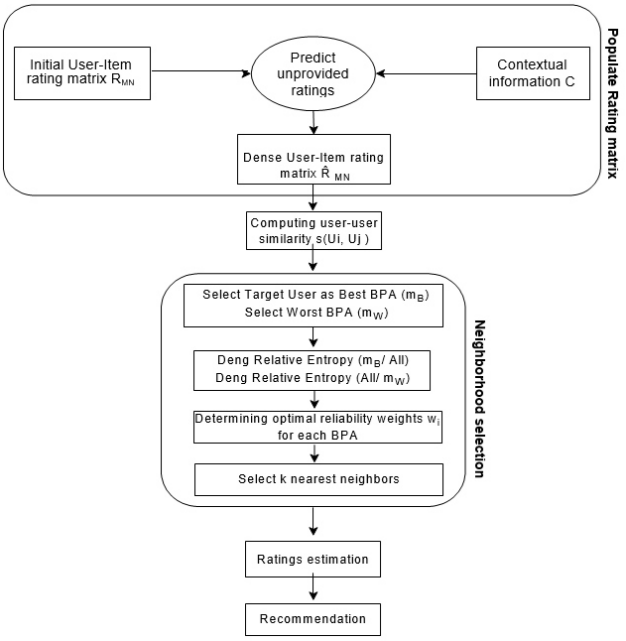


Figure 2. The proposed CA-ECF framework.

The CA-ECF framework, similar to classical ECF, follows five distinct steps, as illustrated in Figure 2. Initially, the unrated entries within the rating matrix R_{MN} are calculated using contextual information C in order to construct a dense rating matrix $\hat{R}_{MN}$. Subsequently, user-user similarities $s(U_i, U_j)$ are calculated using both provided and predicted ratings in $\hat{R}_{MN}$. For each active user U_i, a neighborhood set $knn_{i,k}$ is selected and the user's rating for each item is estimated based on the combined ratings from these selected neighbors. Following this, the estimated ratings for all unrated items are systematically ranked and the most appropriate items are chosen for recommendations to the active user.

In classical ECF, neighborhood sets $knn_{i,k}$ for each unrated item I_k are determined based on similarity scores $s_{i,j}$, which must meet or exceed a specific threshold. This traditional approach, however, does not provide a mechanism to assess the reliability of the selected neighbors. In contrast, the CA-ECF framework selects neighborhood sets based on their corresponding optimal discounting weights. These weights are then utilized to discount the BPAs during the prediction step, thereby refining the accuracy of the recommendations.

In the following sub-sections, we will explore each step of the CA-ECF recommendation process in detail.

3.1 Constructing Dense User-item Rating Matrix

In the classical ECF architecture, each user evaluation is represented as a BPA (m) that spans the evaluation space Θ, enabling it to capture a wide range of user preferences, for instance: uncertain and ambiguous data. The first step of the CA-ECF framework is to predict all the unrated entries $r_{i,k}$ of the user-item matrix using contextual data C in order to mitigate the sparsity issue of CF. Contextual data consists of a set of concepts $C = \{C_1, C_2, \dots, C_P\}$, where each concept p encompasses a set of groups G_p. Both CA-ECF and ECF consider that users who share an interest in a particular group will also have similar choices with respect to that group. The group preference is defined as follows:

- First, consider a concept Cp. For each group $G_{p,q}$ that intersects with $G_p(U_i)$, which is the users' group of interest and $g_q(I_k)$, the items' group of interest, it is assumed that the group's overall preference for item I_k within $G_{p,q}$ reflects the specific group preference of user U_i for item I_k within the same group. Therefore, the concept preference of user U_i for item I_k related

to concept C_p is the result of the combination of all the group preferences, represented as two-probabilities focused-mass functions.
- Second, the overall context preferences are computed as the combination of all concept preferences for the target item, represented as two-probabilities focused-mass functions.
- Finally, the unrated entry $r_{i,k}$ is replaced by the context preference of user U_i for item I_k. If the context information does not allow for making conclusions on the concept preference, then the unrated entry is determined by aggregating the ratings from users who have rated item I_k.

At this stage, all the user-item matrix entries $\hat{R}_{MN}$ are given (provided and predicted) and they all will be used in the subsequent steps.

3.2 Computing User-User Similarity

In contrast to the classical ECF systems, in the CA-ECF we propose to evaluate the similarity between users using Jousselme's distance [22], a decision that directly supports the used conflict-management approach [17].

$$d(U_i, U_j) = \sqrt{\frac{1}{2}(\hat{m}_i - \hat{m}_j)^T \overline{D}(\hat{m}_i - \hat{m}_j)} \tag{9}$$

Since ratings have two sources (provided and predicted), we discount the predicted ratings [29].

$$s(U_i, U_j) = \sum_{k=1} \mu(x_{i,k}, x_{j,k}) * d(U_i, U_j) \tag{10}$$

where $\mu(x_{i,k}, x_{j,k})$ is calculated as follows:

$$\mu(x_{i,k}, x_{j,k}) = 1 - w_1(x_{i,k} + x_{j,k}) - w_2 x_{i,k}, x_{j,k}$$

where w_1 and w_2 are the reliability coefficients [29].

User-user similarities are stored as a matrix. The lower the value of $s(U_i, U_j)$ the more similar user U_i is to user U_j.

3.3 Neighborhood Selection

Consider a target user-item pair, (U_i, I_k). We select a set of the k closest neighbors for U_i, denoted by N_{Ui}, by following four steps, as outlined below:

1) Define best and worst BPA: in order to define those two BPAs, we first define the set of neighbors that have rated the target item I_k, following the equation below:

$$knn_{i,k} = \{U_j \in \mathrm{U} \mid I_k \in R(U_j)\} \tag{11}$$

Then, we set best BPA as the target user $m_B = mU_i$. Additionally, the worst BPA can be determined using the best BPA. The exact definition is provided as follows:

$$m_W = \max_i s(m_B, m_i) \tag{12}$$

2) Compute Deng's relative entropy (best/ others) and (others/ worst): Deng relative entropy is given by the following equation:

$$\sigma = (m_1 \| m_2) = \sum_i m_1(L_i) log \frac{m_1(L_i)}{m_2(L_i)} \tag{13}$$

At this stage, in order to compute the reliability factors, two vectors need to be calculated using the $knn_{i,k}$ set of users.

$$M_B = (\sigma(m_B \| m_1), \sigma(m_B \| m_2), \ldots, \sigma(m_B \| m_n)) \tag{14}$$

which describes the preference of the best BPA m_B over the other BPAs and

$$MW = (\sigma(m1 \| mW), \sigma(m2 \| mW), \ldots, \sigma(mn \| mW))T \tag{15}$$

which describes the preference of BPAs m_i over the worst BPA.

Determining optimal-reliability factors: this step involves determining the optimal weights for BPAs to improve the process of discounting evidence. The consistency ratio (Equation 8) is crucial in this step, as it assesses the consistency of these weights. This step follows a constrained-optimization approach, as formulated in Equation 7, to establish the weights

accurately, relying on BWM. By solving an optimization problem that includes non-linear constraints, the optimal weights for the evidence are obtained.

Select k-nearest neighbors: the selection of neighborhoods is based on reliability factors. We order all members within $knn_{i,k}$ in descending order according to their reliability factors, denoted by w_i. Then, the top K members from this ordered list are chosen to form the neighborhood set N_{Ui}.

3.4 Ratings' Estimation

Rating estimation for each unrated item I_k by an active user U_i is computed using the ratings from the user's neighborhood. Ratings are first adjusted by their respective discounting weights according to Equation 3. Then, the two-probabilities-focused method is used to fuse the evidence to obtain the final fusion result. The steps for preference aggregation are outlined in Algorithm .

Algorithm 1. Preference aggregation for rating estimation in CA-ECF.

procedure EstimateRating(U_i,I_k,Neighborhoods N_{Ui})
    Initialize $\tilde{r}_{i,k} \leftarrow 0$ ▷Initialize the estimated rating for item I_k
    for each neighbor $U_j \in N_{U\tilde{i}}$ **do**
        $r_{j,k} \leftarrow$ rating of U_j on I_k
        $w_{i,j} \leftarrow$ discounting weight
        $\tilde{r}_{j,k} \leftarrow$ discounted BPA according to Equation 3
        $\tilde{r}_{i,k} \leftarrow \tilde{r}_{i,k} \uplus \tilde{r}_{j,k}$ ▷Fusion of discounted BPAs
    end for
    output the estimated rating $\tilde{r}_{ik}$
end procedure

This algorithm synthesizes the weighted contributions of a user's neighbors to predict unrated items. By applying discounting weights, which are optimized during the neighborhood-selection phase, the reliability of each contribution is assessed, ensuring that the final-rating estimation for $\hat{r}_{i,k}$ is not only a reflection of collective-neighborhood opinion, but also of its credibility and relevance to user U_i's preferences.

3.5 Recommendation

Notably, ECF systems can produce both hard (rating as singleton) and soft (rating as sub-sets) recommendations. For a hard recommendation, the pignistic-probability method is employed to select the item with the highest likelihood as the preferred choice. Conversely, for a soft recommendation, the system adopts a maximum-belief strategy with an overlapping interval approach (maxBL) [15], [35]. This method selects an item based on its belief being greater than the plausibility of any alternative, ensuring that a decision can still be made when a direct class label is absent by favoring a composite class label that combines the most believable item with those of higher plausibility.

4. EXPERIMENTS AND RESULTS

Our experiments were performed on Movielens-100K [36], and Flixster [11] datasets. The MovieLens-100K dataset consists of 943 users who have provided 100,000 ratings for 1,682 movies. The ratings are given on a five-point scale, represented as $\Theta = \{1,2,3,4,5\}$. Each user in this dataset has rated at least 20 movies. On the other hand, Flixster dataset includes 535,013 ratings from 3,827 users for 1,210 movies. The rating scale in this dataset is composed of ten possible scores, denoted as $\Theta = \{0.5,1.0,1.5,2.0,2.5,3.0,3.5,4.0,4.5,5.0\}$. Each user has provided at least 15 ratings.

Moreover, in Movielens-100K, the information used to categorize users is the genre, which has 19 values.

$C_1 = \{G_{1,1}, G_{1,2}, \ldots, G_{1,19}\}$ = {Unknown, Adventure, Action, Animation, Children's, Comedy, Drama, Documentary, Crime, Musical, Film-Noir, Fantasy, Horror, Western, Sci-Fi, Romance, Thriller, War,

Mystery}.

$$m_{i,k} = \{\alpha_{i,k}(1-\alpha_{i,k}) \quad for\ A = \theta_l$$

$$m_{i,k} = \begin{cases} \alpha_{i,k}(1-\alpha_{i,k}), & for\ A = \theta_l \\ \alpha_{i,k}\sigma_{i,k}, & for\ A = B; \\ 1-\alpha_{i,k} & forA = \Theta; \\ 0, & otherwise \end{cases} \text{ with } B = \begin{cases} (\theta_1, \theta_2), & if\ l = 1; \\ (\theta_{L-1}, \theta_L), & if\ l = L \\ (\theta_{l-1}, \theta_l, \theta_{l+1}), & otherwise \end{cases} \quad (16)$$

Movielens-100K was transformed into an evidential dataset using Equation 16 as proposed in [15], where $\alpha_{i,k} \in [0, 1]$ and $\sigma_{i,k}$ are trust actor and dispersion factor, respectively. Also, given the absence of specific information regarding the genres that a user prefers, it is presumed that a user's interest spans all genres associated with any item having been rated.

In the Flixster dataset, every hard rating $r_{i,k}$ was converted into a soft rating $m_{i,k}$ using the Dempster-Shafer modeling function [11], as explained below:

$$m_{i,k}(A) = \begin{cases} \alpha_{i,k}(1-\sigma_{i,k}), & for\ A = \{\theta_l\}; \\ \frac{3}{5}\alpha_{i,k}\sigma_{i,k}, & for\ A = B; \\ \frac{2}{5}\alpha_{i,k}\sigma_{i,k}, & for\ A = C; \\ 1-\alpha_{i,k}, & for\ A = \Theta; \\ 0, & otherwise. \end{cases} \quad (17)$$

where $B = \begin{cases} (\theta_1, \theta_2), & if\ l = 1; \\ (\theta_{L-1}, \theta_L), & if\ l = L; \\ (\theta_{l-1}, \theta_l, \theta_{l+1}), & otherwise; \end{cases}$ and $C = \begin{cases} \{\theta_1, \theta_2, \theta_3\}, & if\ l = 1; \\ \{\theta_1, \theta_2, \theta_3, \theta_4\}, & if\ l = 2; \\ \{\theta_{L-3}, \theta_{L-2}, \theta_{L-1}, \theta_L\}, & if\ l = L-1; \\ \{\theta_{L-2}, \theta_{L-1}, \theta_L\}, & if\ l = L; \\ \{\theta_{l-2}, \theta_{l-1}, \theta_l, \theta_{l+1}, \theta_{l+2}\}, & otherwise; \end{cases}$

The available genres in Flixster dataset are as follows:

Genre = {Drama, Comedy, Action & Adventure, Television, Mystery & Suspense, Horror, Science Fiction & Fantasy, Kids & Family, Art House & International, Romance, Classics, Musical & Performing Arts, Anime & Manga, Animation, Western, Documentary, Special Interest, Sports & Fitness, Cult Movies}.

It's important to highlight that the selection of parameters within these systems is primarily influenced by the outcomes analyzed and reported in the published literature [10], [15].

In our study, the choice of baseline for comparison is carefully considered within the context of ECF offering soft ratings, where diversity in baseline methods is limited. Specifically, we have selected the two- probability-focused ECF [15] as our baseline. This ECF variant has not only performed well in prior studies, but also exceeds the performance of earlier baselines, making it a pertinent choice for comparative analysis. The two-probability-focused ECF represents a more advanced iteration, reflecting both the evolution that addresses conflicting preferences in ECF [9] and the state-of-the-art in ECF research.

Additionally, a 10-fold cross-validation approach was adopted for the experiments. Initially, the ratings within the dataset were divided into 10 distinct folds, with each fold comprising a random selection of 10% of each user's ratings. The experimental process was repeated ten times; during each iteration, one fold was designated as the test dataset, while the other ratings were utilized for training purposes. The mean outcomes from these ten iterations are detailed in the subsequent part of this section.

In the field of ECF offering soft ratings, researchers have developed new evaluation methods capable of assessing their performance. These include DS-MAE, DS-Precision, DS-Recall and DS-Fscore [9], [15], [28]. Let $\hat{r}_{i,k}$ be the final estimated rating for user U_i and item I_k and $\widehat{BP}_{i,k}$ represent the pignistic-probability distribution of the mass function $\hat{r}_{i,k}$. The selected evaluation metrics are defined as follows:

$$DS-MAE(\theta_j)=\frac{1}{|D_j|}\sum_{(i,k)\in D_j,\theta_l\in\Theta}\widehat{Bp_{i,k}}(\theta_l)|\theta_j-\theta_i|$$

$$DS-Precision\ (\theta_j)=\frac{TP(\theta_j)}{TP(\theta_j)+FP(\theta_j)}$$

$$DS-Recall\ (\theta_j)=\frac{TP(\theta_j)}{TP(\theta_j)+FN(\theta_j)}$$

$$DS-F_i(\theta_j)=\frac{(i^2+1)(DS-Precision\ (\theta_j)\left(DS-Recall\ (\theta_j)\right)}{i^2(DS-Precision\ (\theta_j)+\left(DS-Recall\ (\theta_j)\right)}$$

where D_j is the test set identifying user-item pairs whose true evaluation is $\theta_j \in \Theta$ and:

$$TP(\theta_j)=\sum_{(i,k)\in D_j}\widehat{Bp_{i,k}}\ (\theta_j)$$

$$FP(\theta_j)=\sum_{(i,k)\in D_j,j\neq 1}\widehat{Bp_{i,k}}\ (\theta_l)$$

$$FN(\theta_j)=\sum_{(i,k)\in D_j}\widehat{Bp_{i,k}}\ (\theta_l)$$

4.1 Results for Movielens-100K Dataset

Tables 3 and 4 provide a comprehensive comparison of the CA-ECF and baseline method across various rating values, assessing their performance through hard metrics, such as MAE, precision, recall and F-score and soft metrics, such as DS-MAE, DS-precision, DS-recall and DS-F-score. The CA-ECF method demonstrates superior precision and recall in both soft and hard recommendations, particularly notable in the middle rating values, where it significantly outperforms the baseline. This trend is consistent across the precision and F-score metrics as well, with CA-ECF showing enhanced accuracy.

Table 3. Comparison in hard decisions for CA-ECF and baseline on Movielens-100K dataset.

Metrics	Rating value					Global
	1	2	3	4	5	
CA-ECF						
MAE	**2.4011**	**1.5072**	**0.7286**	**0.3495**	**1.0153**	**0.8326**
Precision	**0.182**	0.2208	**0.3221**	**0.3935**	**0.4648**	**0.3752**
Recall	**0.0177**	0.0912	**0.3167**	**0.6657**	**0.1863**	**0.3828**
F-score	**0.0322**	0.129	**0.3193**	**0.4946**	**0.2659**	**0.3789**
Baseline						
MAE	2.4075	1.5087	0.7382	0.369	1.0157	0.8343
Precision	0.177	**0.2242**	0.3206	0.3919	0.4484	0.3641
Recall	0.0152	**0.0924**	0.3158	0.6642	0.1851	0.3718
F-score	0.0649	**0.1434**	0.3175	0.4923	0.2592	0.3468

Table 4. Comparison in soft decisions for CA-ECF and baseline on Movielens-100K dataset.

DS-Metrics	Rating value					Global
	1	2	3	4	5	
CA-ECF						
DS-MAE	**2.4057**	**1.4897**	**0.7337**	**0.3702**	**1.0159**	**0.8243**
DS-Precision	**0.1756**	**0.2380**	**0.3191**	**0.3916**	0.4452	**0.36015**
DS-Recall	**0.0159**	**0.0962**	**0.3177**	**0.6612**	0.1787	**0.3722**
DS-F-score	**0.0292**	**0.1370**	**0.3184**	**0.4919**	0.2550	**0.3315**
Baseline						
DS-MAE	2.4066	1.4918	0.7344	0.3713	1.0175	0.8327
DS-Precision	0.1749	0.2300	0.3175	0.3908	**0.4462**	0.3609
DS-Recall	0.0156	0.0949	0.3164	0.6605	**0.1815**	0.3702
DS-F-score	0.0267	0.1329	0.3161	0.4903	**0.2560**	0.3315

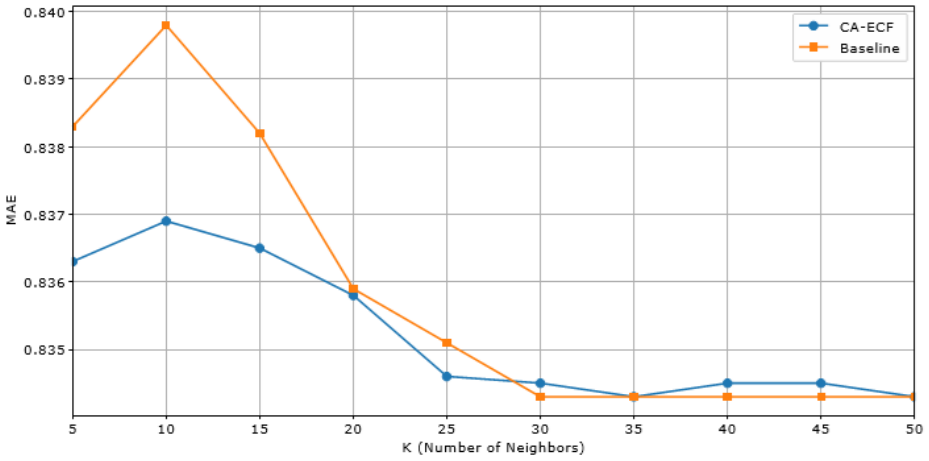


Figure 3. Overall MAE for CA-ECF *versus* baseline on Movielens-100K dataset.

In Figure 3, the data comparing CA-ECF and baseline across varying values of K reveal that both recommendation frameworks show an improvement in DS-MAE as the number of neighbors increases up to K = 20, beyond which the improvement in error rates stabilizes. CA-ECF consistently performs better than baseline at lower values of K, indicating its superior efficiency in scenarios with fewer neighbors. Both methods reach their optimal performance around K = 20. This indicates that increasing K beyond 20 offers no significant benefit, possibly leading to over-specialization and unnecessary computational overhead.

In Figure 4, both CA-ECF and baseline show a trend where the DS-MAE values generally decrease as K increases from 5 to 20. Around K = 20, both CA-ECF and baseline achieve their minimum DS-MAE values, indicating an optimal point for the Movielens dataset. Post this point, both frameworks stabilize, with slight fluctuations in DS-MAE values, suggesting that increasing K beyond this point does not significantly enhance the accuracy. CA-ECF appears to be more robust at lower neighborhood sizes, which could be advantageous in scenarios where the data is sparse or when it is computationally preferable to consider fewer neighbors.

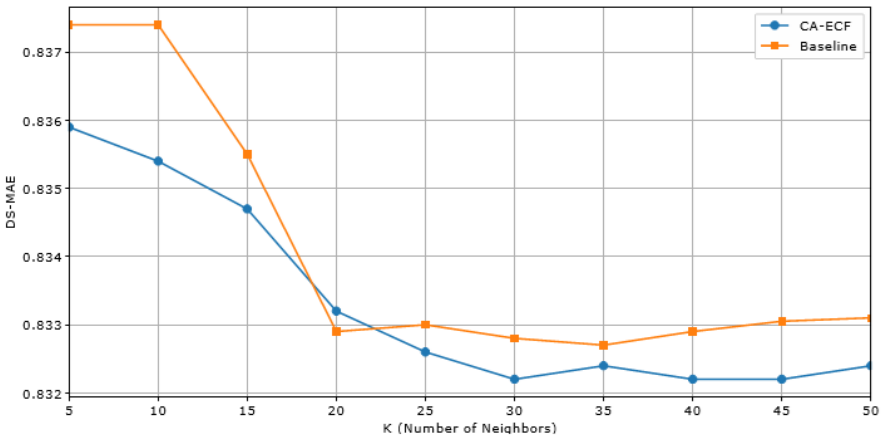


Figure 4. Overall DS-MAE for CA-ECF *versus* baseline on Movielens-100K dataset.

4.2 Results for Flixster Dataset

Based on the presented data from the hard and soft decision comparisons between CA-ECF and the baseline on the Flixster dataset, several insights emerge. As shown in Table 5, for hard decisions, CA-ECF exhibits consistently lower MAE across all rating values compared to the baseline, showcasing its superior accuracy in prediction. Notably, the global MAE for CA-ECF stands at 0.8281, which is lower than the baseline's 0.8503, underscoring the enhanced precision of CA-ECF in handling diverse rating scales from 0.5 to 5.0.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

Table 5. Comparison in hard decisions for CA-ECF and baseline on Flixster dataset.

Metrics	Rating value										Global
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
CA-ECF											
MAE	**3.2204**	**2.6843**	**2.1821**	**1.7530**	**1.2887**	**0.7529**	**0.4068**	**0.1341**	**0.5169**	**0.9056**	**0.8281**
Precision	**0.8790**	0	0	**0.1857**	**0.2727**	**0.2283**	**0.1976**	0.2091	**0.1769**	0.3856	**0.2860**
Recall	**0.0252**	0	0	**0.0017**	**0.0036**	**0.0703**	**0.1489**	0.7767	**0.0287**	0.0852	**0.3160**
F-score	**0.0489**	0	0	**0.0033**	**0.0071**	**0.1074**	**0.1698**	0.3294	**0.0493**	0.1395	**0.3002**
Baseline											
MAE	3.2708	2.7865	2.3006	1.7741	1.3163	0.7806	0.4204	0.1360	0.5264	0.9081	0.8503
Precision	0.8521	0	0	0.1697	0.2435	0.1975	0.1886	**0.2150**	0.1747	**0.3921**	0.2404
Recall	0.0242	0	0	0.0015	0.0031	0.0652	0.1478	**0.7812**	0.0280	**0.0867**	0.3114
F-score	0.0470	0	0	0.0029	0.0061	0.0980	0.1657	**0.3371**	0.0482	**0.1420**	0.2713

Similarly, in the soft-decision results of Table 6, CA-ECF maintains its edge over the baseline, with a global DS-MAE of 0.8190 against the baseline's 0.8381. This precision is further reflected in the metrics of DS-Precision, DS-Recall and DS-F-score, where CA-ECF consistently outperforms the baseline across most rating values, particularly in the mid to high range. These metrics confirm the robustness of CA-ECF in synthesizing evidential data to produce reliable and nuanced recommendations, highlighting its applicability in systems where user preferences are particularly conflicting. In the Flixster dataset, ratings of 1.0 and 1.5 are significantly less frequent compared to higher ratings. Consequently, the columns for ratings 1.0 and 1.5 in the comparison tables sometimes show values as zero, indicating sparse data in these categories.

Table 6. Comparison in soft decisions for CA-ECF and baseline on Flixster dataset.

DS-Metrics	Rating value										Global
	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0	
CA-ECF											
DS-MAE	**3.2137**	**2.6702**	**2.1787**	**1.7467**	**1.2861**	**0.7491**	**0.4002**	**0.1291**	**0.5126**	**0.9014**	**0.8190**
DS-Precision	**0.8702**	**0**	**0**	**0.1767**	**0.2543**	**0.2056**	**0.1998**	0.2165	**0.2036**	0.3891	**0.2561**
DS-Recall	**0.0282**	0	**0**	**0.0018**	**0.0045**	**0.0710**	**0.1502**	0.7689	**0.0312**	0.0821	**0.3189**
DS-F-score	**0.0546**	**0**	**0**	**0.0035**	**0.0088**	**0.1055**	**0.1714**	0.3378	**0.0541**	0.1355	**0.2840**
Baseline											
DS-MAE	3.2360	2.7653	2.2781	1.7482	1.2909	0.7665	0.4187	0.1322	0.5202	0.9061	0.8381
DS-Precision	0.8562	0	0	0.1710	0.2435	0.1998	0.1906	**0.2172**	0.1872	**0.3987**	0.2468
DS-Recall	0.0253	0	0	0.0016	0.0037	0.0667	0.1491	**0.7851**	0.0290	**0.0892**	0.3114
DS-F-score	0.0491	0	0	0.0031	0.0072	0.1000	0.1673	**0.3402**	0.0502	**0.1457**	0.2753

Figure 5 depicts the MAE performance of CA-ECF and the baseline across varying neighborhood sizes (K) on the Flixster dataset. For CA-ECF, there is a consistent enhancement in performance across all K values, showcasing its robustness in managing different neighborhood sizes effectively. In contrast, the baseline exhibits a reduction in MAE as the number of neighbors increases, reaching a plateau at $K = 35$. Beyond this point, no significant gains are observed, indicating that larger neighborhoods do not further contribute to accuracy improvements. This data highlights CA-ECF's superior efficiency, particularly notable at smaller neighborhood sizes.

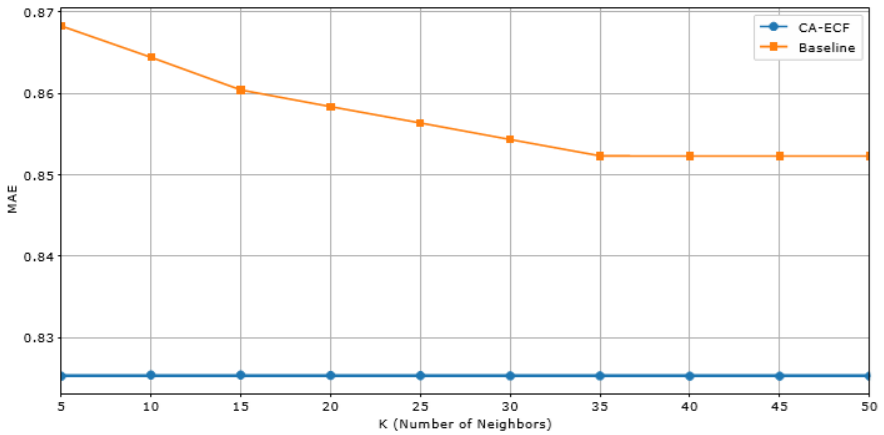


Figure 5. Overall MAE for CA-ECF *versus* baseline on Flixster dataset.

Figure 6 demonstrates the performance trend of CA-ECF and the baseline as the number of neighbors (K) increases within the Flixster dataset. Both frameworks exhibit distinct performance trends. CA-ECF demonstrates stable performance with consistently low DS-MAE values across all K values. In contrast, the baseline framework shows a decrease in DS-MAE from $K = 5$ to $K = 15$, suggesting that accuracy improves with a larger neighborhood up to this point. At $K = 15$, the baseline reaches its lowest DS- MAE, indicating an optimal balance between neighborhood size and predictive accuracy. Beyond $K = 15$, the baseline exhibits negligible improvements and slight fluctuations in DS-MAE, signaling that further increases in K do not yield substantial benefits and may lead to diminishing returns.

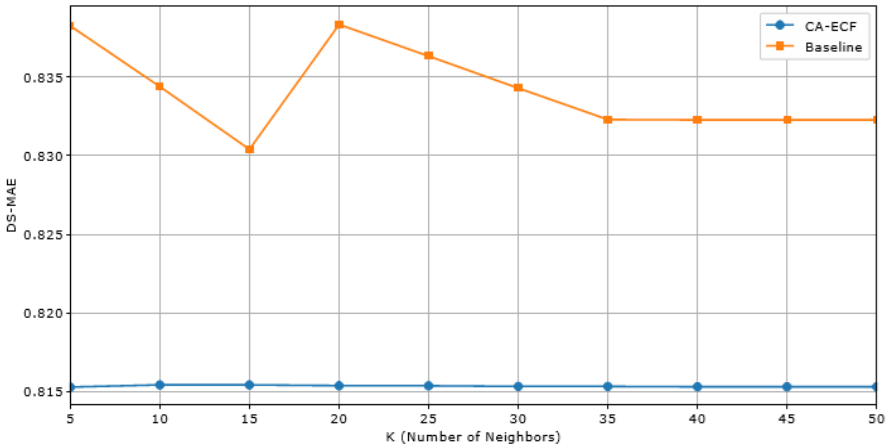


Figure 6. Overall MAE for CA-ECF *versus* baseline on Flixster dataset.

5. DISCUSSION

In the evaluation of the CA-ECF framework, our experiments reveal that the framework demonstrates notably improved performance for rating values that exhibit higher density within the dataset. This enhanced performance can be attributed to the framework's use of reliability factors that judiciously discount ratings. By adjusting the influence of these ratings in the evidential fusion step, the framework not only refines the prediction accuracy, but also effectively manages the inherent uncertainty associated with sparse data. Such a mechanism ensures that the contributions of the neighbors are weighted, which is particularly crucial in sparse datasets where every rating can significantly influence the outcome. This approach underscores the ability of CA-ECF to deliver more reliable and precise recommendations by effectively capturing and utilizing the underlying patterns in user-item interactions.

However, the CA-ECF framework introduces additional computational complexities, primarily from the calculation of Deng's relative entropy and the optimization of reliability factors. The computation of Deng's relative entropy within the neighborhood set $knn_{i,k}$ presents a quadratic complexity, $O(k^2)$, where k represents the number of neighbors who have rated the target item. Further complexity arises during the optimization step to determine optimal reliability factors, potentially extending to $O(k^3)$ depending on the algorithm used. In contrast, classical ECF methodologies typically involve linear operations based on similarity scores, resulting in a lower overall time complexity of $O(n)$. Thus, while the CA-ECF framework incurs a higher computational cost, it leverages this complexity to enhance the accuracy and reliability of recommendations, which is particularly advantageous in applications where the quality of recommendations is critical.

6. CONCLUSIONS

This research introduces a novel Conflict-Aware Evidential Collaborative Filtering framework that significantly advances the management of conflicts in user ratings. By integrating a two-probabilities-focused approach with the advanced conflict-resolution technique based on the Best Worst Method, the framework refines the weighting of user preferences. This precision in handling ratings leads to discernibly improved performance across key metrics, including DS-MAE, DS-precision, DS-recall and DS-F-score, outperforming existing methodologies. While our framework enhances recommendation accuracy and reliability, especially in handling uncertain, imprecise or incomplete user preferences, it introduces complexities related to its computational demands. The detailed calculations required by

Dempster-Shafer theory, along with those needed to optimize reliability factors, can lead to increased computational time, potentially limiting its immediate practicality in real-world scenarios. Looking forward, we plan to explore the potential of distributed computing and Monte Carlo approximations to manage the computational overhead effectively. These techniques aim to reduce the computational intensity while maintaining accuracy, offering scalable solutions for large datasets. We are also keen on investigating alternative fusion rules that can further enhance the framework's ability to handle conflicts. These steps are aimed at extending the scalability and practicality of our framework.

DECLARATION OF COMPETING INTERESTS

The authors declare that they don't have any competing financial interests or personal connections that could have influenced the work reported in this paper.

ACKNOWLEDGEMENTS

We express our gratitude to the Editor-in-Chief and the anonymous reviewers for their dedication and insightful feedback on our manuscript.

REFERENCES

[1] P. B. Anand and R. Nath, "Content-based recommender systems," Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries, pp. 165–195, 2020.

[2] H. Papadakis, A. Papagrigoriou, C. Panagiotakis, E. Kosmas and P. Fragopoulou, "Collaborative Filtering Recommender Systems Taxonomy," Knowledge and Information Systems, vol. 64, no. 1, pp. 35–74, 2022.

[3] R. Seth and A. Sharaff, "A Comparative Overview of Hybrid Recommender Systems: Review, Challenges and Prospects," Chapter 3 in Book: Data Mining and Machine Learning Applications, pp. 57–98, DOI: 10.1002/9781119792529.ch3, 2022.

[4] Y. Wang, P. Wang, Z. Liu and L. Y. Zhang, "A New Item Similarity Based on α-divergence for Collaborative Filtering in Sparse Data," Expert Systems with Applications, vol. 166, p. 114074, 2021.

[5] R. Abdelkhalek, I. Boukhris and Z. Elouedi, "Towards More Trustworthy Predictions: A Hybrid Evidential Movie Recommender System," JUCS: Journal of Universal Computer Science, vol. 285, no. 10, pp. 1003–1029, 2022.

[6] K. Belmessous, F. Sebbak, M. Mataoui and W. Cherifi, "A New Uncertainty-aware Similarity for User-based Collaborative Filtering," Kybernetika, vol. 60, no. 4, pp. 446–474 , 2024.

[7] N. Idrissi and A. Zellou, "A Systematic Literature Review of Sparsity Issues in Recommender Systems," Social Network Analysis and Mining, vol. 10, no. 1, pp. 1–23, 2020.

[8] K. Belmessous, F. Sebbak, M. Mataoui and A. Batouche, "Co-rating Aware Evidential User-based Collaborative Filtering Recommender System," Proc. of the Int. Conf. on Computing Systems and Applications, Advances in Computing Systems and Applications (CSA 2022), vol. 513, pp. 51–60, Springer, 2022.

[9] K. Belmessous, F. Sebbak, M. Mataoui, M. R. Senouci and W. Cherifi, "Dempster-Shafer Theory in Recommender Systems: A Survey," Int. Journal of Uncertainty, Fuzziness and Knowledge-based Systems, vol. 32, no. 5, pp. 747–780, 2024.

[10] V.-D. Nguyen and V.-N. Huynh, "Evidence Combination Focusing on Significant Focal Elements for Recommender Systems," Proc. of the Int. Symposium on Integrated Uncertainty in Knowledge Modeling and Decision Making, vol. 9376, pp. 290–302, Springer, 2015.

[11] V.-D. Nguyen and V.-N. Huynh, "A Community-based Collaborative Filtering System Dealing with Sparsity Problem and Data Imperfections," Proc. of Pacific Rim Int. Conf. on Artificial Intelligence, Trends in Artificial Intelligence (PRICAI 2014), vol. 8862, pp. 884–890, Springer, 2014.

[12] V.-D. Nguyen and V.-N. Huynh, "On Information Fusion in Recommender Systems Based on Dempster-Shafer Theory," Proc. of the 2016 IEEE 28th Int. Conf. on Tools with Artificial Intelligence (ICTAI), pp. 78–85, San Jose, CA, USA, 2016.

[13] F. Sebbak, M. R. Senouci, F. Benhammadi et al., "Towards Cardinality-aware Evidential Combination Rules in Dempster–Shafer Theory," KI-Künstliche Intelligenz, pp. 1–16, DOI: 10.1007/s13218-024-00859-4, 2024.

[14] K. Zhao, L. Li, Z. Chen, R. Sun, G. Yuan and J. Li, "A Survey: Optimization and Applications of Evidence Fusion Algorithm Based on Dempster–Shafer Theory," Applied Soft Computing, vol. 124, p. 109075, 2022.

[15] V.-D. Nguyen and V.-N. Huynh, "Two-probabilities Focused Combination in Recommender Systems," Int. Journal of Approximate Reasoning, vol. 80, pp. 225–238, 2017.

[16] N. Bahri, M. A. Bach Tobji and B. Ben Yaghlane, "ECFAR: A Rule-based Collaborative Filtering System Dealing with Evidential Data," Proc. of the Int. Conf. on Intelligent Systems Design and Applications, Intelligent Systems Design and Applications (ISDA 2021), vol. 418, pp. 944–955, Springer, 2022.

[17] L. Zhou, H. Cui, X. Mi, J. Zhang and B. Kang, "A Novel Conflict Management Considering the Optimal Discounting Weights Using the BWM Method in Dempster-Shafer Evidence Theory," Information Sciences, vol. 612, pp. 536–552, 2022.

[18] A. P. Dempster, "Upper and Lower Probabilities Induced by a Multi-valued Mapping," Chapter in Book: Classic Works of the Dempster-Shafer Theory of Belief Functions, pp. 57–72, Springer, 2008.

[19] G. Shafer, A Mathematical Theory of Evidence, vol. 42, DOI: 10.2307/j.ctv10vm1qb, Princeton University Press, 1976.

[20] P. Smets, "The Combination of Evidence in the Transferable Belief Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 5, pp. 447–458, 1990.

[21] L. A. Zadeh, "A Simple View of the Dempster-Shafer Theory of Evidence and Its Implication for the Rule of Combination," AI Magazine, vol. 7, no. 2, pp. 85–85, 1986.

[22] A.-L. Jousselme, D. Grenier and É. Bossé, "A New Distance between Two Bodies of Evidence," Information Fusion, vol. 2, no. 2, pp. 91–101, 2001.

[23] Y. Song, X. Wang, J. Zhu and L. Lei, "Sensor Dynamic Reliability Evaluation Based on Evidence Theory and Intuitionistic Fuzzy Sets," Applied Intelligence, vol. 48, no. 11, pp. 3950–3962, 2018.

[24] W. Jiang, "A Correlation Coefficient for Belief Functions," Int. Journal of Approximate Reasoning, vol. 103, pp. 94–106, 2018.

[25] F. Xiao, Z. Cao and A. Jolfaei, "A Novel Conflict Measurement in Decision-making and Its Application in Fault Diagnosis," IEEE Transactions on Fuzzy Systems, vol. 29, no. 1, pp. 186–197, 2020.

[26] J. Rezaei, "Best-worst Multi-criteria Decision-making Method," Omega, vol. 53, pp. 49–57, 2015.

[27] L. Fei and Y. Deng, "A New Divergence Measure for Basic Probability Assignment and Its Applications in Extremely Uncertain Environments," Int. Journal of Intelligent Systems, vol. 34, no. 4, pp. 584–600, 2019.

[28] T. L. Wickramarathne, K. Premaratne, M. Kubat and D. Jayaweera, "CoFIDS: A Belief-theoretic Approach for Automated Collaborative Filtering," IEEE Transactions on Knowledge and Data Engineering, vol. 23, no. 2, pp. 175–189, 2010.

[29] V.-D. Nguyen and V.-N. Huynh, "A Reliably Weighted Collaborative Filtering System," Proc. of the European Conf. on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU 2015), vol. 9161, pp. 429–439, Springer, 2015.

[30] V.-D. Nguyen and V.-N. Huynh, "Integrating with Social Network to Enhance Recommender System Based on Dempster-Shafer Theory," Proc. of the Int. Conf. on Computational Social Networks (CSoNet 2016), pp. 170– 181, Springer, 2016.

[31] V.-D. Nguyen and V.-N. Huynh, "Noise-Averse Combination Method," Proc. of the 2016 IEEE 28th Int. Conf. on Tools with Artificial Intelligence (ICTAI), pp. 86–90, San Jose, CA, USA, 2016.

[32] V.-D. Nguyen, S. Sriboonchitta and V.-N. Huynh, "Using Community Preference for Overcoming Sparsity and Cold-start Problems in Collaborative Filtering System Offering Soft Ratings," Electronic Commerce Research and Applications, vol. 26, pp. 101–108, 2017.

[33] V.-D. Nguyen, V.-N. Huynh and S. Sriboonchitta, "Integrating Community Context Information into a Reliably Weighted Collaborative Filtering System Using Soft Ratings," IEEE Transactions on Systems, Man and Cybernetics: Systems, vol. 50, no. 4, pp. 1318–1330, 2017.

[34] Y. Dong, X. Li and Y. Liu, "A Fast Combination Method in DSMT and Its Application to Recommender System," PloS One, vol. 13, no. 1, p. e0189703, 2018.

[35] I. Bloch, "Some Aspects of Dempster-Shafer Evidence Theory for Classification of Multi-modality Medical Images Taking Partial Volume Effect into Account," Pattern Recognition Letters, vol. 17, no. 8, pp. 905–919, 1996.

[36] F. M. Harper and J. A. Konstan, "The Movielens Datasets: History and Context," ACM Transactions on Interactive Intelligent Systems (TIIS), vol. 5, no. 4, pp. 1–19, 2015.

ملخص البحث:

تقدّم هذه الورقة إطارَ عملٍ مُبتكراً لتحسين التّصفية التّعاونية المستندة إلى الأدلّة، يمثّل نظام توصيةٍ حاسماً مصمَّماً خِصّيصاً للمجالات الحسّاسة، مثل الرّعاية الصّحية وتتبُّع الأهداف وغيرهما.

يركّز إطار العمل المقترح على تصفيةِ الكيفية الّتي يتمُّ بها التّعامل مع التّفْضيلات المتناقضة للمستخدمين الّتي تشتمل على مواضِع خللٍ، وبخاصّةٍ من حيث إدارة التّفْضيلات المتناقضة في أثناء انتقاء تفْضيلات الحِوار، وذلك بُغية تحسين جودة التّوصْية.

يجمع النّظام المقترح بين نهجٍ قائمٍ على احتمالين وتقنية متقدّمةٍ لإدارة التّناقض، الأمر الّذي من شأنه تخصيص وزْنٍ موثوقيةٍ أكثر دقّةً لكلّ مُسْتخدم، وذلك عبر استخدام طريقة (Deng) النّسبية وطريقة "الأفضل-الأسوأ". والجديرُ بالذّكر أنّ ذلك يحسِّن من اختيار التّفْضيلات في نظام التّصْفية التّعاونية المستندِ إلى الأدلة.

تمّ تجريب إطار العمل المقترح على عددٍ من مجموعات البيانات، وبيّنت النّتائج أنّه يتفوّقُ على مجموعةٍ من أُطُر العمل المشابهة الواردة في أدبيات الموضوع، وذلك عند مقارنته معها من حيث خطأ التّوقُّع والدّقة والضّبط وغير ذلك من مقاييس الأداء.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

STATE-OF-THE-ART OF MACHINE LEARNING IN NEURO DEVELOPMENT DISORDER: A SYSTEMATIC REVIEW

Lilian Lee Yen Wei [1,2], Ag Asri Ag Ibrahim[1] and Rayner Alfred[1]

(Received: 15-Jul.-2024, Revised: 18-Sep.-2024, Accepted: 2-Oct.-2024)

ABSTRACT

This paper presents a comprehensive literature review focusing on the utilization of machine-learning (ML) and deep-learning (DL) methods for predicting and detecting Neurodevelopmental Disorders (NDDs), such as Intellectual Disability (ID), Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), Dyslexia, among others. While existing reviews often lack detailed discussions on the specific ML algorithms, datasets and performance metrics employed in NDD prediction and detection, this study aims to address this gap by examining two primary aspects: prediction and detection. Objective: The objective of this study is to investigate the current state-of-the-art methodologies, challenges and future directions in leveraging ML and DL techniques for the prediction and detection of NDDs. It aims to categorize the literature based on these two major aspects and provide insights into the various approaches, datasets, parameters and performance measures used in previous research. Methodology: This review encompasses articles published in journals and conference proceedings indexed in Scopus from 2013 to 2023. The search employed terms such as "Predicting Neurodevelopmental Disorder" and/or "Detection of Disorder Using Machine Learning." The analysis focuses on identifying common ML and DL approaches, ensemble models, types of datasets utilized, as well as the parameters and performance metrics employed in NDD-prediction and detection studies. Results: The findings of this review shed light on prevalent ML and DL methodologies, the challenges encountered and potential avenues for future research aimed at enhancing services for the NDD community through improved prediction and detection techniques.

KEYWORDS

Detection, Prediction, Classification, Deep learning, Machine learning, Mental health, Neurodevelopment disorders.

1. INTRODUCTION

Neurodevelopmental Disorders (NDDs), as outlined in the DSM V Diagnostic and Statistical Manual by the American Psychiatric Association, encompass a range of conditions affecting the development of the central nervous system [1]. These conditions manifest in difficulties in behaviours, cognition, social interaction and emotional functioning. Included within NDDs are intellectual disability (ID), communication disorders, Autism Spectrum Disorder (ASD), Attention-Deficit/Hyperactivity Disorder (ADHD), neurodevelopmental motor disorders such as Tic Disorders and Specific Learning Disorders [2]. Despite the prevalence of NDDs, which affects roughly 17% of the general population, many individuals may remain undiagnosed. Factors contributing to NDDs include maternal and fetal genotype, early environmental influences and some causes that are still not fully understood. Particularly concerning is the rising prevalence of NDDs, with autism rates reported by the Centre for Disease Control and Prevalence (CDC) increasing from 1 in 150 children in 2000 to 1 in 36 presently, with around 40% of affected individuals also experiencing ADHD and other comorbidities [3]. NDDs represent a significant mental-health category with profound impacts on daily functioning, potentially jeopardizing the physical and mental well-being of affected individuals as they transition from childhood to adulthood. Given the increasing frequency of NDDs and their substantial impact, it is imperative to address the challenges associated with early identification and intervention. Developing a rapid, reliable and automated method for identifying early signs of mental-health issues is critical in this rapidly evolving world.

Hence, we conduct a systematic review encompassing medical and computer-science literature on the

1. Lilian L.Y. W., A. A. A. Ibrahim and R. Alfred are with Faculty of Computing and Informatics, Universiti Malaysia Sabah, 88400, Kota Kinabalu, Sabah, Malaysia. Emails: lilian_lee_di22@iluv.ums.edu.my, awgasri@ums.edu.my, ralfred@ums.edu.my
2. Lilian L.Y.W. is with Malaysian Administration Modernization and Management Planning Unit, Prime Minister's Department, Malaysia. Email: lilian.phoebe@gmail.com

detection of NDD issues using machine-learning (AI) methodologies. This review spans articles published between 2013 and 2023 sourced from databases such as Scopus, IEEE Explore, PubMed and Web of Science (WOS). Utilizing the Preferred Reporting Items for Systematic Reviews and Meta- Analysis (PRISMA) methodology, we meticulously selected 81 articles from an initial pool of 811.

Our review highlights a notable research gap concerning the utilization of machine learning in interventions for neurodevelopmental disorders (NDDs), particularly in the domain of automated neuro-feedback. We also explore the machine learning techniques utilized in developing EEG-based detection methods for NDDs. Furthermore, we conduct a thorough examination of challenges outlined in existing literature and provide forward-looking recommendations. These recommendations encompass various facets, such as data-fusion techniques, integrating hybrid classification models, emphasizing the importance of publicly available datasets, addressing uncertainties in model predictions, enhancing model interpretability and devising strategies for hardware implementation. Essentially, our systematic review illuminates the current landscape of machine-learning applications in NDD detection and intervention, while also charting a course for future research aimed at bridging existing gaps and overcoming challenges. The objective of this review is to guide future researchers in adopting potential trends or models that can significantly enhance the diagnosis and detection of NDDs.

Responding promptly to a diagnosis is crucial for minimizing the time required for intervention once a diagnosis is detected. Machine-learning techniques can assimilate and analyze integrated data from multiple sources, including population statistics, lifestyle factors and medical records to predict the occurrence and distribution of diagnoses within a specific area. Medical practitioners can utilize machine-learning methods to enhance the implementation of existing interventions and speed up in developing new interventions. For example, deep-learning algorithms can be employed to analyze extensive datasets comprising medical information gathered from hospitals. For example, clinical test data from patients diagnosed with mental health can be utilized as input for machine-learning models, enabling doctors to expedite diagnoses. This research endeavours to explore the current advancements, obstacles and prospective directions in leveraging machine-learning techniques for managing neurodevelopmental disorders, as outlined in the two previously mentioned categories. The study uses the same method of systematic review conducted by Rayner & Obit, 2021, the roles of machine learning methods in limiting the spread of deadly diseases. Thus, the work here is to conduct a comprehensive review of different methodologies, dataset types, parameters or variables, individual and ensemble models, performance metrics and approaches employed in prior studies [4].

We categorized all articles and conference papers based on Scopus Indexed – whether pertaining to prediction or detection strategies. This review's results center on frequently employed machine learning methods, obstacles encountered and future directions aimed at supporting intervention and therapy for neurodevelopmental disorders through both detection and prediction. The trend and distribution of objectives for machine learning and recent works for NDD detection are described in Figure 1.

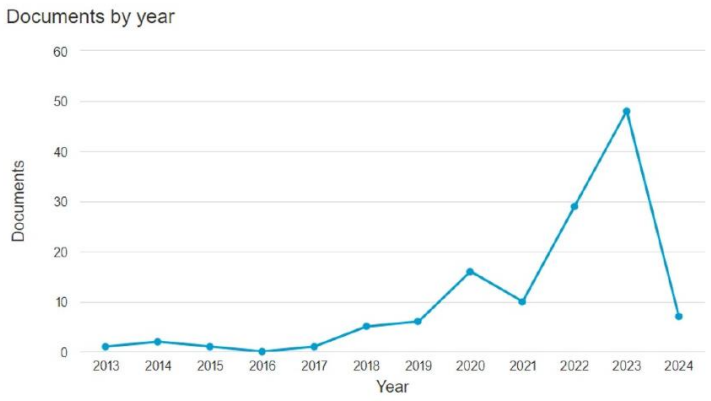


Figure 1. The trend and distribution of objectives for machine learning and recent works for NDD detection.

2. METHODOLOGY

The purpose of this Systematic Literature Review (SLR) is to conduct a sequential process of PRISMA to make available research applicable to machine-learning approaches in assisting medical health diagnosing Neurodevelopment Disorders. Four primary stages of PRISMA are identified to be included in this SLR as shown in Figure 2. They are called: Identification, Screening, Eligibility and Inclusion [5].

2.1 Content Retrieval

Apart from adhering to the PRISMA stages, this literature review underwent two distinct phases: planning and conducting. The initial phase is geared towards defining the prerequisites for a systematic review while mitigating potential researcher biases. It involves crafting a comprehensive review protocol, acting as a blueprint for conducting an unbiased review process. Key elements of this proposed review protocol in our study include delineating research questions, formulating a search strategy to pinpoint relevant studies, specifying inclusion and exclusion criteria, establishing a method for assessing study quality and extracting and synthesizing data, all of which will be detailed in the subsequent section. The planning phase involves crafting research questions centered on employing machine-learning techniques for predicting and detecting neurodevelopmental disorders, followed by setting up suitable search procedures to efficiently execute the research activities. During the conducting phase, several actions are taken, including setting predetermined selection criteria to pinpoint relevant studies and assessing their quality using the predefined quality-assessment procedure outlined in this study. This phase involves extracting information from the selected studies and conducting data synthesis to provide a succinct summary of the reviews. These processes are visually depicted in Figure 2, facilitating the incorporation of new information into the report in the future.

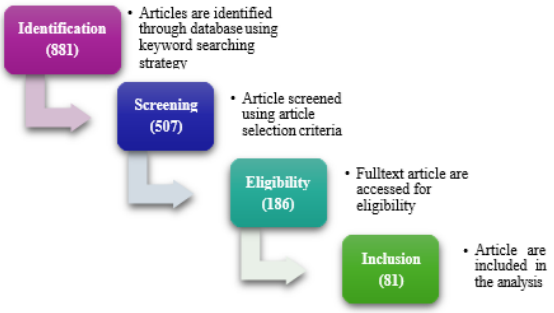


Figure 2. PRISMA method.

2.1.1 Formulating Research Questions

The research questions (RQs) were structured to define the study's boundaries from three distinct angles: population, intervention and outcomes [5]. From the population perspective, the focus is on the domains or functions affected by the intervention, such as detection, prediction and classification. These populations may relate to various aspects, including specific machine-learning methods or types of machine learning models and their applications. The intervention viewpoint centers on machine-learning approaches addressing specific challenges, such as diagnosing, detecting and predicting Neurodevelopment Disorders. Lastly, the outcomes perspective concerns factors significant to practitioners, such as improved prediction accuracy, reduced diagnostic costs for specific disorders and shortened response time in detecting potentially severe disorders. All relevant outcomes must be explicitly stated. For instance, interventions may aim to enhance one aspect of NDD prediction without affecting another, such as improving reliability without increasing costs. The primary goal of this Systematic Literature Review (SLR) is to gather and scrutinize relevant evidence to tackle the defined research questions (RQs). Our motivation for undertaking this endeavor is to obtain responses to a series of seven RQs, aiming to gain deeper insights into key aspects of our research focus. This entails enhancing our comprehension of the roles played by machine-learning technologies in facilitating the prediction and detection of Neurodevelopment Disorders, as well as identifying research constraints to guide future-research directions. The RQs and their rationale are thoroughly elaborated in Table 1.

Table 1. Research questions.

ID	*Research Question*
RQ1	What are the roles of machine-learning models in assisting in screening neurodevelopment disorders?
RQ2	What types of NDD datasets in previous works have been used to build the models? What types of parameters or variables have been used?
RQ3	What types of problems addressed using these models?
RQ4	Which individual models achieved the highest performance?
RQ5	What evaluation metrics and methods are employed to measure the performance of the machine-learning models?
RQ6	What types of ensemble methods are used in machine-learning models?
RQ7	What types of deep-learning approaches used in NDD Detection?

2.1.2 Search Process

In the identification stage, all publications up to Dec. 2023 were compiled from searches made in Scopus, IEEE Explore, PubMed and Web of Science (WOS) databases. The retrieval was performed for articles from journals and conference proceedings published from 2010 to 2023 using the following Boolean search expression: "Prediction" OR "Detection" OR "Classification" OR "Diagnosing" OR "Identification" AND "ADHD" OR "AUTISM" OR "DYSLEXIA" OR "Neuro Development Disorder" AND "Artificial Intelligence" OR "Machine Learning" OR "Deep Learning".

Only final papers were considered in this review. The inclusion and exclusion criteria are shown in Table 2. The search process is designed to thoroughly address all predefined research questions. This involves selecting appropriate digital libraries, setting a time frame for the published articles and defining the search keywords. We will explore six of the most popular and largest online digital libraries in computer science, along with the Medline digital library, which publishes peer-reviewed articles. These digital libraries are listed in Table 3.

Table 2. Criteria of inclusion and exclusion.

Criteria	**Inclusion**	**Exclusion**
Type of Article	Journal articles	Others (Thesis, Handbook, Literature Review and Survey Paper)
Language	English	Non-English
Subjects Covered	Computer Science, Neuroscience, Health Professional and Psychology.	Multidisciplinary
Year of Publications	2013-2023	< 2013
Domain	Mental Health	Other Disorder or Comorbidities
Mental	Neuro Development Disorder (ASD, ADHD, Dyslexia)	Other neurodisorders
Health		Listed in DSM-V (Schizophrenia, psychotic, bipolar, depression, …etc.)
ML Models	Traditional, Deep Learning and Ensemble Model	Transfer Learning
Dataset	• Demographic, Medical, Observation & Behavioural Data	Genetic Data
Type	• Facial Image Data • Eye-Tracking Data • EEG-based Data • Functional Magnetic Resonance Imaging (BMRI) and Functional Magnetic Resonance Imaging (FMRI) Data	Heart-rate Data Handwriting Data Speech Data

Table 3. Online digital libraries and number of studies screened and reviewed.

No.	*Database*	*URL*	*Screened*	*Eligible*	*Inclusion*
1	Elsevier	https://www.sciencedirect.com/	68	14	7
2	Springer	https://link.springer.com/	56	28	2
3	IEEE eXplore	https://ieeexplore.ieee.org/	197	73	35
4	MDPI	https://www.mdpi.com/	92	29	9
5	Wiley	https://onlinelibrary.wiley.com/	12	5	2
6	Medline (PubMed)	https://pubmed.ncbi.nlm.nih.gov/	82	37	26
	TOTAL		507	186	81

Furthermore, we reviewed various independent relevant journals and conference proceedings in the field of artificial intelligence, as detailed in Table 3. The search is limited to articles published between 2013 and 2023. This time frame was chosen, because machine learning has been extensively applied to problems related to Neurodevelopment Disorder (NDD) since the 2010s. Table 4 lists the number of studies reviewed based on year (2013 - 2023). Therefore, this paper aims to systematically summarize artificial-intelligence methodologies, encompassing both machine-learning and deep-learning techniques, applied in the prediction and detection in-response to neurodevelopmental disorders (NDDs).

Table 4. Number of studies reviewed based on year (2013 - 2023).

Year	Studies
2013 - 2017	6
2018	3
2019	4
2020	9
2021	14
2022	29
2023	16
Total	81

Table 5. Type of machine-learning problems and related studies.

Problems	*Roles*	*Related Studies*
Regression	Predict Neurodevelopment Disorder	[6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23]
Classification	Detect Neurodevelopment Disorder	[24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71], [72], [73], [74], [75], [76], [77], [78], [79], [80], [81], [82], [83], [84], [85], [86]

2.2 Content Summarization

Quantitative data was extracted from the chosen studies focusing on the research questions outlined and the findings are presented in Tables 5-12 and Figure 3.

2.2.1 Roles of Machine-learning Model

Predicting and detecting NDDs plays crucial roles in enhancing healthcare systems. The tasks primarily involve predicting NDDs or modeling disorder frequencies using regression methods. Conversely, machine learning models predominantly address classification problems in detecting NDD. Table 5 summarizes all studies focused on NDD prediction and detection.

2.2.2 Type of Datasets and Parameters Used

Table 6 summarizes structured data utilized for predicting and detecting the diagnosis of NDDs, as well as the number of studies conducted. From the collective findings of these studies, six sets of structured datasets were the most employed. Structured databases encompass Demographic Data, Medical Data, Observation & Behavioral Data, Visual Video Data, Meteorological Data, MRI (including BMRI and FMRI), face-expression data, eye-tracking data and EEG-based data. Demographic data includes information on Age, Gender, Race and Ethnicity. Medical data involves a systematic analysis of a child's conditions, incorporating parameters, such as head measurements, weight, height, signs and symptoms of the disorder and treatment information. Observation & Behavioral Data entail numerical representations obtained from responses, speech, cognitive abilities, quotient scores and questionnaire assessments, like M-CHART, Q-Chat, AQ-10, ADI-R and ADOS Screening, UCI repository, IQ score, NCHS survey data, SDS ASDTest, OBTest, UK's National Health Service (NHS), PAAS India and Scale data Questionnaires from Germany Clinic. Some of the observation and behavioral subjects were captured in video for further investigation. Visual Video Data captures activities involving a child during intervention or therapy sessions, focusing on

parameters, such as movement, behavior, sensory perception, angle, direction and speed. Magnetic Resonance Imaging Data (BMRI and FMRI) aids in detecting and monitoring brain characteristics, particularly changes in blood flow. Some BMRI/FMRI datasets are publicly available through ABIDE (ADHD-200, Craddock 400 (CC400), …etc.). Common parameters for detection include normalized region volume, reduced corpus callosum volume and increased amygdala volume. Facial Image Data analyzes emotions through facial expressions, categorizing them as Happy, Sad, Angry or Neutral. Facial analysis employs an arousal- valence model to assess parameters, such as positive-active and negative-passive readings. These datasets can be collected through public databases, like Kaggle (KDEF dataset, …etc.). Finally, EEG-based Data, recorded *via* devices like BCI, captures spectral power of EEG signals, including Beta, Alpha, Theta, Gamma and Delta waves. Analysis often involves spectral, temporal, spatial or time-frequency features, revealing specific brain activities, ERPs, recurrences and transitional states. The EEG dataset is available in Dataport IEEE, …etc. Eye-tracking Data collected through EyeGaze apps, Eye movement, Automatic retinal-image analysis (ARIA) monitors parameters, like retina movement and pupil size.

2.3 Reporting of Review Findings

The summary of findings in the review was derived from the selected studies, focusing on the defined research questions. The overall State-of-the-Art of Neurodevelopment Disorder Prediction and Detection is described in Figure 3.

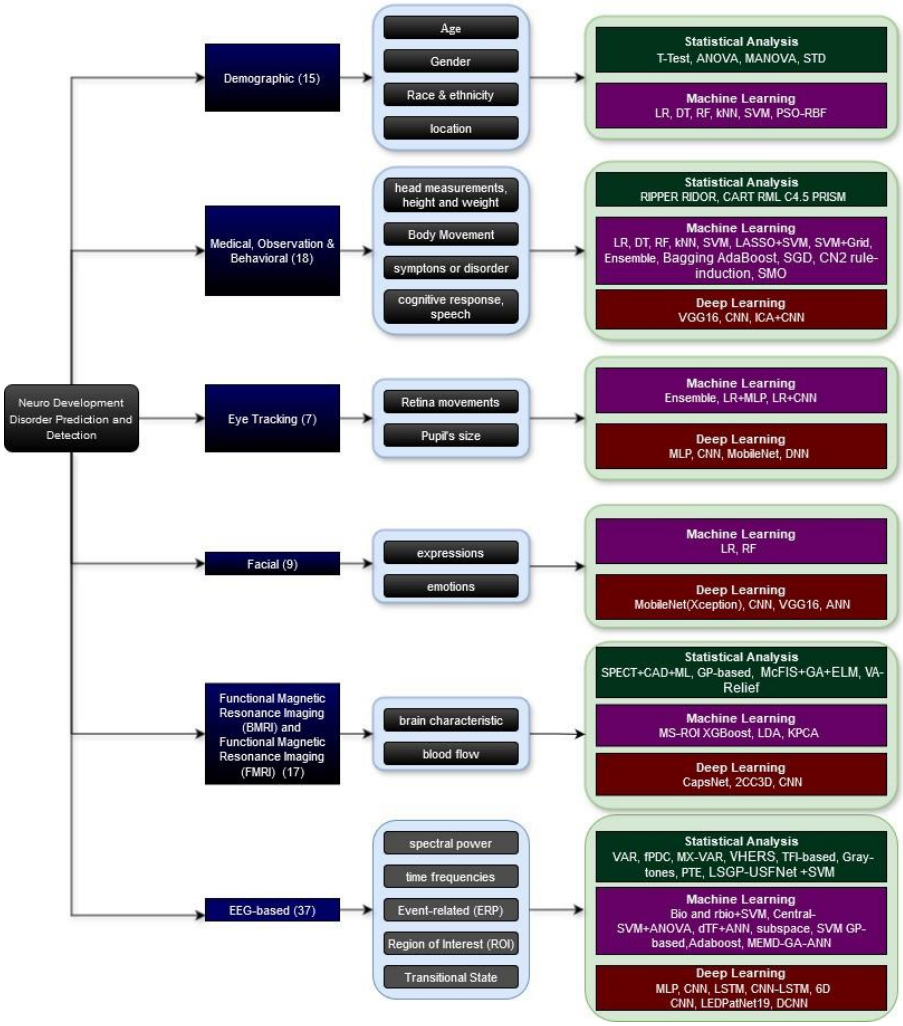


Figure 3. State-of-the-art of neurodevelopment disorder prediction and detection.

2.3.1 Roles of Machine-learning Models

This sub-section summarizes and discusses the findings of RQ1: What are the roles of machine-learning models in assisting in screening neurodevelopment disorders? The roles of machine-learning

models can be categorized into regression and classification.

Table 6. Datasets and parameters used.

Databases	*(Frequency) Features*
Demographic Data (18)	Age, Gender, Race, Ethnicity, Location
Medical Observation & Behavioral Data (16)	Head measurements, Weight, Height, Response, Speech, Cognitive Ability, Quotients scores, Movement, Behavior, Sensory, Angle, Direction, Speed, …etc.
BMRI/FMRI Data (13)	Brain and Functional Magnetic Resonance Imaging: Changing in the blood flow within brain
Facial Image Data (7)	Facial features and expression that can be detect through facial emotion: Happy, Sad, Angry, Null
EEG-based Data (26)	EEG reading: Beta, Alpha, Theta, Gamma, Delta Other features (spectral, time, spatial or time-frequency features, activity, Event Related Potential (ERP), Recurrences, transition state)
Eye Tracking Data (5)	retina movement, pupil size

2.3.2 Regression Problems for Detecting NDDs

While logistic regression is commonly used for binary target variable datasets, it may not be accurate when the sample size is small. The connection between the predictors and a categorical response variable is modeled using logistic regression [58]. Regression problems are commonly addressed in the task of predicting or modeling the disorders as shown in Table 7.

Table 7. Regression: Types of machine-learning approaches and individual models used.

Study	Model	Best Model with Accuracy	Additional Performance Metrics
[6]	LDA, Ensemble (SVM-RBF, regression, Fuzzy sets, SVM+RBF	Cubic Regression method+SVM+RBF 98%	_
[7]	LR, SVM Polynomial, SVM RBF, NB, DT, RF, XGB, RF	LR 90.27%(ASD), RF 80.89% (Dyslexia)	PCS 92.30%, REC 90%, AUC 91.60%, F1-S 91.13%, CV 99.77% (ASD) PCS 83.56%, REC 77.21%, AUC 79.78%, F1-S 80.26%, CV 95.05% (Dyslexia)
[8]	DT, LR, RF	LR +CNN 81%	-
[9]	NB, LR, SVM	LR 95.87% (Adolescent), LR 99.82% (Adult), SVM 97.82% (Toddler), SVM 99.61% (Child)	KP 91.74%, F1-S 95.90%, AUROC 99.00% (Adolescent) KP 99.59%, F1-S 99.90%, AUROC 99.80% (Adult) KP 94.87%, F1-S 97.80%, AUROC 99.70% (Toddler) KP 99.21%, F1-S 99.60%, AUROC 99.60% (Child)
[10]	LR, MLP, CNN	∗MLP+LR 81%	PCS 77%, REC 78%, F1-S 76%
[11]	LR, SVM, polynomial regression, RF, MLP	Polynomial Regression 92.6%	PCS 91%, REC 89%, F1-S 92%
[12]	SVM, NB, DT, VGG16, DenseNet, AlexNet	LR 97.15%, RF 82%	-
[13]	LR, kNN, SVM, NB, DT, RF	LR 98.11%, SVM 98.11%, kNN 96.22%, NB 96.22%	F1-S, SP, SE
[14]	LR, MLP and SVC	LR 82.26%, SVC + PSO 93.55%	F1-S, SP, SE
[16]	NB, LR	LR 94.23% (Adults), LR 99.85% (Adolescent), LR 97.94% (Child)	SE 99.90%, SP 99.70% (Adults), SE 92.20%, SP 92.68% (Adolescent), SE 98%, SP 97.35% (Child)
[15]	RIPPER RIDOR Bagging AdaBoost CART RML C4.5 PRISM	RML LR-based 94.0% (Adult), 88% (Adolescent), 92% (Child)	SE 94.00%, SP 97.00% (Adults), SE 87.00%, SP 80.00% (Adolescent), SE 91.00%, SP 91.00% (Child)
[23]	PSO -RBF, PSO-BPNN	∗PSO+RBF 97%	SE 90%, SP 89%

Models: Cubic Regression Model, Logistic Regression (LR), Polynomial Regression, Naïve-Bayes (NB), Random Forest (RF), K Nearest Neighbour (kNN), Convolution Neural Networks (CNN), Multi-Layer Perceptron (MLP), Scala Vector Classifier (SVC), RIPPER RIDOR Bagging AdaBoost CART, Rules Machine Learning (RML), C4.5 PRISM, Back Propaganda Neural Network (BPNN), Particle Swam Optimizer (PSO), Radial Basis Function (RBF). Additional Performance Matrix: Precision (PREC), Recall (REC), F1-Score(F1-S), Confusion Matrix(CM), Specificity (SP), Sensitivity (SE), Area under Curve (AUC), Cohen Kappa (KP), Matthew Correlation Coefficient (MCC), False Discovery Rate (FDR), Cross Validation (CV), Negative Predictive Value (NPV), Positive Predictive Value (PPV), False Positive Value (FPV), Area under Receiver Operating Characteristic (AUROC)
Note: ∗Belongs to the Neural Network family.

For instance, Shilaskar et al. showed that logistic regression is the most accurate type of regression for autism, while the random forest is the most accurate type for dyslexia. However, given that the source data is highly biased and several performance indicators tend to zero, the results are not all that encouraging and their work needs to apply the Synthetic Minority Oversampling Technique (SMOTE) to handle imbalance data [7]. Besides, Thabtah et al. illuminated recent research that utilizes machine learning for ASD classification and investigated the utility of machine learning with Decision Tree and Random Forest for ASD prediction. They developed an ASDTest application to detect symptoms of ASD based on AQ-10 scoring data (cognitive, behavioral and social skill test) [15]-[16]. The results obtained using the Machine Learning with DT algorithm in WEKA were compared to the results obtained by other statistical models, such as Logistic Regression, showed superiority in detecting autistic traits over probabilistic classifiers derived by Naïve Bayes; however, the performance

decreased when data collected through the application gets bigger [15]. Accordingly, the work on using regression models to detect NDDs seems to decrease in its suitability due to its inefficiency in handling big datasets.

2.3.2.1 Classification Problems for Detecting NDDs

Classification is a common task in supervised learning, utilized when predicting categorical outcomes and determining whether a given example belongs to a specific category or not. This is distinct from regression, which is employed for predicting continuous values. However, although it can help categorize results based on certain tasks, it may not be able to handle all complex tasks within a timeframe. Most classification problems address the task of detecting NDDs as shown in Table 6. For instance, Alice Jacob et al. focused on ADHD detection from Time-Frequency images (TFIs) to identify the frequency band with higher ADHD-related data resulting in the TFI-based CNN model and GLCM-based KNN classifier supported higher ADHD-related information at the theta band compared with the upper- beta and lower- gamma bands [24]. The results showed the best performance on specific features in a short timeframe, but not with bigger TFI. Wang et al. (2023) proposed a model of ICA-CNN and achieved an accuracy of 67%; however, the training dataset comprises only 168 observations, which is insufficient for thoroughly training the parameters [25]. On the other hand, Omar et al. (2023) in their work on LSTM-CNN model in detecting epilepsy which is common in ADHD's children achieved a high accuracy of 97%. They also emphasized the significance of temporal dependencies in EEG signals, which reflect the connectivity and evolving state of the subject's cognition [26]. Due to the multi-dimensional and different datasets used in previous studies, there is an inconsistent affirmation stating which of the five bands has a significant effect on discriminating ADHD. In the U-Profile, resting-state power graph highlighted theta and beta best to detect ASD and ADHD [87], supported by Alim (2023) who indicated that signals less than 30Hz or the first four sub-bands are significant [80]. On the other hand, Parashar (2021) pointed out that all bands from each regional cortex are significant. Furthermore, different feature selections cause mixed statements about which nodes affect ADHD the most. Holker (2022) mentioned only six nodes; namely, FP2 (right pre-frontal), O2 (right occipital), F7 (left frontal), F8 (right frontal), T7 (left temporal) and P8 (right post-temporal) are the most important nodes [74]. Chen (2018) agreed on that all nodes are equally important [49]. Previously, a classical ML-based classifier was used to identify ADHD by extracting the features manually. Although the contributions have already been proved, they cannot achieve multiple-class classification with automated feature extraction. Meanwhile, the identifiable EEG segments of ADHD are too long to limit the real-time ADHD detection. Furthermore, methods of extraction that involved complex time-series features have not been extensively explored for ADHD [32]. The Deep Neural Network Framework has more layers (more depth) and each layer adds complexity to the architecture while enabling the framework to process the inputs concisely for outputting the ideal solution. LSTM can be applied when there is a long series with a sequence prediction that's required and some long- term dependency of data to go parallel with it. The CNN-LSTM framework proposed by Wang et al. (2022) incorporated features extracted by the CNN across various frequency bands and intricate ERP waveforms. However, despite this comprehensive feature set, the framework struggled to identify the ultimate key activities due to spatial feature-extraction problems [34].

2.3.3 Type of Datasets and Parameters Used

This sub-section summarizes and discusses the findings of RQ2: What type of NDD datasets in previous works have been used to build the models? and RQ3: What type of parameters or variables have been used? Table 9 shows the type of neurodevelopment disorders, dataset sources and related studies working on the prediction and detection of NDD.

There are different diagnosing way being performed to collect the datasets. Existing diagnostic tools to detect NDD include Clinical Observation, Statistical Evaluation, ML Classification, IOT/Robotics …etc. For instance, for ASD, most studies have used the Demographic, Medical, Facial Image Data, BMRI, Eye Tracking data, BMRI/FMRI and EEG data to perform the predictions and detections based on their representation stated in DSM-V. For structured datasets, the most frequently used databases include Observation and Behaviorial, BMRI/FMRI and EEG-based data. This is due to their availability on published well-known websites, like IEEE and Kaggle, for advanced research [87].

Incorporating multiple sources of data can be useful if there is a lack of data availability to predict and detect NDD Disorder. For instance, the dynamics of certain disorders, (e.g., cerebral palsy, epilepsy, GDD) could be associated with other information (e.g. neuron-defect density, population density and mobility) and this information should be incorporated in the process of modeling the classification of NDDs to reduce the residual errors of the models [83].

2.3.4 Type of Problems Addressed by Machine-learning Models

This sub-section summarizes and discusses the findings of RQ4: What types of problems are addressed using these models? and RQ5: Which models achieved the highest performance? Table 8 tabulates and summarizes the regression problems and all the individual machine-learning models applied to achieve the objectives of each study. On the other hand, Table 9 tabulates and summarizes the classification problems and all the relevant individual machine-learning models applied to solve these classification problems. The best models and their performances for each study are also tabulated in these tables. The details of the findings are discussed in the next sub-section 2.3. Based on the results shown in Table 8 and Table 9, for time-series data, VAR and LSTM are the most common machine-learning algorithms used to perform detection and prediction [17], [18], [31], [32], [83], [85], [87]. On the other hand, the family of ANN, *kk*NN, MLP and CNN algorithms are widely used in solving classification tasks [24], [30], [33]-[34], [36], [39], [47], [84].

Table 8. Classification: Types of machine-learning approaches and individual models used.

Study	Model	Best Model with Accuracy	Additional Performance Metrics
[24]	CNN, GLCMbased, KNN Time-Frequency image (TFIs)	∗Deep-CNN 99.75%	PREC 96.33%, REC 96.74%, F1-S 96.54%, CM FN 3.26% FP 3.75%
[25]	LR, SVM, RF, CNN	∗ICA+CNN 67%	SP 89%, SE 42%, PREC 77%, AUC 0.65
[26]	EEGNet, DeepConvNet and ShallowConvNet	∗DeepConvNet (LSTM-CNN) 96%	PREC 96%, REC 96%, F1-S 96%, KP 95.20%, MCC 95.23%, Robustness Difference 39%
[27]	SVM, kNN, RF, DT, CNN	SVM 88%	-
[28]	GPC, RF, kNN, MLP, DT, LR	GP-based 97.53%	REC98.46%, PREC 96.92%, AUC 0.99
[29]	SVM, LR, NB, CNN, CNN+LSTM	∗CNN+LSTM 98.03%	SP 98.97%, SE 99.25%, F1-S 99.13%, FDR 71%
[30]	SVM+RBF, expEEGNetwork-LSTM	∗expEEGNetwork-LSTM 99.06% and 98.68%	F1-S 99.14% and 99.24%, MCC 98%
[31]	EEGNET, ConvNets, LSTM	∗LSTM 90.50%	-
[32]	Graph FMRI, FCNet, fusion FMRI, Deep FMRI, SVM RFE	SVM RFE 75%	-
[33]	NLSVM, LR, RF, GNB, kNN, CNN	∗CNN+LR 95.83%	PREC 100%, REC 92%, F1-S 96%, AUC 0.96
[34]	LSTM, LessCNN+LSTM, DeepCNN+LSTM,CNN+LSTM	∗CNN 84.44%	SE 85.39%, SP 80.57%
[35]	ANN, SVM, kNN, MPL, LR, RF, GPC	LASSO+SVM 94.2%	SE 93.3%, SP 90.2%, AUC 0.96
[36]	CNN, 4D CNN, 6D CNN	∗4D CNN 98.56%, *6D CNN 98.85%	PREC 98.69%, REC 98.81%, F1-S 98.75% (4D CNN) PREC 98.75%, REC 99.25%, F1-S 99% (6D CNN)
[37]	VGGFace, ResNet50, VGG19, MobileNet (Xception)	∗MobileNet(Xception) 91%	SP 94%, SE 88%, CM FN 26%, FP 14%
[38]	VGGFace, MobileNet, resNet50, VGG19	∗MobileNet 97.60%	PREC 97.50%, SE 97%, SP 97%, AUC 0.97
[39]	RF, LR, DNN, CNN resting state	∗CNN 97%	CV 93-96%
[40]	CNN, VGG16, VGG19, ResNet50, ResNet101, ResNet152, AutoML	∗AutoML 96%	CV 94%
[41]	DSVM, DT, BDT, DNN	∗DNN 93.3% (AUC)	AUC 0.97, SE 93.28%, SP 91.38%, CV NPV 94.46%, PPV 90.06%
[42]	BQC, FF-NN, IF-SVM, kNN, LDA, SCNN, MBCNN, SVM-MLP, IPSO-NN, RF, SVM-RBF	∗1DCNN 99.70%-99%	PREC 98-99%, REC 98-99%, F1-S 98-99%
[43]	LR, kNN, SVM, NB, ALexNet, GoogleNet, SqueezeNet	∗CNN +LASSO (SqueezeNet) 88.33%	PREC 83%, AUC 0.83, FPV 16%
[44]	NB, kNN, LR	kNN 86%	PREC 100%, REC 78%
[45]	kNN, SVM, MLPNN, LEDPatNet19	∗LEDPatNet19 99.29% (Arousal 94.58%, Dominance 92.86% and Valance 94.44%)	PREC 99.29%, REC 99.30, F1-S 99.29 (Arousal FC6) PREC 94.43%, REC 94.63%, F1-S 94.53% (Valance F7)
[46]	SVM, kNN, J48, Bagging, Stacking, AdaBoost, NB	kNN 99.1%	CV 98.6%-99.2%
[47]	MLP, RF, CNN	∗CNN 92.31%	AUC 0.96 F1-S 91.54%, PREC 89.72%, REC 93.45%
[48]	SVMLinear, SVM+RBF, SVM+Grid,RF, RF+Grid	SVM+Grid 97.42%	PREC 96%, REC 91.4%, F1-S 93.4%
[49]	SVM, LR, NB, RF, DT, kNN	DT and NB 79.71%	AUROC 0.83, SP 96.4%, PPV 20.5%, SE 40%
[50]	rbio1.1 +kNN	rbio1.1 +kNN 99.17%	CV
[51]	CNN, VGG16	∗VGG16 68.54%	CV
[52]	LR, SVM, NB, kNN, ANN, CNN	∗CNN 96.88%	SP 100%, SE 93.35%
[53]	SVM, kNN, RF, CNN	*CNN 70.20%	SP 61%, SE 77% CV
[54]	Stacked autoencoders, Stacked autoencoders+MLP	∗MLP 85.06%	SE 81%, SP 89%
[81]	RF, SVM, DNN, CapsNet	∗CapsNets 71%	SE 73%, SP 66%
[55]	MLPNN, DeepCNN	∗Deep CNN 98.48%	PREC 97.48%, REC 97.47%, F1-S 97.47%, CV 99.06%
[56]	SVM, LDA, DT, RF, kNN+RKF	kNN+RKF 88.37%	SP 91.3%, SE 85%, AUC 0.88
[57]	SVM, SVM+RBF	SVM+RBF 91.3%	CV
[58]	kNN, Efficient Net, LR	ANN 97%	CV
[59]	LR, SVM, SVM+RBF	SVM RBF 98.62%	PREC 89%, REC 89%, F1-S 89%, CV 59.78%
[82]	SVM, RF, LR, 2CC3D	∗2CC3D 89%	F-Score 89%
[60]	MLP + DISR, MLP + mRMR	∗MLP+DISR 93.65%, ∗MLP+ mRMR 92.28%	Variance 0.7%
[61]	SVM	SVM 59-66.3%	SP 68%-87.7% SE 22.9%-55.6%

Models: Gray level co-occurrence matrix (GLCM)-based, Long Short Term Memory (LSTM), Gaussian Processes (GP), Naïve-Bayes (NB), Locations of Sophie Germain's Primes on Ulam's Spiral-Based Features (LSGP-USFNet), Expert EEG Network (expEEGNetwork), Least Absolute Shrinkage and Selection Operator with Support Vector Machine (LASSO +SVM), One Dimension Convolutional Neural Network (1D+CNN), extreme inception (Xception), Artificial Neural Network (ANN), LED Pattern Feature Extraction (LEDPatNet19), Back Propagation Neural Network (BPNN), Decision Tree (DT), Linear Regression (LR), *kk*-Nearest Neighbour (*kk*-NN), Support Vector Machine RBF kernel (SVM+RBF), Support Vector Machine (SVM), GoogleNet, AlexNet, Residual Neural Network (RNN), 2 Channel Convolutional 3 Deep Neural Network (2CC3D), Double Input Symmetrical Relevance (DISR), minimum Redundancy Maximum Relevance (mRMR), BQC: Bayesian quadratic classifier, FF-NN: Feed forward neural network, IF-SVM: Immune feature weighted SVM, QDA: Quadratic discriminant analysis, KNN: K nearest neighbor, SVM-RBF: SVM-radial basis function, SVM-RFE: SVM-Recursive Feature Elimination, Deep Belief Network (DBN), Decision Tree (BDT), Deep Support Vector Machine (DSVM). Additional Performance Matrix: Precision (PREC), Recall (REC), F1-Score(F1-S), Confusion Matrix(CM), Specificity (SP), Sensitivity (SE), Area under Curve (AUC), Cohen Kappa (KP), Matthew Correlation Coefficient (MCC), False Discovery Rate (FDR), Cross Validation (CV),

Negative Predictive Value (NPV), Positive Predictive Value (PPV), False Positive Value (FPV), Area under Receiver Operating Characteristic (AUROC).
Note: ∗Belongs to the Neural Network family.

Table 9. Disorders, database sources and studies.

NDD Disorder	Database Sources or Parameters
ASD	Demographic Data [12], [13], [15], [16], [19], [30], [52], [56], [67], [84] Medical Observation and Behavioral Data [9], [13], [15], [16], [40], [46], [49], [52], [69], [75] Eye Tracking Data [8], [10], [38], [41], [76] Facial Images Data [12], [37], [40], [47], [51], [58], [70], [72], [85] BMRI/FMRI Data [34], [39], [53], [54], [81], [82] EEG Data [71]
ADHD	Demographic Data [25], [61], [62], [78], [79] Medical Observation & Behavioral Data [25], [61], [62], [28], [35] Eye Tracking Data [11] BMRI/ FMRI Data [21], [22], [25], [32], [61], [62], [68], [78], [79], [86] EEG Data [6], [17], [18], [20], [23], [24], [25], [27], [28], [30], [31], [33], [35], [36], [42], [43], [44], [45], [50], [55], [57], [59], [60], [63], [64], [66], [73], [74], [80], [83], [84]
Dyslexia	Medical Observation & Behavioral Data [7], [48], [65] Eye Tracking Data [77] BMRI/FMRI Data [65] EEG Data[14], [27]
Others	EEG Data [26], [29], [87]

(Please, refer to Table 6 and Figure 3)

2.3.4.1 Approaches to Solve Regression Problems

The approaches to solve regression problems in detecting and predicting the occurrence of NDDs can be divided into statistical and machine learning approaches. Based on the information tabulated in Table 9, for the statistical approaches, several models have been used to perform the detection and prediction of NDDs, including the Cubic Regression Model [6], LR [7], [8], [11]-[12], [14], [16], [45], [60], [64] and ANN [38]. In multivariate and time-series modeling, Cubic Regression combined with SVM-RBF by Delisle et al. (2023) outperformed the statistical approach MX-VAR model by Redondo. Based on the review, deep-learning algorithms have outperformed the statistical approaches in detecting and predicting the disorders with multi-vatraite approaches, such as, Locations of Sophie Germain's Primes on Ulam's Spiral-based (LSGP-USFNet) [17], mixed-effect functional-coefficient autoregressive (MX-FAR) [18], Single Photon Emission Computed Tomography (SPECT) [62], Variational Mode Decomposition and Hilbert Transform-based (VHERS) [63], Multi-layer Perceptron (MLP), Phase-transfer Entropy (PTE) [20], Deep Variational Autoencoder (DVAE), Attention Attribute-enhanced Network (AAEN), Metaheuristic Spatial Transformation (MST) , Graph Signal Processing (GSP), Graph Learning (GL), Meta-cognitive Neuro-fuzzy Inference System (McFIS) (International Conference on Cognitive Computing and Information Processings 1. 2015 Noida et al., n.d.), Local Binary Encoding Method (LBEM), Linear Discriminate Analysis (LDA) [51], Kernel Principal-component Analysis (KPCA) [68], [79].

2.3.4.2 Approaches to Solve Classification Problems

Based on the information tabulated in Table 8, neural network methods have been found to be very effective in detecting NDDs. This review reports that the neural network-based methods have achieved 27 best results out of 81 studies [24]-[26], [29]-[34], [36]-[43], [45], [47], [51]-[55], [58], [60], [64]. These classification approaches use different methods of extraction and selection depending on the type of datasets represented for the purposes of their studies. Few researchers have used T-test and LASSO [28], [73], while few others used ICA and PCA to select the most discriminate features to optimize the multi-dimensional features within their datasets before being fed into their proposed models [20], [51]. Some authors applied the grid method to improve accuracy performance, such as Pralhad et al. (2021) who compared SVM and RF models using the grid method in dyslexia detection through Video on Observation and Behavioral datasets, resulting SVM using grid achieved the highest accuracy of 97.42% [48].

Various studies have investigated autism classification using diverse methodologies and datasets. For instance, Alsaade et al. (2022) evaluated deep-learning models' performance in detecting ASD *via* facial features, highlighting Xception's effectiveness [37]. Elshoky (2022) employed deep learning (VGG16), achieving a remarkable accuracy of approximately 96% compared to other deep-learning models, such as VGG19 and ResNet [40]. Kanhirakadavath and Chandran (2022) utilized eye-tracking datasets along with deep-learning models, while Kanhiraka, Rashid and Lin (2022) employed machine-learning techniques on eye-tracking data for early autism screening in children [41], [68]. Studies like Shilaskar et al. (2023) and Delisle-Rodriguez et al. (2023) utilized observation and

behavioral (AQ-10) datasets with supervised-learning models, noting SVM's superior performance [6]-[7]. In the area of medical research datasets, such as ABIDE and fMRI, were commonly used alongside machine-learning and deep-learning models, like MLP, NB, RF, CNN, ResNet and GoogleNet. Researchers like Attlah et al. and D.Wang et al. (2023) observed improved accuracies with their trained models compared to pre-trained ones [25]. Rabbi et al. (2021) compared various models, finding CNN to be highly accurate in detecting autism from facial images [47]. Ahmed et al. (2022) developed a web application using deep learning, achieving 95% accuracy with models like MobileNet [38]. Ahmed et al. (2022) adopted a deep transfer-learning approach, with MobileNet exhibiting the highest accuracy of 97% in detecting autism from children's facial images. Deep-learning algorithms offer significant benefits over statistical methods when it comes to uncovering inherent patterns for prognosis or diagnosis in neuropsychiatry [29]. In recent decades, research in neuropsychiatric diagnosis using EEG has primarily centered on addressing the "multi-dimensional problem" of localizing the complex brain-activity measurements. EEG-based models have seen extensive research in investigating dysfunctions across various neuropsychiatric disorders such as depression, Alzheimer's disease, epilepsy, phobias, conduct disorder, schizophrenia and NDD. Often, these methods are combined with artificial intelligence or machine-learning approaches, as shown in Table 10.

Table 10. Statistical methods used for classification and regression problems.

Study	Model	Best Model	Additional Performance Metrics
[17]	kNN, CNN, LSTM, SVM, NB, LSGP-USFNet	LSGP-USFNet +SVM 97.5%-98.9%	SE 90.57%-99.06%, PREC 91.45%-98.49%, F1-S 93.03%-98.77%
[18]	FAR, VAR	MX-VAR 95% (fPDC)	Mean fPDC 95%
[62]	SPECT, CAD, ML (Different Brain Region)	SPECT+CAD+ML 80% (frontal cortex)	F-Measure 79.95%
[63]	CNN, MLP, VHERS	ELM VHERS 99.95% (delta)	SE 100%, SP 99.89%, KP 99.9%, PREC 99.91%, F1-S 99.9%, MCC 99.9%
[64]	MLP	∗MLP 90.01% (Trend)	SE 90.55%, SP 89.84%
[20]	gECV+ANN+GA (PTE Brain Region)	*ANN+gECV 89.7%	PTE p<0.01 dPTE 0.5<dPTExy≤1
[19]	AAEN	∗AAEN 86.22%	SE 44.45%-98.18% SP 66.66%-97.14%
[22]	McFIS+GA+ELM	McFIS+GA+ELM (63 voxels taken from Top- 50 best binary solutions)	PREC 92%, REC 90%, F1-S 90%
[21]	VA-Relief	VA-Relief 98.04%	-
[78]	Functional connectivity, resting state	LDA 80.08%	SE 80.7%, SP 79.47%
[79]	KPCA-SVM	KPCA-SVM 81%	-

Models: Locations of Sophie Germain's Primes on Ulam's Spiral-Based (LSGP-USFNet). mixed-effects functional-coefficient autoregressive (MX-FAR), functional Partial Directed Coherence (fPDC), Single Photon Emission Computed Tomography (SPECT), Variational Mode Decomposition and Hilbert Transform-Based (VHERS), extreme learning machine(ELM), Multi-Layer Perceptron (MLP), Phase Transfer Entropy (PTE), Genetic Algorithm (GA), Global Effective Connectivity Vector (gECV), Deep Variational Autoencoder (DVAE), attention attribute-enhanced network (AAEN), Graph Signal Processing (GSP), Graph Learning (GL), Meta- Cognitive Neuro-Fuzzy Inference System (McFIS), Extreme Learning Model (ELM), linear discriminate analysis (LDA), kernel principal component analysis (KPCA). Additional Performance Metrics: Sensitivity (SE), Specificity(SP), Precision (PREC), Kohen's kappa (KP), F1-Score (F1-S), Mathews Correlation Coefficient (MCC), Functional Partial Directed Coherence (f PDC), Multivariate Analysis Of Variance (MANOVA).
Note: ∗Belongs to the Neural Network family.

This line of inquiry offers considerable potential for revealing neural correlates of NDD, enhancing diagnostic methods and progressing treatment strategies. This entails employing sophisticated statistical techniques, like low-resolution electromagnetic tomography (LORETA), Phase Transfer Entropy (PTE), Variational Mode Decomposition and Hilbert Transform-Based (VHERS), optimization methods, among others, to overcome the inherent spatial resolution limitations of EEG [20], [63]. Furthermore, there was a drastic increasing amount of research conducted with EEG-based datasets. Studies and methods of feature extraction and selection are shown in Table 12.

2.3.5 Assessment Measures and Methods

This sub-section summarizes and discusses the findings of RQ6: What evaluation metrics and methods are employed to measure the performance of the machine-learning models? (e.g. Accuracy, Precision, Recall, F-Measure, ROC, AUC, Kappa) of the proposed machine learning algorithms for prediction and detection models? In most regression problems, all the proposed methods or algorithms are measured by using Autoregressive (VAR), mean, standard deviation (STD), mean functional partial directed coherence(fPDC), Root Mean Square Error (RMSE), t-test, two-way ANOVA analysis, average shortest path (d) and betweenness centrality (Cbetweenness), Friedman test, Nemenyi test, 10-fold metrics (Recurrence, Determinism, Entropy, Laminarity, Trapping Time and Trend), permutation statistical test, VOXELS'COUNTS and high testing efficiency (fitness value), nested cross-validated accuracy and kappa score. On the other hand, Accuracy and ROC are mostly used for evaluating the performance of the classifiers proposed in those studies. In this paper, 27 out of 81 (33%) studies found that the individual models that belong to the neural-network family performed better when compared to other linear and non-linear methods. Tables 9 and 10 show that machine-learning models

achieved lower Mean Absolute Error (MAE) and Mean Squared Error (MSE) measurements compared to other statistical models (e.g. VAR and MX-VAR) [18]. However, for long-term trend, it can be observed from these tables that deep-learning approaches improve RMSE readings in non-linear models' classification which achieved above 98% of accuracy [36], [44], [55]. As we have noticed, based on summaries stated in previous sub-sections, machine-learning approaches performed better than statistical approaches. Deep-learning and ensemble algorithms consistently exhibit a trend of achieving higher accuracy measurements [24], [55], $FF1$ Score measurement [11]-[12], [28], [33], [53], [55], [59], [74], AUC [35], [41] and ROC measurement [67], [73], in comparison to other statistical and machine-learning models evaluated in this study.

2.3.6 Ensemble Method

This sub-section summarizes and discusses the findings of RQ7: What types of ensemble methods are used in machine-learning models?

Various ensemble approaches have been introduced for predicting NDDs. Table 11 provides an overview of these methods used for predicting and detecting disorder outcomes, along with summarizing the evaluation techniques and metrics employed in ensemble learning. Further exploration is warranted to assess the potential of ensemble or hybrid models based on deep-learning techniques utilizing multi-source data, as they have demonstrated enhancements in base-model performance. An ensemble method refers to a strategy that employs multiple independent models or weak learners, which may be similar or diverse, to generate an output. Ensemble methods are typically classified into boosted trees, bagged trees, subspace kNN and stacked approaches [73]. Bagging involves employing homogeneous weak learners arranged independently in parallel and aggregating their predictions to determine the final output.

Table 11. Ensemble methods used for classification problems.

Study	Dataset	Best Model with Accuracy	Additional Performance Metrics
[65]	MS-ROI XGBoost	MS-ROI XGBoost 99.87%	PREC 92.36%, REC 91.65%, SE 99.89%, SP 99.91%
[74]	ANN, RF, SVM	RF 81.82%	F1-S 81.79%, PREC 81.95%, REC 81.82%
[66]	SVM, RF, AdaBoost	Adaboost 82%	SE 75%, SP = 86%
[73]	Ensemble	Ensemble 98.33%	CF
[67]	Boosting, DT, NN, NB	RF+SMOTE 98%(ROC)	CF TPR 88% TNR 93%
[75]	SVM, RF, SMO	RF 87% (ROC)	TPR 88.5%
[68]	CDAE+AdaDT	CDAE+AdaDT 90% (AUC)	SE 76.92%, SP 73.08%, CF
[76]	DT, NB, kNN, SVM, Stacking	Ensemble(stacking) 89.82%	SE 89.21%, SP 90.31%, KP 0.33%
[69]	SVM, kNN, RF, NB, AdaBoost, SGD, CN2 rule inducer	SGD 99.6% (Adult), RF 97.2% (Adolescent) RF & SGD 99.7% (Toddler)	F1-S, PREC & REC (90%-100%)
[72]	SVM, RF, LR, kNN, SVM+PSO	SVM-PSO 95.6%, RF 90.45%	-
[70]	DT, CNN, AdaBoost	Adaboost 98.77% (Toddler), 97.20% (Child), 93.89% (Adolescent), 98.36% (Adult)	SE 99.39%, SP 99.39%, KP 97.10%, AUROC 99.98%, Logloss 3.01% (Toddler),SE 98.40%, SP 98.46%, KP 94.41%, AUROC 99.89%, Logloss 9.62% (Toddler) SE 97.50%, SP 98.33%, KP 89.37%, AUROC 98.61%, Logloss 15.80% (Toddler),SE 99.30%, SP 96.11%, KP 96.02%, AUROC 99.95%, Logloss 5.64% (Toddler)
[71]	RF, LR, Bagging, CNN	RF 97%	PREC 97%, REC 97%, F1-S 97%
[77]	kNN,	kNN 53.4%	MANOVA p-value<0.01

Models: MS-ROI XGBoost, AdaBoost, Ensemble, Random Forest (RF), Random Forest Based (RF-based), Stochastic Gradient Descent (SGD), Partical Swam Optimization (PSO), Gradient Boosting Machine (GBM), Sequential Minimal Optimization (SMO), Convolutional Denoising Autoencoder (CDAE), Adaptive Boosting Decision Trees (AdaDT). Additional Performance Metrics: Precision (PREC), Recall (REC), Sensitivity(SE), Specificity(SP), Confusion Matrix(CF), True Positive Rate(TPR), True Negative Rate(TNR).

For instance, in their study on classifying ASD *versus* control groups, M. Rakic and M. Cabezac combined data from functional and structural MRI and assessed it on a sizable multi-site dataset. Their quantitative analysis was conducted on 817 cases from the International Autism Brain Imaging Data Exchange I (ABIDE I) dataset, comprising 368 ASD patients and 449 control subjects. They achieved a classification accuracy of 85.06% with a standard deviation of 3.52% when employing an ensemble of classifiers. Combining information from both functional and structural sources resulted in significantly improved performance compared to using an individual pipeline [54]. Sangeetha et al. (2022) showed that ensemble methods, especially MS-ROI with XGBoost, are capable optimizing computational time in detecting dyslexia within smaller data sizes [65]. Hamedi et al. (2021) used the stacking method in detection for ASD with rs-MEG signals and achieved an accuracy of 89.82%,

showing that the left central (LC) of the brain can discriminate the ADHD group [76]. Thus, ensemble methods have proven to improve predictive performance using an individual model and multiple learning algorithms although they are time and space-consuming compared to other machine-learning models [67], [74], [76]. Efforts need to be directed towards harnessing the potential of ensemble methods in future-research endeavors, in order to bolster their applications for addressing various disorders.

Table 12. Feature-extraction and machine-learning models from related EEG-based studies.

	Features Extraction										Features Selection								ACC	Model
Paper	3FD	PSO	ICA	SPM(MSP)	LASSO	T-Test	RFE	3EDAs	VMD-HT	DISRmRMR	Bandsδ θ/β	ROI	Recurrence	Resting State	Changing state	ERP	PCA	Time Series		
[60]	✓									✓	✓								93.65%	MLP
[23]		✓									✓								90%	SVM (RBF)
[57]											✓		✓	✓					91.3%	SVM (RBF)
[55]											✓					✓			98.48%	CNN (DCNN)
[73]	✓				✓		✓				✓								98.33	Ensemble
[66]											✓	✓							84%	Ensemble (Adaboost)
[20]			✓	✓													✓		98%	ANN (dTF+ANN)
[35]					✓	✓					✓								94.2	SVM (LASSO)
[36]				✓							✓								98.85%	CNN (6D CNN)
[33]											✓								95.83%	CNN (CNN+LR)
[74]				✓							✓								81.82%	RF
[28]					✓	✓					✓								97.53%	SVM (GP-based)
[30]								✓			✓								96.16%	ANN (MEMD-GA-ANN)
[80]											✓						✓		94%	SVM (Gaussian)
[63]									✓		✓								99.81%	DNN+ELM (VHERS)
[59]	✓										✓								98%	SVM (RBF)
[55]				✓							✓								98.48%	CNN
[42]				✓							✓								98%	DCNN
[6]				✓							✓								81.37%	SVM (RBF)
[51]			✓								✓			✓					85%	CNN
[50]											✓							✓	99.17	SVM (RBF) kNN Bio and rbio
[73]	✓										✓								98.33%	Ensemble (subspace)
[45]				✓							✓								99.29%	LEDNet (LEDPatNet19)
[83]				✓							✓							✓	97.75%	LSTM
[24]				✓							✓								99.75%	CNN (TFI-based)
[34]				✓							✓			✓				✓	98.23%	CNN-LSTM
[17]				✓							✓							✓	97.46%	Gray-tones (LSGP-USFNet)
[26]			✓								✓							✓	96%	CNN (DeepConvNet)
[31]	✓										✓							✓	90.50%	LSTM

Notes: 3FD: Higuchi, Katz and Petrosian fractal dimensions Largest Lyapunov Exponent (LLE),PSO: Partical Swam Optimization, ICA: Independent Component Analysis,SPM: Statistical Parametric Mapping applied multiple sparse priors (MSP) algorithm, LASSO: Least absolute shrinkage and selection operator,T-Test: T score = (difference between the group)/(difference within the groups),RFE: Recursive feature elimination,3EDAS: three multivariate EDAs (MEMD, MEWT and MVMD), VMD-HT: variational mode decomposition (VMD) and Hilbert transform (HT),DISR: Double Input Symmetrical Relevance (DISR),mRMR: minimum Redundancy Maximum Relevance, ROI: Region of Interest, ERP: Event Related Potential, PCA: Performance Component Analysis

2.3.7 Deep Learning Method

This sub-section summarizes and discusses the findings of RQ7: What types of deep-learning approaches are used in NDD detection?

Within the emergence of machine learning, the most effective methods identified for predicting neurodevelopment disorders are predominantly associated with the neural-network family. The experimental results showed consistent performance improvements by the proposed deep-learning approaches over other representative linear and non-linear methods on multiple real-world datasets. These algorithms include the Long Short-Term Memory (LSTM) [30]-[32], [83], Convolutional Neural Network (CNN)[24], [27], [33], [34], [36], [39], [47], [55], Multi-layer Perceptron (MLP) [11], [14], [45], [60], [64], Neural Network [31], [34], [42]-[43], [47], [53], [55], Hybrid Neural Network (HNN) [30], [45], [81], [82] and combinations of statistic and deep-learning approaches. LSTM algorithms were shown to be superior in detecting ADHD, which supports long sequential data, like EEG [29], [31], [34]. A feature selection-based time-series modeling has been proposed for predicting future disorders [24], [26], [87]. The work proposed a multi-objective evolutionary algorithm to find the best neural-network algorithm (deep learning) for detection differences. Although the Convolutional Neural Network (CNN) is the best model when it comes to process image data, as it is capable to excel local features and is good in pattern recognition [47], it has limited effectiveness for

sequential data. For large datasets, training takes a long time to complete. In previous studies, Kaur et al. (2021), Moghaddari et al. (2020), Mafi & Radfar (2022), TaghiBeyglou et al. (2022) and Saini et al. (2022) have conducted their work on ADHD detection using CNN model [33], [36], [51], [55]. Moghaddari in his work to tackle the ERP fatigue problem using deep CNN achieved an accuracy of 98.48%. Mafi & Radfar (2022) used 4D and 6D connectivity tensors as a convolutional neural network input, achieving an accuracy of 98.85%. While Taghibeyglou et al. (2022) achieved an accuracy of 95.83% on their CNN+LR model, the framework suffers from a time-consuming training procedure, since the method focuses only on raw time series in both spatial and temporal domains. Furthermore, in the work of Saini et al. (2022) on the evaluation of their proposed architecture 1DCNN on three databases, the best accuracy was achieved on the database with few features compared to the database with more features [44]. Hence, an improved model is required to be able to overcome the limitation of CNN model in processing more features for ADHD detection.

Long Short-Term Memory (LSTM) is a deep recurrent neural-network architecture utilized for the classification of time-series data, a crucial aspect of time-series analysis focusing on comprehending and predicting sequential data points over time [68]. Within deep learning, LSTM models are applied to regression analysis, addressing issues of non-linearity and data interdependence to enhance traditional regression models. These networks are trained to classify sequence data, leveraging LSTM's capability to retain information from previous inputs over extended periods. This characteristic renders LSTM particularly effective for handling sequences with prolonged dependencies, where earlier time steps significantly influence subsequent ones. Sharma & Singh (2023) in their novel approach on expEEGNetwork-LSTM achieved an accuracy of about 98.02% [30]. In other works, Huang et al. (2022) with their objectives to solve time window issues in deep learning, they achieved an accuracy of 90.50% with their LSTM model. One drawback of LSTM models is their computational intensity, requiring more processing time compared to alternative methods [31]. While LSTM models can achieve high accuracy, there remains room for improvement with certain datasets. Notably, LSTM overcomes the limitations of traditional RNNs by employing separate memory cells capable of storing long-term information independently of current inputs or outputs [30], [31], [34], [83]. This property enables LSTM to learn and retain long-term dependencies while mitigating issues like the vanishing or exploding-gradient problem. Another way to optimize the LSTM model is to use hyper-parameter optimization, which is a process that involves searching for the best combination of values for the parameters that control the behavior and performance of the model, such as the number of layers, units, epochs, learning rate or activation function like sigmoid, hyperbolic tangent and rectifier.

A CNN-LSTM network on the other hand uses convolutional and LSTM layers to learn from the training data. Huang et al. (2022) and Zhang et al. (2023) showed that the proposed EEG-based LSTM networks can extract the varied temporal characteristics of high-resolution electrophysiological signals to differentiate between ADHD and NT children and bring new insights to facilitate the diagnosis of ADHD [26], [31]-[32] by leveraging LSTM’s ability to capture temporal dynamics and Convolutional Neural Network (CNN) capability to detect spatial patterns. The proposed method proved successful in enhancing EEG classification by outperforming existing models developed for similar EEG-based classification tasks. Wang et al. (2022) in their work with the LSTM-CNN model to process multiple frequency bands and complex ERP waveforms achieved an accuracy of 98.23% [25]. Somehow, this did not help the network find the final key activities. An improved deep-learning model that can extract more spatial feature information from multi-channel EEG signals could be employed to identify commonalities and sub-types [34]. Omar et al., 2022 in their work on detecting epilepsy applied Convolutional Neural Networks (CNNs) for extracting spatial features and Long Short-Term Memory (LSTM) for identifying temporal dependencies, achieving an accuracy of 96% focusing on scalability and efficiency. However, their result suggests that even models with fewer trainable parameters may still require many epochs or batch sizes to achieve optimal performance, highlighting the importance of careful model selection and hyper-parameter tuning [26].

Table 13 illustrates brief description of methods and techniques: their principles, advantages and limitations, in terms of each machine-learning model and technique used in this study.

Table 13. Description methods of each machine-learning model and technique used.

Methods/Techniques	Principles	Advantages	Limitations
Logistic Regression	• Linear model used for binary classification. • Predicts the probability of a binary outcome by applying the logistic (sigmoid) function to a linear combination of input features.	• Coefficients can be interpreted to understand the relationship between features and the probability of the outcome • Computationally efficient with a closed-form solution. • Provides probability estimates for classification	• Assumes a linear relationship between features and the log-odds of the outcome, which may not capture complex patterns • Requires proper feature scaling for optimal performance. • Best suited for binary outcomes, though variations exist for multiclass classification
Decision Trees	• Non-linear model that splits data into subsets based on feature values • Creates a tree-like structure where each internal node represents a feature (or attribute), each branch represents a decision rule and each leaf node represents an outcome. • Constructs the tree using criteria such as Gini impurity or information gain to make splits.	• Easy to visualize and interpret decision rules • Handles both numerical and categorical data without scaling • Can model complex relationships through hierarchical splits	• Prone to overfitting, especially with deep trees. • Small changes in data can lead to different tree structures • May create biased trees if some classes dominate
Random Forest	• Ensemble method using multiple decision trees • Aggregates predictions from multiple decision trees to improve accuracy and robustness • Builds a multitude of trees using bootstrapped samples and random feature subsets, then averages (regression) or votes (classification) to make the final prediction	• Typically, more accurate than a single decision tree due to averaging and reducing variance • Less prone to overfitting compared to individual decision trees • Can provide insights into the importance of different features	• Less interpretable compared to single decision trees • More computationally intensive, requiring more memory and processing power • Making predictions can be slower due to the need to aggregate results from multiple trees
Support Vector Machine (SVM)	• Supervised learning algorithm for classification and regression • Finds the hyperplane that best separates classes in a high-dimensional space. For regression, it finds the hyperplane that best fits the data within a specified margin of tolerance. Allows the algorithm to operate in higher-dimensional spaces using kernel functions (e.g., polynomial, RBF)	• Works well in high-dimensional spaces and with a clear margin of separation • Especially effective in cases with a clear margin of separation • Can use different kernels for non-linear classification	• Training can be time-consuming, especially with large datasets • Performance heavily depends on the choice of kernel and hyperparameters • May not perform well with very large datasets compared to other methods
Multi-Layer Perceptron (MLP)	• Type of artificial neural network with multiple layers of neurons. • Consists of an input layer, one or more hidden layers and an output layer. Uses non-linear activation functions (e.g., ReLU, sigmoid) to model complex relationships • Trained using backpropagation and gradient descent to minimize a loss function	• Capable of modelling complex non-linear relationships • Can be used for various types of tasks, including classification, regression and more • Automatically learns features from raw data	• Can be slow to train, especially with large networks and datasets • Prone to overfitting, especially with a large number of parameters • Performance can be sensitive to hyperparameters and network architecture
Convolutional Neural Networks (CNN)	• Specialized neural network for processing grid-like data (e.g., images). • Uses convolutional layers to automatically learn spatial hierarchies of features (edges, textures, etc.) and pooling layers to reduce dimensionality • Comprises convolutional layers, activation functions, pooling layers and fully connected layers	• Automatically learns and extracts features from images or spatial data • Reduces the number of parameters and computational load through convolutional filters • Performs exceptionally well in tasks like image classification and object detection	• Requires significant computational power and memory • Can be slow to train, especially with large networks • Typically needs large amounts of labelled data for effective training
Recurrent Neural Network (RNN)	• Neural network designed for processing sequential data • Uses loops to maintain a state across sequences, allowing it to handle temporal dependencies • Contains recurrent connections that process sequences one element at a time and update the internal state	• Suitable for tasks involving sequential data, such as time series or text • Can maintain context over sequences to some extent	• Struggles with long-term dependencies due to vanishing gradient issues • Difficult to train on long sequences; often requires more sophisticated architectures like LSTMs or GRUs • Can be computationally demanding, especially for long sequences
Long Short-Term Memory (LSTM)	• A type of Recurrent Neural Network (RNN) designed to handle long-term dependencies and sequential data • Uses gates (input, forget and output) to control the flow of information and manage long-term dependencies in sequences • Comprises LSTM cells that maintain a memory cell to remember information over long periods	• Effectively captures long-term dependencies in sequential data • Mitigates the vanishing gradient problem common in traditional RNNs • Used in various applications like time series forecasting, language modelling and sequence prediction	• Training can be resource-intensive due to the complexity of the model • More complex to understand and tune compared to simpler mod

3. CONCLUSIONS

Based on previous studies on Neurodevelopment Disorder, a summarization included in this review shows strengths, limitations and future directions for research on this domain.

This literature review endeavors to identify and examine various methodologies, datasets, parameters, individual models, ensemble models, performance metrics and approaches utilized in prior research on employing machine-learning techniques to mitigate the escalation of Neurodevelopment Disorder. Six

online digital libraries were utilized to retrieve pertinent peer-reviewed articles, resulting in the selection of 81 studies published between 2013 and 2023. The primary objective of this systematic literature review (SLR) was to assess and curate all pertinent research studies concerning the detection and prediction of Neurodevelopment Disorder using machine learning, guided by the mentioned seven questions. The contributions of this paper can be summarized as follows:

- Recognition of the improvement in predicting NDDs by leveraging diverse data sources.
- Acknowledgement of the superior efficacy of neural-network algorithms over alternative linear and non-linear machine-learning approaches.
- Validation of the efficacy of deep learning and hybrid methodologies, showcasing their superior performance and appropriateness in predicting and detecting ND Disorders.

3.1 Significance, Limitations and Future Directions

From this review, we have identified limitations that affect previous works on detection and prediction of Neurodevelopment Disorder using machine learning. Autism Spectrum Disorder (ASD) detection is well studied and achieved maximum performance and highlighted the strengths of signal fusion utilizing Signal-processing and Decision-making techniques. The review cautions that the focus on detecting ASD may overshadow research into other diseases, despite the promising results achieved in ASD studies. While signal-fusion techniques have been extensively explored, other Neurodevelopment Disorders (NDDs) have not received as much research attention. Sustainable ML models are suggested for future work to provide models with feature fusion able to merge different extracted features from various sources and compressed to a single layer before being fed into ML models. Therefore, fusing only important features and suppressing the others will reduce time complexity, thus improving the model's performance. The limited research on signal fusion for NDDs is due in part to challenges in information technology and computer science. A new approach is needed to manage and integrate signals from multiple sensors using artificial intelligence to create a single, optimized feature for meaningful analysis. Although current cloud technologies, such as Google Colab and Kaggle, enable researchers to upload and test datasets, collaboration is often hindered by issues related to credentials and copyrights. Additionally, the limited number of investigations conducted on NDD prediction based on multi-source data underscores the potential for obtaining a more comprehensive understanding of the disorder by integrating such data sources. Analyzing the complex relationships among multi-source data can yield more robust modeling outcomes. To address this issue, researchers need to collaborate openly and be properly credited for their contributions. This would allow signal fusion for NDDs to receive the attention that it deserves and facilitate more effective investigations.

These studies also explored multiple validations that prove the accuracy of each prediction. However, due to the limitations of public datasets, average testing can be performed to varying performances of ML and DL techniques. The analysis highlights that the limitations of publicly available datasets often undermine the effectiveness of machine-learning and deep-learning techniques. This restriction hampers thorough testing and leads to inconsistent performance results across different research studies. Some professionals may face challenges in sharing datasets online due to limited access to technology or varying levels of expertise. To bridge this gap and enhance research outcomes, greater collaboration between medical professionals and data scientists is essential.

Future directions should be ready for the paradigm shift through the emerging technologies which require models in handling big datasets that allow fusion of features to be processed simultaneously. This study underscores the need for future research to embrace new technologies that can manage vast amounts of data. Current algorithms may struggle to handle the simultaneous integration of multiple features, which is crucial for enhancing detection and prediction accuracy. As the volume of data continues to grow, it is essential to develop technologies that can process large datasets while effectively merging various data characteristics. Advanced cloud solutions capable of intelligently integrating these features are needed. Approaches such as genetic algorithms, sentiment analysis and Large Language Models (LLMs) have made strides in this area, but further innovation is required to address the challenges of data integration. On the other hand, studies on ADHD detection have increased the research exposure, especially research related to neurons which acquired deeper feature explorations and sustainable approaches.

Despite the increasing number of studies on Attention Deficit Hyperactivity Disorder (ADHD), more research into its characteristics and the development of sustainable strategies is still needed. This suggests that, while some progress has been made, there remains a significant gap in understanding and managing ADHD through machine learning. To address this gap, clear guidelines on the application of machine learning and deep learning for ADHD are essential to enhance researchers' knowledge and comprehension. Researchers can access handbooks and other resources online, including detailed explanations on platforms like YouTube. Forums on GitHub, Kaggle and Ubuntu also provide opportunities for discussion. Additionally, platforms such as Medium.com and blogs featuring data science can help bridge the gap in understanding and treating ADHD using machine learning.

There is also a need for further exploration of the capacity of deep-learning models or hybrid models in leveraging multi-source data, given their demonstrated ability to enhance the performance of base models. Although several studies have applied cross-fold validations and proven models to be powerful, models are tested on single datasets and are non-data driven. This research also highlights the fact that existing literature frequently lacks in-depth descriptions of specific machine-learning algorithms, datasets and performance indicators. When comparing the predictive and identification efficacy of various approaches, this discrepancy creates a challenge. Some approaches involved data augmentation or ablation approach to train the models. A new performance matrix is required to complement the current evaluation metrics, like accuracy, RMSE, Confusion Matrix and k-fold validation. The new performance metrics should be able to encompass the differences between models which applied different machine-learning algorithms, signal fusions and overfitting/underfitting regardless small/large capacity of data.

Furthermore, to improve the uncertainty and explainability of proposed models, it is essential to explore publicly available datasets with diverse modalities. Enhancing model interpretability is crucial for industry professionals, as understanding how models generate predictions is vital for trust and effective use of these technologies. Approaches such as Explainable AI, Interpretable AI, Responsible AI and Generative AI provide valuable tools and frameworks to facilitate the understanding and interpretation of machine-learning predictions. Integrated with various Google products and services, these approaches help in troubleshooting and refining model performance while also aiding in comprehending how models function. Applying these methods to each testing model can address the challenge of model interpretability.

ACKNOWLEDGEMENTS

The first author would like to thank and acknowledge the financial support from the Public Service Department (PSD) for their generosity in giving the sponsorship of “Hadiah Latihan Persekutuan”. The authors take responsibility for the integrity of the data and the accuracy of the data analysis. The authors also thank the anonymous reviewers and the editor for their helpful comments.

REFERENCES

[1] Diagnostic and Statistical Manual of Mental Disorders (DMS-5-TR), 5th edn. Arlington, American Psychiatric Association, 2013.

[2] D. J. Morris-Rosendahl and M. A. Crocq, "Neurodevelopmental Disorders-the History and Future of a Diagnostic Concept," Dialogues in Clinical Neuroscience, vol. 22, no. 1, pp. 65–72, Mar. 2020.

[3] M. J. Maenner et al., "Prevalence and Characteristics of Autism Spectrum Disorder among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2020," MMWR, Surveillance Summaries, vol. 72, no. 2, pp. 1–14, Mar. 2023.

[4] R. Alfred and J. H. Obit, "The Roles of Machine Learning Methods in Limiting the Spread of Deadly Diseases: A Systematic Review," Heliyon, vol. 7, no. 6, Elsevier Ltd., Jun. 01, 2021.

[5] B. Kitchenham et al., "Systematic Literature Reviews in Software Engineering: A Systematic Literature Review," Information and Software Technology, vol. 51, no. 1. pp. 7–15, Jan. 2009.

[6] D. Delisle-Rodriguez et al., "Multi-channel EEG-based BCI Using Regression and Classification Methods for Attention Training by Serious Game," Biomedical Signal Processing and Control, vol. 85, DOI: 10.1016/j.bspc.2023.104937, Aug. 2023.

[7] S. Shilaskar et al., "Prediction of Autism and Dyslexia Using Machine Learning and Clinical Data Balancing," Proc. of the 2023 IEEE Int. Conf. on Advances in Intelligent Computing and Applications (AICAPS 2023), DOI: 10.1109/AICAPS57044.2023.10074161, 2023.

[8] K. F. Kollias et al., "Autism Detection in High- functioning Adults with the Application of Eye-tracking Technology and Machine Learning," Proc. of the 2022 11th IEEE Int. Conf. on Modern Circuits and Systems Technologies (MOCAST 2022), DOI: 10.1109/MOCAST54814.2022.9837653, 2022.

[9] M. Bala, M. H. Ali, M. S. Satu, K. F. Hasan and M. A. Moni, "Efficient Machine Learning Models for Early Stage Detection of Autism Spectrum Disorder," Algorithms, vol. 15, no. 5, DOI: 10.3390/a15050166, May 2022.

[10] R. M. Kannan and R. Sasikala, "Predicting Autism in Children at an Early Stage Using Eye Tracking," Proc. of the 2nd IEEE Int. Conf. on Vision towards Emerging Trends in Communication and Networking Technologies (ViTECoN 2023), DOI: 10.1109/ViTECoN58111.2023.10157663, 2023.

[11] A. R. Khan et al., "EXECUTE: Exploring Eye Tracking to Support E-learning," Proc. of the IEEE Global Engineering Education Conf. (EDUCON), pp. 670–676, Tunis, Tunisia, 2022.

[12] Ramanjot, D. Singh, M. Rakhra and S. Aggarwal, "Autism Spectrum Disorder Detection Using the Deep Learning Approaches," Proc. of IEEE Int. Conf. on Technological Advancements in Computational Sci. (ICTACS 2022), pp. 761– 766, Tashkent, Uzbekistan, 2022.

[13] N. Zaman, J. Ferdus and A. Sattar, "Autism Spectrum Disorder Detection Using Machine Learning Approach," Proc. of the 2021 IEEE 12th Int. Conf. on Computing Communication and Networking Technologies (ICCCNT 2021), DOI: 10.1109/ICCCNT51525.2021.9579522, 2021.

[14] V. Chakraborty et al., "Intelligent Dyslexia Prediction Empowered with Optimization Techniques," Proc. of the 2021 6th IEEE Int. Conf. on Recent Trends on Electronics, Inf., Comm. and Tech. (RTEICT 2021), pp. 412–415, DOI: 10.1109/RTEICT52294.2021.9573823, 2021.

[15] F. Thabtah and D. Peebles, "A New Machine Learning Model Based on Induction of Rules for Autism Detection," Health Informatics J., vol. 26, no. 1, pp. 264–286, Mar. 2020.

[16] F. Thabtah, "An Accessible and Efficient Autism Screening Method for Behavioural Data and Predictive Analyses," Health Informatics J., vol. 25, no. 4, pp. 1739–1755, Dec. 2019.

[17] O. Atila et al., "LSGP-USFNet: Automated Attention Deficit Hyperactivity Disorder Detection Using Locations of Sophie Germain's Primes on Ulam's Spiral-based Features with Electroencephalogram Signals," Sensors, vol. 23, no. 16, DOI: 10.3390/s23167032, Aug. 2023.

[18] P. V. Redondo, R. Huser and H. Ombao, "Functional-coefficient Models for Multi-variate Time Series in Designed Experiments: With Applications to Brain Signals," arXiv: 2208.00292, [Online], Available: http://arxiv.org/abs/2208.00292, 2022.

[19] S. Subramaniam, "Enabling Innovative Technologies for Global Healthcare," 42nd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society Proc., Institute of Electrical and Electronics Engineers and Canadian Medical and Biological Engineering Society, Montreal, Canada, July 2020.

[20] A. Ekhlasi, A. M. Nasrabadi and M. R. Mohammadi, "Analysis of Effective Connectivity Strength in Children with Attention Deficit Hyperactivity Disorder Using Phase Transfer Entropy," Iran J. Psychiatry, vol. 16, no. 4, pp. 374–382, DOI: 10.18502/ijps.v16i4.7224, Oct. 2021.

[21] D. Hai Shi Da Xue, J. Da Xue, M. IEEE Systems, M. IEEE Systems, B. International Federation of Automatic Control. Technical Committee on Economic and Institute of Electrical and Electronics Engineers, 2017 Int. Conf. on Information, Cybernetics and Computational Social Systems (IEEE ICCSS 2017), Dalian, Liaoning, China, July 24-26, 2017.

[22] Galgotias University, Universitatea "Aurel Vlaicu" din Arad, IEEE Industry Applications Society and Institute of Electrical and Electronics Engineers, 2020 IEEE 5th Int. Conf. on Computing Communication and Automation (ICCCA), Galgotias University, Noida, India, Oct. 30-31, 2020.

[23] S. S. Beriha, "Computer Aided Diagnosis System to Distinguish ADHD from Similar Behavioral Disorders," Biomedical and Pharmacology Journal, vol. 11, no. 2, pp. 1135–1141, Jun. 2018.

[24] L. Alice, B. K. S Jacob and S. Ramachandran, "ADHD Detection Based on Time Frequency Image and Deep- learning CNN from Event-related EEG," DOI: 10.21203/rs.3.rs-2986442/v1, 2023.

[25] D. Wang, D. Hong and Q. Wu, "Attention Deficit Hyperactivity Disorder Classification Based on Deep Learning," IEEE/ACM Trans. Comput. Biol. Bioinform., vol. 20, no. 2, pp. 1581–1586, Mar. 2023.

[26] S. M. Omar, M. Kimwele, A. Olowolayemo and D. M. Kaburu, "Enhancing EEG Signals Classification Using LSTM-CNN Architecture," Engineering Reports, DOI: 10.1002/eng2.12827, 2023.

[27] N. Gupte, M. Patel, T. Pen and S. Kurhade, "Early Detection of ADHD and Dyslexia from EEG Signals," Proc. of the 2023 IEEE 8th Int. Conf. for Convergence in Technology (I2CT 2023), DOI: 10.1109/I2CT57861.2023.10126272, 2023.

[28] M. Maniruzzaman et al., "Optimal Channels and Features Selection Based ADHD Detection from EEG Signal Using Statistical and Machine Learning Techniques," IEEE Access, vol. 11, pp. 33570–33583, DOI: 10.1109/ACCESS.2023.3264266, 2023.

[29] A. Afzali et al., "Automated Major Depressive Disorder Diagnosis Using a Dual-input Deep Learning Model and Image Generation from EEG Signals," Waves in Random and Complex Media, pp. 1–16, DOI: 10.1080/17455030.2023.2187237, Mar. 2023.

[30] Y. Sharma and B. K. Singh, "Classification of Children with Attention-Deficit Hyperactivity Disorder Using Wigner-Ville Time-Frequency and Deep expEEGNetwork Feature-based Computational

Models," IEEE Trans. Med. Robot Bionics, vol. 5, no. 4, pp. 890–902, Nov. 2023.

[31] I. W. Huang et al., "Optimal EEG Data Segmentation in LSTM Networks for Learning Neural Dynamics of ADHD," Proc. of the IEEE 2022 Int. Conf. on System Science and Engineering (ICSSE 2022), pp. 33–38, DOI: 10.1109/ICSSE55923.2022.9948260, 2022.

[32] Y. Zhang et al., "ADHD Classification by Feature Space Separation with Sparse Representation," Proc. of the 2018 IEEE 23rd Int. Conf. on Digital Signal Processing (DSP), DOI: 10.1109/ICDSP.2018.8631658, Shanghai, China, 2019.

[33] B. Taghi Beyglou et al., "Detection of ADHD cases using CNN and Classical Classifiers of Raw EEG," Computer Methods and Programs in Biomedicine Update, vol. 2, p. 100080, Jan. 2022.

[34] C. Wang, "Indentification of Autism Spectrum Disorder Based on an Improved Convolutional Neural Networks," Proc. of the 2021 IEEE 3rd Int. Conf. on Machine Learning, Big Data and Business Intelligence (MLBDBI 2021), pp. 235–239, DOI: 10.1109/MLBDBI54094.2021.00051, 2021.

[35] M. Maniruzzaman et al., "Efficient Feature Selection and Machine Learning Based ADHD Detection Using EEG Signal," Computers, Materials and Continua, vol. 72, no. 3, pp. 5179–5195, 2022.

[36] M. Mafi and S. Radfar, "High Dimensional Convolutional Neural Network for EEG Connectivity-based Diagnosis of ADHD," J. of Biomedical Physics and Engineering, vol. 12, no. 6, pp. 645–654, 2022.

[37] 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, DOI: 10.1155/2022/8709145, 2022.

[38] I. A. Ahmed et al., "Eye Tracking-based Diagnosis and Early Detection of Autism Spectrum Disorder Using Machine Learning and Deep Learning Techniques," Electronics (Switzerland), vol. 11, no. 4, DOI: 10.3390/electronics11040530, Feb. 2022.

[39] U. B. Mahadevaswamy and C. Manjunath, "f-MRI Based Detection of Autism Using CNN Algorithm," Proc. of the 2022 IEEE 2nd Mysore Sub Section Int. Conf. (MysuruCon 2022), DOI: 10.1109/MysuruCon55714.2022.9972394, 2022.

[40] B. R. G. Elshoky et al., "Comparing Automated and Non-automated Machine Learning for Autism Spectrum Disorders Classification Using Facial Images," ETRI J., vol. 44, no. 4, pp. 613–623, 2022.

[41] M. R. Kanhirakadavath et al., "Investigation of Eye-tracking Scan Path As a Biomarker for Autism Screening Using Machine Learning Algorithms," Diagnostics, vol. 12, no. 2, 2022.

[42] M. Saini, U. Satija and M. D. Upadhayay, "One-dimensional Convolutional Neural Network Architecture for Classification of Mental Tasks from Electroencephalogram," Biomed Signal Process Control, vol. 74, DOI: 10.1016/j.bspc.2022.103494, Apr. 2022.

[43] S. Altun, A. Alkan and H. Altun, "Automatic Diagnosis of Attention Deficit Hyperactivity Disorder with Continuous Wavelet Transform and Convolutional Neural Network," Clinical Psychopharmacology and Neuroscience, vol. 20, no. 4, pp. 715–724, 2022.

[44] S. Saini, R. Rani and N. Kalra, "Prediction of Attention Deficit Hyperactivity Disorder (ADHD) Using Machine Learning Techniques Based on Classification of EEG Signal," Proc. of the IEEE 8th Int. Conf. on Advanced Comp. and Comm. Systems (ICACCS 2022), pp. 782–786, Coimbatore, India, 2022.

[45] T. Tuncer, S. Dogan and A. Subasi, "LEDPatNet19: Automated Emotion Recognition Model Based on Non-linear LED Pattern Feature Extraction Function Using EEG Signals," Cognitive Neurodynamics, vol. 16, no. 4, pp. 779–790, Aug. 2022.

[46] N. A. Mashudi, N. Ahmad and N. M. Noor, "Classification of Adult Autistic Spectrum Disorder Using Machine Learning Approach," IAES Int. J. of Artificial Intelligence, vol. 10, no. 3, pp. 743–751, 2021.

[47] M. F. Rabbi et al., "A Convolutional Neural Network Model for Early-stage Detection of Autism Spectrum Disorder," Proc. of the 2021 IEEE Int. Conf. on Information and Communication Technology for Sustainable Development (ICICT4SD 2021), pp. 110–114, Dhaka, Bangladesh, Feb. 2021.

[48] G. P. Pralhad et al., "Dyslexia Prediction Using Machine Learning," Proc. of the 2021 1st IEEE Int. Conf. on Artificial Intelligence and Machine Vision (AIMV 2021), DOI: 10.1109/AIMV53313.2021.9671004, Gandhinagar, India, 2021.

[49] Y. H. Chen et al., "Early Detection of Autism Spectrum Disorder in Young Children with Machine Learning Using Medical Claims Data," BMJ Health Care Inform, vol. 29, no. 1, p. 100544, Sep. 2022.

[50] S. Karimi-Shahraki and M. Khezri, "Identification of Attention Deficit Hyperactivity Disorder Patients Using Wavelet-based Features of EEG Signals," JIPET J. of Intelligent Procedures in Electrical Technology, vol. 12, no. 47, pp. 29–40, [Online], Available: http://jipet.iaun.ac.ir/, 2021.

[51] N. Kaur and G. Gupta, "Refurbished and Improvised Model Using Convolution Network for Autism Disorder Detection in Facial Images," Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 2, pp. 883–889, DOI: 10.11591/ijeecs.v29.i2.pp883-889, Feb. 2023.

[52] S. Raj and S. Masood, "Analysis and Detection of Autism Spectrum Disorder Using Machine Learning Techniques," Procedia Computer Science, pp. 994–1004, DOI: 10.1016/j.procs.2020.03.399, 2020.

[53] Z. Sherkatghanad et al., "Automated Detection of Autism Spectrum Disorder Using a Convolutional Neural Network," Frontiers in Neuroscience, vol. 13, DOI: 10.3389/fnins.2019.01325, Jan. 2020.

[54] M. Rakić, M. Cabezas, K. Kushibar, A. Oliver and X. Lladó, "Improving the Detection of Autism

Spectrum Disorder by Combining Structural and Functional MRI Information," NeuroImage: Clinical, vol. 25, DOI: 10.1016/j.nicl.2020.102181, Jan. 2020.

[55] M. Moghaddari, M. Z. Lighvan and S. Danishvar, "Diagnose ADHD disorder in Children Using Convolutional Neural Network Based on Continuous Mental Task EEG," Computer Methods and Programs in Biomedicine, vol. 197, DOI: 10.1016/j.cmpb.2020.105738, Dec. 2020.

[56] Z. Zhao et al., "Applying Machine Learning to Identify Autism with Restricted Kinematic Features," IEEE Access, vol. 7, pp. 157614–157622, DOI: 10.1109/ACCESS.2019.2950030, 2019.

[57] R. Yaghoobi, S. Azadi and P. Keshavarzi, "Loss Detection of Recurrence Rate in the EEG Signals of Children with ADHD," NIScPR Online Periodicals Repository, JSIR, vol.78, no. 04, April 2019.

[58] G. A. Senthil et al., "A Novel Analysis and Detection of Autism Spectrum Disorder in Artificial Intelligence Using Hybrid Machine Learning," Proc. of the IEEE Int. Conf. on Innovative Data Communication Technologies and Application (ICIDCA), pp. 291–296, Uttarakhand, India, 2023.

[59] A. E. Alchalabi et al., "FOCUS: Detecting ADHD Patients by an EEG-based Serious Game," IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 7, pp. 1512–1520, Jul. 2018.

[60] M. R. Mohammadi et al., "EEG Classification of ADHD and Normal Children Using Non-linear Features and Neural Network," Biomedical Engineering Letters, vol. 6, no. 2, pp. 66–73, May 2016.

[61] A. J. Hao, B. Lianghua He and H. Yin, "Discrimination of ADHD Children Based on Deep Bayesian Network," Proc. of 2015 IET Int. Conf. on Biomedical Image and Signal Processing (ICBISP 2015), DOI: 10.1049/cp.2015.0764, Beijing, China, 2015.

[62] M. De Oliveira Meira et al., "Evaluating Brain Regions that Characterize Attention Deficit/Hyperactivity Disorder Based on SPECT Images and Machine Learning Models," Proc. of the IEEE Int. Joint Conf. on Neural Networks, DOI: 10.1109/IJCNN55064.2022.9892968, 2022.

[63] S. K. Khare, N. B. Gaikwad and V. Bajaj, "VHERS: A Novel Variational Mode Decomposition and Hilbert Transform-based EEG Rhythm Separation for Automatic ADHD Detection," IEEE Transactions on Instrumentation and Measurement, vol. 71, DOI: 10.1109/TIM.2022.3204076, 2022.

[64] M. Aceves-Fernandez and M. A. Aceves-Fernandez, "Methodology Proposal of ADHD Classification of Children Based on Cross Recurrence Plots Methodology Proposal of ADHD Classification of Children based on Cross Recurrence Plots," Non Linear Dynamics, vol. 104, pp. 1491-1505, 2021.

[65] S. Sangeetha, R. Uma and R. Valarmathi, "Dyslexia Biomarker Detection Using Machine Learning," Proc. of the 2022 Int. Conf. on Communication, Computing and Internet of Things (IC3IoT 2022), DOI: 10.1109/IC3IOT53935.2022.9767973, 2022.

[66] A. Parashar et al., "Machine Learning Based Framework for Classification of Children with ADHD and Healthy Controls," Intelligent Automation and Soft Computing, vol. 28, no. 3, pp. 669– 682, 2021.

[67] B. Wingfield et al., "A Predictive Model for Paediatric Autism Screening," Health Informatics J., vol. 26, no. 4, pp. 2538–2553, DOI: 10.1177/1460458219887823, Dec. 2020.

[68] S. Liu et al., "Deep Spatio-temporal Representation and Ensemble Classification for Attention Deficit/Hyperactivity Disorder," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1–10, DOI: 10.1109/TNSRE.2020.3019063, 2021.

[69] R. Sujatha et al., "A Machine Learning Way to Classify Autism Spectrum Disorder," International Journal of Emerging Technologies in Learning, vol. 16, no. 6, pp. 182–200, 2021.

[70] T. Akter et al., "Machine Learning-Based Models for Early Stage Detection of Autism Spectrum Disorders," IEEE Access, vol. 7, pp. 166509–166527, DOI: 10.1109/ACCESS.2019.2952609, 2019.

[71] Y. Jayawardana et al., "Analysis of Temporal Relationships between ASD and Brain Activity through EEG and Machine Learning," Proc. of the 2019 IEEE 20th Int. Conf. on Information Reuse and Integration for Data Science (IRI), DOI: 10.1109/IRI.2019.00035, Los Angeles, USA, 2019.

[72] S. Pulido-Castro et al., "Ensemble of Machine Learning Models for an Improved Facial Emotion Recognition," Proc. of the 2021 IEEE URUCON, pp. 512–516, Montevideo, Uruguay, 2021.

[73] P. Thi Viet Huong et al., "Ensemble Learning in Detecting ADHD Children by Utilizing the Non-linear Features of EEG Signal*," Proc. of the 2nd Int. Conf. on Human-centered Artificial Intelligence (Computing4Human 2021), [Online], Available: http://ceur-ws.org, 2021.

[74] R. Holker and S. Susan, "Computer-aided Diagnosis Framework for ADHD Detection Using Quantitative EEG," Part of Book: Lecture Notes in Computer Science (Including Sub-series Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 229– 240, DOI: 10.1007/978-3-031-15037-1_19, Springer, 2022.

[75] S. Batsakis et al., "Data-driven Decision Support for Adult Autism Diagnosis Using Machine Learning," Digital, vol. 2, no. 2, pp. 224–243, DOI: 10.3390/digital2020014, Jun. 2022.

[76] N. Hamedi et al., "An Effective Connectomics Approach for Diagnosing ADHD Using Eyes-open Resting-state MEG," Proc. of the 11th IEEE Int. Conf. on Computer Engineering and Knowledge (ICCKE 2021), pp. 110–114, DOI: 10.1109/ICCKE54056.2021.9721443, Mashhad, Iran, 2021.

[77] A. JothiPrabha, R. Bhargavi and B. V. Deepa Rani, "Prediction of Dyslexia Severity Levels from Fixation and Saccadic Eye Movement Using Machine Learning," Biomed Signal Process Control, vol. 79, DOI: 10.1016/j.bspc.2022.104094, Jan. 2022.

[78] S.-F. Liang et al., "Differentiation between Resting-state fMRI Data from ADHD and Normal Subjects: Based on Functional Connectivity and Machine Learning," Proc. of the 2012 Int. Conf. on Fuzzy Theory and Its Applications (iFUZZY2012), DOI: 10.1109/iFUZZY.2012.6409719, Taiwan, 2012.
[79] Y. Ding, Y. Li, S. Li, Z. Fan and L. Wang, 2011 8th Int. Conf. on Fuzzy Systems and Knowledge Discovery (FSKD 2011) Proceedings, pp. 1349-2073, Shanghai, China, 26-28 July 2011.
[80] A. Alim and M. H. Imtiaz, "Automatic Identification of Children with ADHD from EEG Brain Waves," Signals, vol. 4, no. 1, pp. 193–205, DOI: 10.3390/signals4010010, Mar. 2023.
[81] EEE, IEEE Engineering in Medicine and Biology Society and IEEE Signal Processing Society, IEEE ISBI 2020 Int. Symposium on Biomedical Imaging (2020 Symposium Proc.), Iowa City, Iowa, 2020.
[82] IEEE, International Symposium on Biomedical Imaging, 2018 IEEE 15th Int. Symposium on Biomedical Imaging (ISBI 2018), DOI: 10.1109/ISBI39256.2018, 2018.
[83] H. Alkahtani et al., "Developing System-based Artificial Intelligence Models for Detecting the Deficit Hyperactivity Disorder," Mathematics, vol. 11, no. 22, DOI: 10.3390/math11224698, Nov. 2023.
[84] G. Sharma and A. M. Joshi, "SzHNN: A Novel and Scalable Deep Convolution Hybrid Neural Network Framework for Schizophrenia Detection Using Multichannel EEG," IEEE Transactions on Instrumentation, vol. 71, DOI: 10.1109/TIM.2022.3212040, 2022.
[85] A. Sharma and P. Tanwar, "Identification of Autism Spectrum Disorder (ASD) from Facial Expressions Using Deep Learning," Proc. of the 2022 IEEE Int. Conf. on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON 2022), pp. 478–484, Faridabad, India, 2022.
[86] S. Dalai, Jadavpur University, Institute of Electrical and Electronics Engineers, Kolkata Section and IEEE Signal Processing Society, Kolkata Chapter, Proc. of 2018 IEEE Applied Signal Processing Conference (ASPCON), Venue: Jadavpur University Main Campus, Kolkata, India, Dec. 7-9, 2018.
[87] A. A. Torres-García et al., Biosignal Processing and Classification Using Computational Learning and Intelligence: Principles, Algorithms and Applications, 1st Edn., ISBN: 9780128201251, Elsevier, 2022.

ملخص البحث:

تقدّم هذه الدراسة مراجعة شاملة للأدبيات السابقة تركّز على استخدام تقنيات التّعلّم الآلي والتّعلّم العميق لتوقُّع اضطرابات النّموّ العصبي والكشف عنها؛ مثل التّخلّف العقلي، واضطراب طيْف التّوحُّد، واضطراب نقْص الانتباه وفرط النّشاط، والديسليكسيا، وغيرها. حيث تفتقر المراجعات المتوافرة إلى النّقاشات التّفصيلية لخوارزميات التّعلّم الآلي ومجموعات البيانات ومؤشّرات الأداء المستخدمة في توقُّع اضطرابات النّموّ العصبي والكشف عنها، وتسعى هذه الدّراسة إلى جَسْر هذه الفجوة عبر تناول جانبين منفصلين هما التّوقّع والكشْف. كما تهدف الدّراسة إلى البحث في آخر ما توصّلت إليه الأبحاث العلمية بالمنهجيات والتّحدّيات واتّجاهات البحث المستقبلية بشأن استخدام تقنيات التّعلّم الآلي والتّعلّم العميق في توقُّع اضطرابات النّموّ العصبي والكشف عنها. وتهدف إلى تصنيف الدّراسات السابقة تبعاً للجانبين الرئيسيّين المتمثّلين في توقّع اضطرابات النّموّ العصبي والكشف عنها، إلى جانب التّطرُّق إلى المنهجيات ومجموعات البيانات والمتغيرات ومؤشّرات الآداء الّتي استخدمتها الدراسات السابقة حول الموضوع.

شملت المراجعة الدّراسات المنشورة في المجلات والمؤتمرات المتخصصة والمفهرسة في Scopus في الفترة من عام 2013 إلى عام 2023. واستخدمت المراجعة مصطلحات للبحث، مثل: توقّع اضطرابات النّموّ العصبي، والكشف عن اضطرابات النّموّ العصبي، باستخدام التّعلّم الآلي. وركّز التّحليل على تحديد منهجيات التّعلّم الآلي والتّعلّم العميق، والنّماذج المجمَّعة، وأنواع مجموعات البيانات، بالإضافة إلى المتغيرات ومؤشرات الأداء المستخدمة في الدراسات السابقة. ولقد ألقت نتائج المراجعة الضّوء على أكثر تقنيات التّعلّم الآلي والتّعلّم العميق انتشاراً، والتّحدّيات المرتبطة بالبحث في هذا المجال، واتّجاهات البحث المستقبلية الرّامية إلى تحسين الخدمات المقدَّمة إلى مجتمع اضطرابات النّموّ العصبي؛ من أجل تطوير الرّعاية الصّحية للمصابين بهذه الاضطرابات عبر تقنيات توقّعٍ وكشْفٍ أفضل.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

A NEW APPROACH COMBINING RSA AND ELGAMAL ALGORITHMS: ADVANCEMENTS IN ENCRYPTION AND DIGITAL SIGNATURES USING GAUSSIAN INTEGERS

Yahia Awad[1], Douaa Jomaa[1],Yousuf Alkhezi[2], Ramiz Hindi[1]

(Received: 16-Jul.-2024, Revised: 11-Sep.-2024, Accepted: 2-Oct.-2024)

ABSTRACT

This article introduces a novel approach that integrates the ElGamal and RSA algorithms to advance the security and efficiency of public-key cryptosystems. By combining these two established asymmetric-key algorithms, our method leverages their individual strengths and addresses the limitations of traditional systems, particularly in relation to the integer-factorization and discrete-logarithm problems. The application of Gaussian integers enhances the robustness of both encryption and digital signature processes, offering a more secure cryptographic framework. Our study involves a comprehensive analysis of the integrated algorithms, including practical implementations and extensive cryptanalytic evaluations focused on the integer-factorization and discrete-logarithm challenges. Quantitative assessments are provided to evaluate the effectiveness and computational efficiency of the proposed system. While key generation is slightly slower compared to using RSA or ElGamal individually, our approach delivers comparable performance in encryption and decryption, with notable improvements in robustness and versatility. In contrast to existing research predominantly focused on optical-image processing, our work extends the scope to a broader range of applications, enhancing both theoretical insights and practical implementations of cryptographic schemes. Future research will focus on optimizing key generation, exploring integration with existing security frameworks and evaluating performance in diverse real-world scenarios to further refine and validate the proposed approach.

KEYWORDS

Combined RSA-ElGamal public-key cryptosystem, RSA, ElGamal, Digital signature, Gaussian integers.

1. INTRODUCTION

Cryptography, an intricate fusion of art and science, has long been fundamental to ensuring secure communication throughout human history. From early simple ciphers to today's sophisticated digital-encryption techniques, the field has continually adapted to meet increasing demands for data security. In the contemporary digital era, where massive volumes of information are exchanged and stored globally, the urgency for robust and adaptable encryption solutions has never been greater. Public-key cryptography represents a significant breakthrough, revolutionizing security protocols with its dual-key system: a public key for encryption and a private key for decryption. This innovative approach allows for secure communication even when the encryption method is known, relying on the mathematical intricacies of cryptographic processes to maintain confidentiality and trust. For further details, see [4][8][11][21][25][36] and the references therein.

As computational power advances and cyber-threats become more sophisticated, the field of public-key cryptography continues to evolve. Recent research has made significant strides in several key areas. Extensions of classical systems, such as RSA, ElGamal and Rabin, have been explored through their application in Gaussian integers and finite fields, enhancing their security and resilience against attacks [6]-[7], [13]-[15]. Hybrid encryption systems, like the one introduced by Kuppuswamy et al. [24], combine public and private-key algorithms to enhance security and authentication. Novel hybrid algorithms, including the HRSA proposed by Panda et al. [28], use multiple prime numbers to complicate factorization, while Iswari et al. [22] and Ahmed et al. [3] have combined RSA with ElGamal and integrated integer factorization with discrete logarithms to improve efficiency and security. Additionally, Adeniyi et al. [2] have focused on integrating RSA and ElGamal with hash functions to bolster data integrity through enhanced digital signatures. Meanwhile, numerous studies have addressed public-key cryptosystems' application in optical-image processing, tackling specific

1. Y. Awad, D. Jomaa and R. Hindi are with Department of Mathematics and Physics, Lebanese Int. University, Faculty of Arts and Sciences, Bekaa Campus, Lebanon. Emails: yehya.awad@liu.edu.lb, o.a.douaa@gmail.com and r.math090@gmail.com
2. Y. Alkhezi is with College of Basic Education Mathematics Department, Public Authority for Applied Education and Training, Kuwait. Email: ya.alkhezi@paaet.edu.kw

challenges and opportunities in this specialized field [5][19][21][29][36]. These contributions advance security measures, but are often confined to particular applications.

The novelty of our research lies in the innovative integration of RSA and ElGamal algorithms, which are traditionally viewed as distinct entities in cryptographic practice. By strategically merging these two algorithms, we have developed a combined RSA-ElGamal public-key cryptosystem that harnesses their individual strengths while mitigating their respective weaknesses. This novel approach not only enhances the overall security of the system, but also provides a versatile framework adaptable to various cryptographic functions, including encryption, decryption and digital signatures. Our work is distinguished by a thorough analysis of the mathematical foundations of this new approach, rigorous cryptographic evaluations and a comprehensive comparative study. These elements collectively advance the field of cryptography, offering deeper insights and new possibilities for future developments in secure communication protocols.

The structure of this paper is as follows: Section 1 introduces the research objectives and context. Section 2 provides an overview of the essential mathematical concepts relevant to our work. In Section 3, we present our public-key generation, encryption and decryption algorithms, supported by formal proofs and numerical examples. Section 4 introduces our combined RSA-ElGamal algorithms and ElGamal digital-signature scheme, detailing key generation, signature creation and verification processes. Section 5 focuses on the security analysis of our combined RSA-ElGamal cryptosystem, evaluating its efficiency and comparing it with classical RSA and ElGamal schemes. This section also includes a comparative complexity analysis, offering insights into the computational costs and advantages of our proposed system.

2. PRELIMINARIES

In this section, we provide a concise overview of the mathematical concepts required for our work. For additional details, please refer to [9], [10] and [25].

2.1 Arithmetic in Z

In algebra, it is widely known that if we consider a group G and an element g within that group, the order of g, represented as,$|g|$ refers to the smallest positive integer t for which $g^t \equiv e$. Furthermore, if there exists an element g in a group G such that G can be generated entirely by g, denoted as $G = \langle g \rangle = \{g^n | n \in Z\}$, we say that G is a cyclic group and g is known as the generator of G, where the order of g is equal to the order of G (i.e., $|g| = |G|$). Euler's phi function, represented as $\phi(n)$, denotes the count of positive integers that are both relatively prime to n and less than n. Additionally, the set of $\phi(n)$ integers that are relatively prime to n and do not contain different elements congruent to each other modulo n is referred to as a reduced residue system modulo n, denoted as Un. This set U_n is cyclic if and only if n takes on the values 2, 4, p^k or $2p^k$, where p is an odd prime and $k \geq 1$. For more information, we refer to [9] and the references therein.

Theorem 2.1 [25] (Euler's Theorem) If n is a positive integer and a is an integer relatively prime to n, then $a^{\phi(n)} \equiv 1 (\text{mod } n)$.

Theorem 2.2 [25] (Fermat's Theorem) Let a be a positive integer and p be any prime number. If p doesn't divide a, then $a^{p-1} \equiv 1 (\text{mod } p)$.

2.2 Arithmetic in Z[i]

The domain of Gaussian integers is the subring $|Z_i| = \{x + iy | a, b \in Z \text{ and } i^2 = -1\}$. It is well known that Z[i] is an Euclidean domain of norm $N(z) = x^2 + y^2$. Let γ be a Gaussian integer. If γ divides 1, then γ is called a unit. As γ is a unit, we call $\gamma\alpha$ an associate of the Gaussian integer α. An element $\gamma \in$ Z[i] is said to be a unit if and only if $N(\gamma) = 1$. This implies that the only units in Z[i] are 1, -1, i and $-i$. If a non-zero non-unit Gaussian integer π is divisible only by units and associates, then it is called a Gaussian prime. The only Gaussian primes are $1 \pm i$, those Gaussian integers π such that $N(\pi) = \pi\bar{\pi}$ which is a natural prime number of the form $4k + 1$ and those natural prime numbers of the form $4k + 3$. For more information, see [10], [20] and [27].

Definition 2.1 [10] The complete residue system modulo β ∈ Z[i] is the set $A(\beta) = \{z \mid z \in \langle\beta\rangle\}$.

Theorem 2.3 [10] Suppose that γ and β are any two non-zero relatively prime Gaussian integers. Then, $A(\gamma\beta) = \{s + r\gamma : s \in A(\gamma),\ r \in A(\beta)\}$.

Theorem 2.4 [10] For any positive integer n, if we consider α =1+i, p as a Gaussian prime in the form $4k + 3$ and π as a Gaussian prime where $N(\pi) = \pi\bar{\pi}$ is a natural prime number q in the form $4k + 1$, then the complete residue systems modulo prime powers in Z[i] are given as follows:

1. $A(\alpha^{2n}) = \{x + iy : 0 \le x \le 2^n - 1,\ 0 \le y \le 2^n - 1\}$ and it has an order of 2^{2n}.
2. $A(\alpha^{2n+1}) = \{x + iy : 0 \le x \le 2^{n+1} - 1,\ 0 \le y \le 2^n - 1\}$ and it has an order of 2^{2n+1}.
3. $A(p^n) = \{x+iy : 0 \le x \le p^n - 1, 0 \le y \le p^n - 1\}$ and it has an order of p^{2n}.
4. $A(\pi^n) = \{x : 0 \le x \le q^n - 1\}$ and it has an order of q^n.

Theorem 2.5 [10] For any positive integer n, if we consider α =1+i, p as a Gaussian prime in the form $4k + 3$ and π as a Gaussian prime where $N(\pi) = \pi\bar{\pi}$ is a natural prime number q in the form $4k + 1$, then the reduced residue systems modulo prime powers in Z[i] are given as follows:

1. $R(\alpha^n)=\{x + iy \in A(\alpha^n) : x \not\equiv y \pmod 2\}$ *and* it has an order of $\phi(\alpha^n) = 2^n - 2^{n-1}$.
2. $R(p^n)=\{x+iy \in A(p^n)$: $\gcd(x, p)\sim 1$ *or* $\gcd(y, p)\sim 1\}$and it has an order of $\phi(p^n)=p^{2n-2}(p^2-1)$.
3. $R(\pi^n)=\{x \in A(\pi^n)$: $\gcd(x, q)\sim 1\}$ *and* it has an order of $\phi(\pi^n) = q^{n-1}(q-1)$.

Remark 2.1 [10] Let β be a Gaussian integer, then the factor ring of Z[i] modulo $\langle\beta\rangle$ is the set of all cosets of $\langle\beta\rangle$ denoted by $G\beta$ or Z[i]/ $\langle\beta\rangle$. Its elements are the equivalence classes of the form $[x+iy] = (x+iy)+\langle\beta\rangle$. The operations are defined by $[\alpha]+[\gamma] = [\alpha + \gamma]$ and $[\alpha][\gamma] = [\alpha\gamma]$, for every $\alpha, \gamma \in$ Z[i]/ $\langle\beta\rangle$. Note that the order of a factor ring modulo $\langle\beta\rangle$ is equal to the number of elements in $A(\beta)$. G_β is a complete residue system modulo β and of order $q(\beta)$. In addition, the units form a group under multiplication, denoted by $U(\beta)$ or G_β^*, which is the reduced residue system modulo β.

Definition 2.2 [10] Let β be a Gaussian integer, then the order of G_β^* is defined as $\phi(\beta)$, which is the extension of Euler's phi function to be the domain of Gaussian integers Z[i].

Theorem 2.6 [10] G_β^* is cyclic if and only if β is of the form $\alpha, \alpha^2, \alpha^3, \pi^n, p, \alpha\pi^n$ *or* αp.

Theorem 2.7 [10] Suppose that $\eta = \beta_1\beta_2$ is a composite Gaussian integer such that both β_1 and β_2 are odd prime integers of the form $4k_1 + 3$ and $4k_2 + 3$, respectively. Then, the complete residue system modulo η is the set $G_\eta = \{x + iy : 0 \le x \le \beta_1\beta_2 - 1,\ 0 \le y \le \beta_1\beta_2 - 1\}$.

2.3 Classical RSA Public-key Cryptosystem

The RSA public-key cryptosystem is widely recognized as one of the most prominent cryptographic systems, initially introduced by Ronald Rivest, Adi Shamir and Leonard Adleman in 1977 (refer to [32]). The security of RSA is rooted in two fundamental problems: the integer-factorization problem and the RSA problem. The integer-factorization problem involves finding the prime factorization of a positive integer $n = p_1^{e_1}p_2^{e_2} \ldots p_k^{e_k}$, where p_i's are distinct primes and $e_i \ge 1$. On the other hand, the RSA problem entails finding an integer m that serves as the e^{th} root of c modulo a composite integer n. In this scenario, n is a product of two distinct odd primes p and q and e is a positive integer satisfying $\gcd(e, (p-1)(q-1))=1$. It is widely acknowledged that while the integer-factorization problem and the RSA problem share similarities, this resemblance has not been formally proven yet (see [8] and [25]).

The RSA cryptosystem operates through the following steps: Entity A generates two large, distinct random primes, p and q (approximately of the same size). They compute $n = pq$ and $\phi(n) = (p-1)(q-1)$ and then choose a random integer e such that $1 < e < \phi(n)$ and $\gcd(e, \phi(n)) = 1$. Entity A also computes the multiplicative inverse of e modulo $\phi(n)$ and obtains the value d. The resulting public key is denoted as (n, e), while the private key is denoted as (p, q, d). To encrypt a plaintext $m \in Zn$, entity B employs the public key (n, e) to compute the ciphertext $c \equiv m^e \pmod n$ and transmits it to entity A. Subsequently, entity A utilizes the private key d to recover the original plaintext by computing $m \equiv c^d \pmod n$.

2.4 Classical ElGamal Public-key Cryptosystem

The ElGamal public-key cryptosystem, introduced by Taher ElGamal in 1985 (refer to [12]), stands as a widely adopted and robust cryptographic technique. Its security is fundamentally based on the discrete logarithm problem (DLP), which poses the challenge of finding an integer k within the range of $0 \leq k \leq p-1$, such that $\alpha^k \equiv \beta (\text{mod } p)$, with p denoting a prime, α serving as a generator of Z_p^* and β representing an element in Z_p^*.

The ElGamal cryptosystem unfolds as follows: Entity A initiates the process by generating a large random prime integer p, along with a generator α of the multiplicative cyclic group Z_p^*. Subsequently, a random integer a is selected, adhering to the condition $1 \leq a \leq p-2$. Entity A then computes $\alpha^a (\text{mod } p)$. The resulting public key is represented as (p, α, α^a), while the private key remains as a.

To encrypt a plaintext $m \in Z_p$, entity B proceeds by choosing another random integer k, satisfying $1 \leq k \leq p-2$. Subsequently, $\gamma \equiv \alpha^k (\text{mod } p)$ and $\delta \equiv m(\alpha^a)^k (\text{mod } p)$ are computed. The resulting ciphertext is then given by $c = (\gamma, \delta)$. Finally, for the decryption and recovery of the plaintext, entity A applies the private key a to compute $\gamma^{p-1-a} (\text{mod } p)$, from which the original message m is obtained as $m = (\gamma^{-a}).\delta (\text{mod } p)$.

2.5 RSA and ElGamal Digital Signatures

Let's define some notations before discussing the RSA and ElGamal signature algorithms, including key generation, signature and verification algorithms (refer to [25]).

2.5.1 Prerequisite Notations

1. M (Message Space): This represents a collection of elements to which a signer can attach a digital signature.
2. M_S (Signing Space): It refers to a collection of components on which the signature transformations are applied.
3. S (Signature Space): It denotes a collection of items in M that are associated with messages. These components establish a link between the signer and the message.
4. R (Redundancy Function): It represents a one-to-one mapping from M to M_S. It is important that R is not multiplicative, meaning that $R(ab) \neq R(a)R(b)$ for all pairs of relatively prime elements a and b in M.
5. M_R: It refers to the image of R.
6. R^{-1}: It represents the inverse of R and maps elements from M_R back to M.
7. h (Hash Function): It is a one-way function with its domain defined as M.
8. M_h (Hash Value Space): If $h : M \rightarrow M_h$, then M_h is a sub-set of M_S.

2.5.2 Hash Function

The hash function [25] is a fundamental cryptographic tool widely employed in protocols. It generates a hash value denoted as $\tilde{m} = h(m)$, a concise, fixed-length bit string used to represent a specific message (e.g. fingerprints). To ensure the security of the hash function, three fundamental properties must be satisfied:

1. Preimage Resistance (or the one-way property): This ensures computing the original message m given that the hash value m is computationally infeasible.
2. Weak Collision Resistance: A form of pre-image resistance, making it computationally infeasible to find two distinct messages $m_1 \neq m_2$ that produce the same hash values; i.e., $m_1 = m_2$.
3. Collision Resistance: It ensures it's challenging to find two distinct inputs $m_1 \neq m_2$ that hash to the same value; i.e., $h(m_1) = h(m_2)$.

Thus, it's crucial to highlight that when dealing with the hash-value representation of a message, both signature generation and verification operate on the hash value itself rather than the original message. Moreover, digital signatures are broadly categorized into two main types.

2.5.3 Digital Signature

There are two types of digital signatures

1. Digital signature with an appendix: This type of signature requires the original message as an input during the verification process. It utilizes cryptographic hash functions instead of custom redundancy functions, making it less vulnerable to existential forgery attacks. One example of this method is the ElGamal signature, introduced by Taher ElGamal in 1985. It is a digital-signature scheme that relies on the discrete-logarithm problem (DLP). It is a probabilistic algorithm used to generate digital signatures for messages of any length. The scheme requires a hash function, denoted as h, which maps messages to integers modulo a large prime number p. It is described as follows: Entity B signs the message $m \in Z_p$ by selecting a random secret integer k, such that $1 \leq k \leq p\text{-}2$ with $\gcd(k, p\text{-}1) = 1$. Then, entity B computes $r \equiv \alpha^k \pmod{p}$, $k^{-1} \pmod{p\text{-}1}$ and $s \equiv k^{-1}\,(h(m)\text{-}ar) \pmod{p\text{-}1}$. The signature is (r, s). Now, entity A verifies B's signature by verifying that $1 \leq r \leq p\text{-}1$, otherwise the signature is rejected. Then, entity A computes $v_1 \equiv (\alpha^a)^r r^s \pmod{p}$ and $v_2 \equiv \alpha^{h(m)} \pmod{p}$. The signature is accepted if and only if $v_1 = v_2$. For further information, please see [25].
2. Digital signature with message recovery: Unlike the previous type, this method does not need the original message for the verification process. The original message can be extracted from the signature. The RSA signature is an example of a technique that provides digital signatures with message recovery. It was introduced in 1978 and it is the most commonly used digital-signature system in practice since its verification process is fast and easy. Its security is also based on the integer-factorization problem. It is described as follows: Entity B signs a message $m \in M$ by computing $\tilde{m} = R(m)$, where $m \in [0, n-1]$ and computes the signature $s \equiv \tilde{m}^d \pmod{n}$. Now, entity A verifies B's signature by computing $\tilde{m} \equiv s^e \pmod{n}$, which should be in M_R and recovers $m = R^{-1}(m)$. For further information, please see [25].

Remark 2.2 We will employ the hash function $h(m) = m^3$ with a specified modulus depending on the cryptographic context. In the case of ElGamal encryption, we take mod p, where p is a prime, while in RSA encryption, we take mod n, where n is a composite integer. This hash function is chosen for multiple reasons. Firstly, it maintains pre-image resistance, making it computationally difficult to find any pre-image m given $\tilde{m}$, such that $h(m) \equiv \tilde{m}$ mod modulus. Secondly, it upholds second preimage resistance since, given a pre-image m_1, it is computationally infeasible to generate another distinct preimage m_2 such that $h(m_1) \equiv h(m_2)$ mod modulus and $m_1 \neq m_2$. Thirdly, it preserves collision resistance, making it computationally impracticable to discover any two distinct inputs m_1 and m_2 where $m_1 \neq m_2$ and $h(m_1) \equiv h(m_2)$ mod modulus.

3. COMBINED RSA-ELGAMAL ALGORITHMS AND ELGAMAL CRYPTOSYSTEM

In this section, we present a novel combined RSA-ElGamal public-key encryption scheme that combines the RSA and ElGamal encryption schemes. We provide the algorithms for public-key generation, encrypion and decryption, along with accompanying proofs. Additionally, we illustrate the concepts with a numerical example.

3.1 Methodology

The ElGamal public-key cryptosystem relies on the discrete-logarithm problem, while the strength of the RSA public key cryptosystem lies in the difficulty of the integer-factoring problem. In this proposal, we present a novel algorithm that combines both RSA and ElGamal public-key cryptosystems. To achieve this, we first implement a modified ElGamal scheme using Gaussian integers and then utilize the RSA scheme in the domain of Gaussian integers.

Here is a brief overview of the process: We start by generating a large prime number p along with a generator α for the group G_p^*. Next, we select a random positive integer $a < p^2\text{-}1$ and compute $\alpha^a \pmod{p}$. Following that, we choose two Gaussian primes q and r and find their product $\eta = qr$. Subsequently, we select a random integer e and using the extended Euclidean algorithm, we determine its unique inverse $d \in G_\eta$, ensuring that $\gcd(e, \phi(\eta)) = 1$ and $1 < e, d < \phi(\eta)$. The resulting public key is given by $(p, \alpha, \alpha^a, \eta, e)$ and the private key is represented as (a, q, r, d).

To encrypt a message $m \in G_p$, we randomly choose a positive integer $k< p^2 -1$ and compute the ciphertext $c \equiv \text{Me}(\text{mod } \eta)$, where $M = \gamma + \delta i$ with $\gamma \equiv \alpha^k$ and $\delta \equiv m(\alpha^a)^k$, which are elements in G_p. For the decryption of the sent ciphertext c, we utilize the private keys a and d to recover the original message. This is achieved by computing

$$m = \left[\left(\left(Re\left(c^d mod\ \eta\right)\right)^{q(p)-1-a}(modp)\right).\left(Im\left(c^d mod\eta\right)(modp)\right)\right](mod\eta). \quad (1)$$

3.1.1 Choice of the Gaussian Primes

In the following discussion, we will present an analysis of the primes p, q and r that will be selected in our novel approach. Initially, ElGamal scheme will be applied within the complete residue system G_p, which is defined as mentioned in Theorems 2.4 and 2.7 as follows:

1. If p is any natural prime integer, then $G_p = Z_p$.
2. If p is a Gaussian prime such that $p\bar{p}$ is a natural prime of the form $4k$+1, then $G_p = \{x: 0 \leq x \leq p\bar{p} - 1\}$ and it has an order of $q(p) = p\bar{p}$.
3. If p is a Gaussian prime of the form $4k$+3, then $G_P = \{x + iy: 0 \ll x \leq p - 1, 0 \leq y \leq p - 1\}$ and it has an order of $q(p) = p^2$.

For the sake of simplicity, we can utilize the initial implementation. Nevertheless, we shall employ the third implementation.

Second, the RSA scheme will be implemented in the complete residue system G_η such that η is a product of two Gaussian primes q and r; i.e., $\eta = qr$, where we have three possible cases:

1. If $q = \pi_1$ and $r = \pi_2$, where $\pi_1\overline{\pi_1}$ and $\pi_2\overline{\pi_2}$ are two prime integers of the form $4k + 1$, then the complete residue system modulo η is $G_\eta = \{x + qy : x \in G_q, y \in G_r\}$ and of order $q(\eta) = qr$. But, this case will be neglected due to its similarity to the classical settings.
2. If $q = \pi_1$ is a Gaussian prime such that $\pi_1\overline{\pi_1}$ is a prime integer of the form $4k + 1$ and r is a prime integer of the form $4k + 3$, then the factorization of $\eta = \pi_1 r$ which has the form $x+ yi$ could be easily solved by simply finding the gcd(x, y) which will be equal to r. Hence, this case will be also neglected, since our aim is to ensure the infeasibility of the factorization of η.
3. If q and r are both Gaussian primes of the form $4k + 3$, then the complete residue system modulo η is $G_\eta = \{x + qy: x \in G_q, y \in G_r\}$ and of order $q(\eta) = q^2r^2$, which is huge enough to enhance the security of our approach compared to that of the classical one. Hence, this case will be chosen, since it is the best choice for the new implementation of the RSA scheme.

Thus, to provide a clearer justification: when using Gaussian primes of the form $4k + 3$ for both q and r, the order of G_η is $q(\eta) = q^2r^2$, meaning that the message space is not just doubled, but squared. This increase in size is crucial, because it exponentially expands the variety of possible plaintexts, making brute-force attacks, including exhaustive search methods, computationally infeasible. The complexity of deciphering the original message from the ciphertext becomes exponentially harder, requiring much more effort than in classical RSA systems with the same prime numbers.

Moreover, by increasing the size of the message space, the number of possible combinations of plaintexts grows exponentially. This means that any adversary attempting to recover the plaintext would face a significantly more difficult task, as the size of the problem space grows much larger. Traditional algorithms for factorization or solving the discrete-logarithm problem become less effective, further strengthening the cryptographic security of our approach.

3.1.2 Choice of Plaintext *m*

The plaintext, denoted as $m \in G_p$, can be expressed in two possible forms. The first form is $m=x+iy$, where both $x, y \in Z_p$ and $y \not\equiv 0 \pmod p$. The second form is $m = x$, where $x \in Z_p$.

3.2 Combined RSA-ElGamal Algorithms and ElGamal Public-key Scheme

In the subsequent sub-sections, we present a comprehensive explanation of our novel concept for

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

the "Combined RSA-ElGamal public-key cryptosystem." We elucidate the procedures for key generation, encryption and decryption in the following manner:

Algorithm 3.1 Key generation for the combined RSA-ElGamal public-key scheme by entity A.

1. Generate three distinct large random odd prime integers p, q and r of the form $4k+3$ and approximately the same size.
2. Find a generator α of G_p^*.
3. Select a random integer a, such that $2 \leq a \leq p^2-2$ *and* then compute $\alpha^a \pmod p$.
4. Compute $\eta = qr$ *and* $\phi_\eta = (q^2-1)(r^2-1)$.
5. Select a random integer e such that $1 < e < \phi_\eta$ *and* $\gcd(e, \phi_\eta) = 1$.
6. Use the extended Euclidean division algorithm to compute d, such that $ed \equiv 1 \pmod{\phi_\eta}$.
7. The public key is $(p, \alpha, \alpha^a, \eta, e)$ *and* the private key is (a, q, r, d).

Algorithm 3.2 Combined RSA-ElGamal public-key encryption by entity *B*.

1. Obtain A's public key $(p, \alpha, \alpha^a, \eta, e)$.
2. Choose a random integer k, such that $2 \leq k \leq p^2-2$.
3. Compute the ciphertext $c \equiv M^e \pmod \eta$, where $M = \gamma + \delta i$, $\gamma \equiv \alpha^k \pmod p$ and $\delta \equiv m(\alpha^a)^k \pmod p$.
4. Send the ciphertext c to entity A.

Algorithm 3.3 Combined RSA-ElGamal public-key decryption.

By using the private keys a and d, entity A recovers the plaintext m such that:

$$m \equiv (\mathrm{Re}(c^d \bmod \eta))^{p^2-1-a} \pmod p \; (\mathrm{Im}(c^d \bmod \eta)) \pmod p \pmod \eta.$$

Theorem 3.1 The original message m is recovered by reducing

$$\left[\left(\left(Re(c^d mod\ \eta)\right)^{p^2-1-a}(mod\ p)\right).\left(Im(c^d mod\ \eta)(mod\ p)\right)\right](mod\ \eta).$$

Proof 3.1 Consider the Gaussian integer $m' \in G_p$ such that

$$m' \equiv \left[\left(\left(Re\left(c^d mod \eta\right)\right)^{p^2-1-a}(mod\ p)\right).\left(Im\left(c^d mod\ \eta\right)(mod\ p)\right)\right](mod \eta) \tag{2}$$

Since $ed \equiv 1 \pmod{\phi(\eta)}$, then there exists an integer k', such that $ed = 1 + k'\phi(\eta)$. Hence, there are two cases:

1. Suppose that the $\gcd(M, q)=1$. Then, by using the modified Euler's theorem to the domain of Gaussian integers, we have $M^{\phi(\eta)} \equiv 1 \pmod \eta$. After raising both sides of the congruence to the power of k' and then multiplying them by M. We get,
$$M^{1+k'\phi(\eta)} \equiv M^{ed} \equiv c^d \pmod \eta \equiv M \pmod \eta. \tag{3}$$
2. Suppose that $\gcd(M, q) = q$. Then, we have $M \equiv 0 \pmod q$. Hence, $M^{k'(q^2-1)(r^2-1)} \equiv 0(mod\ q)$. After multiplying both sides by M, we get $M^{1+k'(q^2-1)(r^2-1)} \equiv 0(mod\ q)$ and hence, $M^{1+k'\emptyset(\eta)} \equiv M^{ed} \equiv c^d \equiv 0(mod\ q)$, since $M \equiv 0 \pmod q$. Then, $c^d \equiv M \pmod q$. By the same argument, we also get $c^d \equiv M \pmod r$. Since q and r are two distinct Gaussian primes, we obtain that $c^d \equiv M \pmod \eta$.

Hence, for any Gaussian integer M, we have $c^d \equiv M \pmod \eta$. Therefore,

$$\begin{aligned} m' &\equiv \left[\left(\left(Re(c^d mod\ \eta)\right)^{p^2-1-a}(mod\ p)\right).\left(Im(c^d mod\ \eta)(mod\ p)\right)\right](mod\ \eta) \\ &\equiv \left[\left(\left(Re(M)\right)^{p^2-1-a}(mod\ p)\right).\left(Im(M)(mod\ p)\right)\right](mod\ \eta) \end{aligned} \tag{4}$$

But, $M = \gamma + \delta i$. Then,

$$\begin{aligned} m' &\equiv \left[\left(\gamma^{p^2-1-a}(mod\ p)\right).\left(\delta(mod\ p)\right)\right](mod\ \eta) \equiv \left[\left(\alpha^{-ak}(mod\ p)\right).\left(\delta(mod\ p)\right)\right](mod\ \eta) \\ &\equiv [\alpha^{-ak}.m.\alpha^{ak} \pmod p] \pmod \eta \;\equiv\; m \pmod{qr}. \end{aligned} \tag{5}$$

Example 3.1 (Combined RSA-ElGamal public-key scheme) Entity A generates the keys as follows: If $p = 3$, and $\alpha = 2$ is a generator of G_3^*, then entity A chooses the private key $a = 2$ and computes $\alpha^a = 1(\text{mod } 3)$. Also, if $q = 7$ and $r = 11$, then entity A computes $\eta = 77$ and $\phi(\eta) = 5760$. After that, entity A chooses $e = 971$ and by using the extended Euclidean division algorithm, finds $d = 611$ such that $ed \equiv 1(\text{mod } \phi(\eta))$. The public-key is (3, 2, 1, 77, 971) and the private key is (2, 7, 11, 611). Now entity B encrypts the message $m = 2$ by selecting a random integer $k = 6$ and computing $\gamma \equiv 1(\text{mod } 3)$ and $\delta \equiv 2(\text{mod } 3)$. Then, entity B assumes that $M = 1+2i$ and computes $c = 1-24i$. Entity B then sends c to entity A, which decrypts and recovers the message m by computing $m \equiv [((\text{Re}(c^d \text{ mod } \eta))^{p2-a-1}(\text{mod } p)). (\text{Im}(c^d \text{ mod } \eta)(\text{mod } p))](\text{mod } \eta) \equiv 2$.

3.3 Security of the Proposed Combined RSA-ElGamal Cryptosystem

As the new proposed scheme combines elements of both the modified ElGamal and RSA schemes, each relying on distinct mathematical problems (the discrete-logarithm problem and the integer-factorization problem, respectively), the security of our combined RSA-ElGamal public-key scheme is predicated on both of these cryptographic challenges. To decrypt a message encrypted using this new scheme, one must first solve the integer-factorization problem, followed by solving the discrete-logarithm problem to obtain the plaintext. Consequently, the time required to compromise the new proposed scheme is influenced by the hacking times of both classical ElGamal and RSA schemes, as demonstrated in the comparative study outlined in Section 5. Additionally, the new scheme implements RSA in the domain of Gaussian integers by generating two odd primes, designated as q and r, in the form of $4k + 3$. This choice results in the complete residue system $A(\eta)$ containing q^2r^2 elements, as opposed to just qr elements in the classical scheme. Moreover, if we implement the ElGamal in the domain of Gaussian integers modulo a Gaussian prime p of the form $4k + 3$, the cyclic group G_p^* has p^2-1 elements and the private key a can range from 2 to p^2-1. In contrast, the cyclic group of the classical scheme, Z_p^*, has p-1 elements and the private key a can range from 2 to p-1. Consequently, with equivalent effort to that in classical settings, our new scheme offers an expanded set of choices for plaintext and private keys by more than the square of the choices in the classical case. This extension bolsters the security provided by the new proposed scheme without necessitating any additional efforts.

4. COMBINED RSA-ELGAMAL SIGNATURE SCHEME

In this section, we introduce our proposed signature called the combined RSA-ElGamal signature scheme, where the key generation, signature and verification algorithms are given with proofs and a numerical example.

4.1 Description of the Combined RSA-ElGamal Signature

The concept behind our proposed signature arises from the necessity to enhance the security of our cryptosystem. Our signature approach combines elements from the classical ElGamal signature and the modified RSA signature within the domain of Gaussian integers. Its security is dependent on both the discrete-logarithmic and integer-factorization problems. In our proposed signature scheme, the message space, denoted as M, is represented by Z_p, while the ciphertext signing and signature spaces are all denoted as G_η. The redundancy function, denoted as $R : Z_p \rightarrow G_\eta$, can be made public and the hash function, denoted as $h : M \rightarrow Z_p$, is selected in a manner such that p represents a large prime number.

The procedure is as follows: Firstly, a natural prime integer p is chosen, along with a generator α for Z_p^*. Then, a random positive integer a is selected such that $a < p$-1 and $\alpha^a(\text{mod } p)$ is computed. In the next step, two Gaussian primes, q and r, are chosen in the form $4k + 3$ and their product $\eta = qr$ is determined. Following this, a random integer e is selected and its unique (up to associates) inverse $d \in G_\eta$ is calculated using the extended Euclidean algorithm, satisfying gcd $(e, \phi(\eta)) = 1$ and $1<e, d<\phi(\eta)$. The public key comprises $(p, \alpha, \alpha^a, \eta, e)$, while the private key comprises (a, q, r, d). To sign a message $m \in Z_p$, a random positive integer k is chosen such that $k<p$-1 and ς is computed as $\varsigma \equiv z^d(\text{mod } \eta)$, where $z = r' + si = R(\tilde{m})$, with $r' \equiv \alpha^k(\text{mod } p)$, and $s \equiv k^{-1}(\tilde{m} - ar')(\text{mod } p\text{-}1)$. To verify the signature ς and recover the original message m, z is calculated as $z \equiv \varsigma^e(\text{mod } \eta)$, where it should belong to M_R and $\tilde{m}$ is recovered such that $\tilde{m} = R^{-1}(z)$.

Finally, $h(m)$, v_1 and v_2 are computed so that $h(m) \equiv ks+ar' (\text{mod } p\text{-}1)$, $v_1 \equiv y^{\text{Re}(z)}. \text{Re}(z)^{\text{Im}(z)} (\text{mod } p)$, where $1 \leq \text{Re}(z) \leq p\text{-}1$ and $v_2 \equiv \alpha^{h(m)} (\text{mod } p)$. The signature is accepted if $v_1 = v_2$.

The aforementioned description is presented in a step-by-step manner in the following algorithms.

Algorithm 4.1 Key generation for the combined RSA-ElGamal signature scheme:

1. Generate a random large odd prime p and a generator α *of* Z_p^* *and* choose a random integer a, where $1 \leq a \leq p-2$.
2. Compute $y \equiv \alpha^a (\text{mod } p)$.
3. Generate two large, distinct odd primes, q and r, each of roughly the same size.
4. Compute $\eta = qr$ *and* $\phi(\eta) = (q^2\text{-}1)(r^2\text{-}1)$.
5. Select a random integer e such that $1 \leq e \leq \phi(\eta)$ with $\gcd(e, \phi(\eta)) = 1$.
6. Use the extended Euclidean algorithm to compute the unique integer d, such that $ed \equiv 1 (\text{mod } \phi(\eta))$.
7. The public key is (p, α, y, η, e) *and* the private key is (a, q, r, d).

Algorithm 4.2 Combined RSA-ElGamal signature generation by entity B.

1. Select a random secret integer k, such that $1 \leq k \leq p\text{-}2$ with $\gcd(k, p\text{-}1) = 1$.
2. Compute $r' \equiv \alpha^k (\text{mod} p)$, $k^{-1} (\text{mod} p\text{-}1)$, $h(m) = \tilde{m} = m^3 (\text{mod} p)$ *and* $s \equiv k^{-1}(\tilde{m} - ar') (\text{mod} p\text{-}1)$.
3. Take $z = r' + si = R(\tilde{m})$ *and* compute $\varsigma \equiv z^d (\text{mod} \eta)$.
4. B*'s* signature for m *is* ς.

Algorithm 4.3 Combined RSA-ElGamal verification by entity A.

1. Obtain B's authentic public key (p, α, y, η, e).
2. Compute $z \equiv \varsigma^e (\text{mod} \eta)$.
3. Verify that $z \in M_R$, if not, reject the signature.
4. Recover $\tilde{m} = R^{-1}(z)$.
5. Verify that $1 \leq \text{Re}(z) \leq p\text{-}1$, if not, reject the signature.
6. Compute $v_1 \equiv y^{\text{Re}(z)}. \text{Re}(z)^{\text{Im}(z)} (\text{mod} p)$ *and* $v_2 \equiv \alpha^{h(m)} (\text{mod} p)$.
7. Accept signature *if* $v_1 = v_2$.

Theorem 4.1 The signature verification method works.

Proof 4.1 Let $\varsigma \equiv z^d (\text{mod } \eta)$ such that $z = r' + si$. Since $ed \equiv 1 (\text{mod } \phi(\eta))$, we have had $\varsigma^e \equiv z^{ed} \equiv z (\text{mod} \eta)$. Then, $R^{-1}(z) = R^{-1}(R(\tilde{m})) = \tilde{m} = h(m)$. Hence, $s \equiv k^{-1}(h(m) - ar') (\text{mod} p - 1)$. Multiply both sides by k, $ks \equiv h(m) - ar' (\text{mod } p - 1)$. Then, $h(m) \equiv ks + ar' (\text{mod } p - 1)$. Hence, $\alpha^{h(m)} \equiv \alpha^{ar'+ks} \equiv (\alpha^a)^{r'}.r^{s'} \equiv y^{r'}.r^{s'}$. Therefore, $v_1 = v_2$.

Theorem 4.2 The redundancy function $R(\tilde{m}) = r' + si = \alpha^k + i\ [k - 1(\tilde{m} - ar') (\text{mod}(p - 1))]$ is a $1-1$ mapping from M *to* M_S.

Proof 4.2 Suppose that $R(\tilde{m}_1) = R(\tilde{m}_2)$ such that $R(\tilde{m}_1) = \alpha^k + i[k^{-1}(\tilde{m}_1 - ar') (\text{mod}(p - 1))] \in G_\eta$ *and* $R(\tilde{m}_2) = \alpha^k + i\ [k^{-1}(\tilde{m}_1 - ar') (\text{mod}(p - 1))] \in G_\eta$. *Then,*

$$\alpha^k + i\ [k^{-1}(\tilde{m}_1 - ar') (\text{mod}(p - 1))] = \alpha^k + i[k^{-1}(\tilde{m}_2 - ar') (mod(p - 1))]. \quad (6)$$

Thus,

$$i[k^{-1}(\tilde{m}_1 - ar') (mod(p - 1))] = i[k^{-1}(\tilde{m}_2 - ar') (mod(p - 1))]. \quad (7)$$

Multiplying both sides by (ik), we get $\tilde{m}_1 - ar' (mod(p - 1)) = \tilde{m}_2 - ar' (mod(p - 1))$, which implies that $\tilde{m}_1 = \tilde{m}_2$.

Theorem 4.3 The redundancy function $R(\tilde{m}) = r' + si = \alpha^k + i[k^{-1}(\tilde{m} - ar') (\text{mod}(p\text{-}1))]$ is not multiplicative.

Proof 4.3 It is clear that

$$R(\tilde{m}_1).R(\tilde{m}_2) = (\alpha^k + i\ [k^{-1}(\tilde{m}_1 - ar') (\text{mod}(p - 1))]).(\alpha^k + i[k^{-1}(\tilde{m}_2 - ar') (mod(p - 1))] =$$

$$[\alpha^{2k} + i\alpha^k k^{-1}(\tilde{m}_1 + \tilde{m}_2 - 2ar') - k^{-2}(\tilde{m}_1 - ar')(\tilde{m}_2 - ar')]\ (mod(p - 1)) \quad (8)$$

But, $R(\tilde{m}_1\tilde{m}_2) = \alpha^k + i[k^{-1}(\tilde{m}_1\tilde{m}_2) - ar') (mod(p - 1))]$. Therefore, R is not multiplicative.

Example 4.1 (Combined RSA-ElGamal Signature) Entity B generates the keys as follows: If p=61 and α=33 is a generator of Z_{61}^*. Then, entity B chooses the private key a=58 and computes $y \equiv 33^{58} (\text{mod } 61) \equiv 27$. After that, entity B selects q=9871 and r=5107 and computes both η=50411197 and $\phi(\eta)$=2541288659454720. Entity B chooses e=1844480063626867 and solves $ed \equiv 1844480063626867$. $d \equiv 1 (\text{mod } 50411197)$, yielding d=993514318001083. Hence, the public-key is: (p=61, α=33, α^a =27, η=50411197, e=1844480063626867) *and* the private key is (a = 58, q = 9871, r = 5107, d = 993514318001083). Assume that the hash function is $h(m) = \tilde{m} = m^3$. To sign a message m=42, entity B selects a random integer k=7 and computes $r' \equiv 33^7 (\text{mod } 61) \equiv 38$, $k^{-1} \text{ mod}(p\text{-}1) \equiv 43$ *and* $\tilde{m}$=74088. Finally, entity B computes $s \equiv 34(71884) (\text{mod } 60) \equiv 34$ and assumes z=38+34i to compute $\varsigma \equiv 23812157\text{-}23285899i$. As a result, the signature for m *is* ς. Now, to verify the signature, entity A first computes z=38+34i, then computes $\tilde{m} = R^{-1}(z) = 74088 \in M_R$. After that, entity A computes $v_1 \equiv 52$, $h(m)$=74088 *and* $v_2 \equiv 52$. Entity A accepts the signature since $v_1 = v_2$.

5. COMPARATIVE STUDY

In this section, we undertake a comparative analysis to position our novel cryptosystem against existing methodologies.

5.1 Security Evaluation and Comparative Analysis

In this study, we evaluate the security of our novel cryptographic scheme through assessments of attack, encryption and decryption times, supported by numerical simulations to measure its efficacy. Experimental investigations were conducted using an ALIENWARE laptop, specifically the Alienware 15 R4 model, equipped with an Intel(R) Core(TM) i7-8750H CPU, 16384MB RAM and BIOS version 1.20.0 (UEFI type). The laptop's robust specifications, including compatibility with Windows 11 Pro 64-bit, DirectX 12 and UEFI BIOS, along with features like Miracast Support and Microsoft Graphics Hybrid Compatibility, make it well-suited for computationally intensive experiments. Following this experimental setup, we perform a comparative analysis involving traditional RSA and ElGamal schemes alongside our novel hybrid approach, followed by a discussion of the identified strengths and weaknesses of our proposed cryptosystem.

5.2 Data Collection and Cryptanalysis

In this study, we employed Mathematica 10 to implement the algorithms for key generation, encryption, decryption and cryptanalysis of our new scheme. The prime numbers p, q and r were randomly selected from nineteen distinct intervals. These intervals, numbered from 1 to 19, encompass the ranges 10^1 to 10^{38} for both q and r and 10^1 to 10^{11} for p.

Due to the computational limitations of our current hardware, we were unable to explore higher exponents beyond 10^{38}. Our personal computer, despite its capabilities, was unable to efficiently handle the larger key sizes required for more advanced cryptanalysis. In future work, we plan to leverage high-performance computing resources or cloud platforms to extend our analysis to larger primes, which will allow for a more thorough evaluation of the scheme's performance and security with larger key sizes.

5.2.1 Key Generation

It has been observed that the total time needed for key generation in the new scheme is roughly equal to the sum of the individual times required for generating RSA and ElGamal keys.

5.2.2 Encryption and Decryption

Regarding the encryption processes, the time slots required for all three cryptosystems are approximately the same, which is very significant, because our proposed cryptosystem does not require excessive durations to be done compared to the classical ones. The same applies to the decryption process.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

5.2.3 Hacking Time Results

Hacking time denotes the period taken by an unauthorized entity to successfully decrypt or compromise the security of the combined RSA-ElGamal public-key cryptosystem under consideration. In our manuscript, this temporal measure is quantified in seconds and serves as a crucial metric for evaluating the system's resistance to potential breaches. The subsequent table (Table 1) provides comprehensive results for Hacking Time (HT) in seconds across each cryptosystem, delineating the time necessary to solve either the integer-factorization problem or the discrete-logarithm problem.

The first column of the table represents the data sizes, which are expressed as the i^{th} power of 10, ranging from 10^1 to 10^{19}. These values correspond to bit lengths ranging from approximately 4 bits (for 10^1) to approximately 63 bits (for 10^{19}). The y-axis signifies the hacking time measured in seconds.

The table provides insights into the key-generation time for RSA, ElGamal and the proposed combined RSA-ElGamal scheme, elucidating the temporal dynamics of each cryptographic system across diverse data sizes. This analysis highlights that the key-generation time of the combined RSA-ElGamal scheme is the sum of the times required to generate keys for both RSA and ElGamal. This detailed examination underscores the performance attributes of the proposed cryptographic methodology and its implications for practical applications.

In addition, the figures presented below visually portray the obtained results, providing a graphical representation of the hacking time (HT) needed to initiate an attack on each cryptosystem, measured in seconds. This hacking time pertains to the duration taken to resolve either the integer-factorization problem or the discrete-logarithm problem.

Table 1. Time required (in seconds) to compromise RSA, ElGamal and the Combined RSA-ElGamal scheme through hacking attempts. This figure illustrates the comparative performance of each encryption scheme based on its respective vulnerability to attacks.

Data Size	RSA	ElGamal	Combined RSA-ElGamal Scheme
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0.015	0.016
7	0	0.015	0.016
8	0.031	0.109	0.125
9	0.047	0.015	0.062
10	0.14	0.032	0.172
11	0.297	0.86	1.125
12	2.797	2.36	5.157
13	5.359	10.5	16.25
14	11.016	6.89	18.047
15	60.922	117.016	168.5
16	307.61	213.563	519.218
17	661.328	8826.2	9737.52
18	2635.02	23808.9	26482.3
19	10993.7	71228.1	82294

5.2.4 Observations

The analysis of Figures 1, 2 and 3 reveals an interesting observation regarding the impact of data size on the time required to attack the combined RSA-ElGamal scheme in comparison to the classical RSA and ElGamal schemes. Initially, when the data size is relatively small, there is no noticeable difference between the three schemes. However, as the data sizes increase, significant differences arise, with the time required to attack the combined RSA-ElGamal scheme surpassing that of the RSA and ElGamal schemes by several thousands of seconds. Furthermore, Figure 4 provides

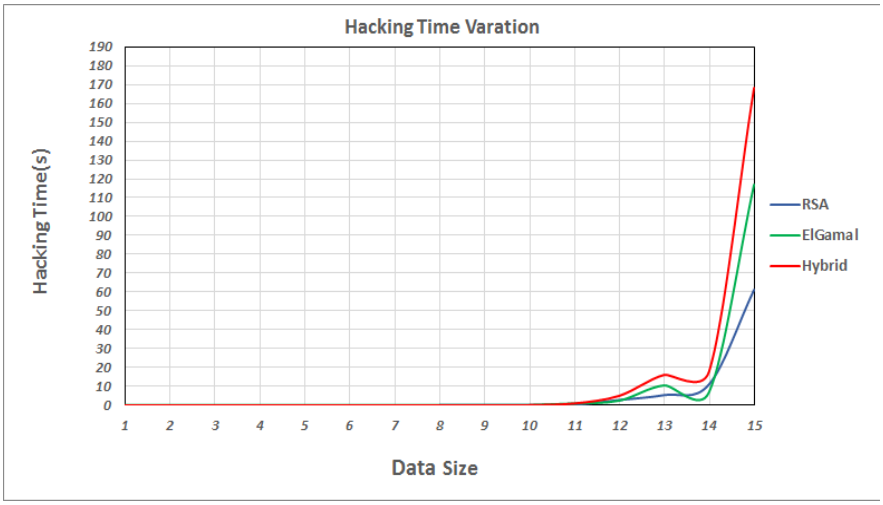


Figure 1. Comparison of performance and efficiency between the classical RSA and ElGamal cryptosystems and the innovative combined RSA-ElGamal cryptosystem, with data sizes ranging from 10^1 to 10^{15} (approximately 4 to 50 bits). This figure illustrates how each system performs across different data sizes.

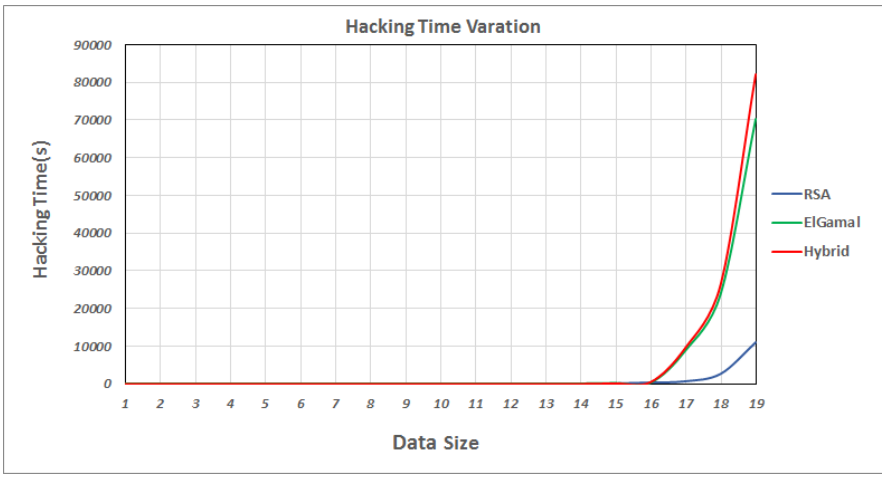


Figure 2. Comparison of performance and efficiency of the classical RSA and ElGamal cryptosystems *versus* the innovative combined RSA-ElGamal cryptosystem across data sizes ranging from 10^1 to 10^{19} (approximately 4 to 67 bits).

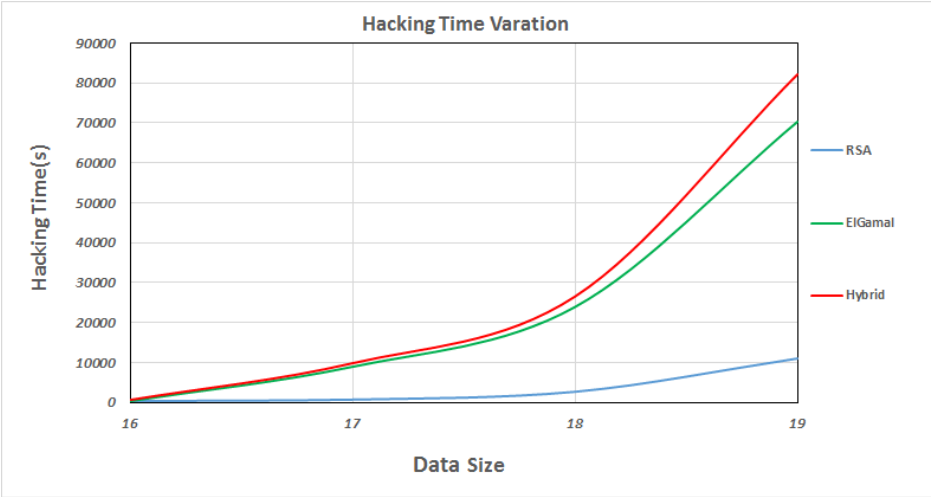


Figure 3. Comparison of performance and efficiency of the classical RSA and ElGamal cryptosystems with the innovative combined RSA-ElGamal cryptosystem across data sizes ranging from 10^{16} to 10^{19} (approximately 54 to 64 bits).

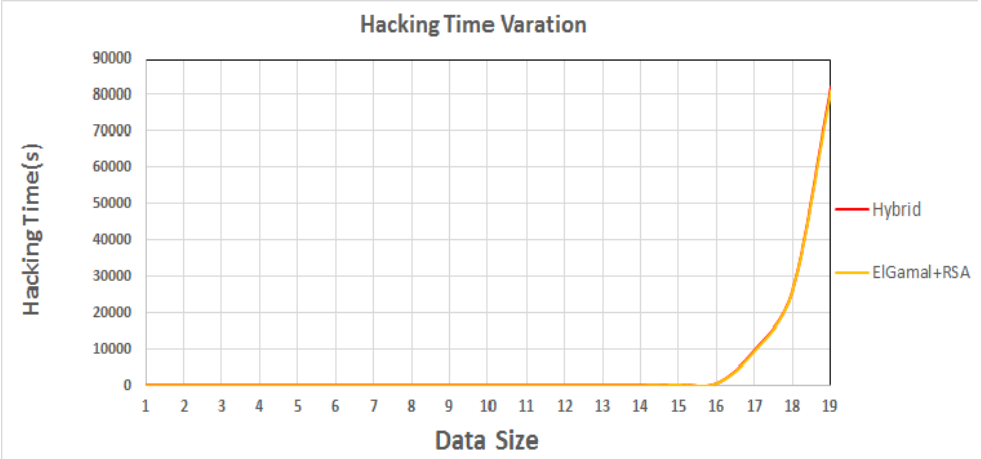


Figure 4. Analysis of the combined RSA-ElGamal algorithm's hacking time (HT) compared to the combined hacking time of the classical RSA and ElGamal algorithms. This figure illustrates the performance differences between the combined RSA-ElGamal approach and the sum of individual classical algorithms' hacking times.

additional insight, demonstrating that the attacking time of the new scheme is notably greater and equivalent to the cumulative attacking times of the classical RSA and ElGamal cryptosystems combined. Based on this evidence, we can deduce that the combined RSA-ElGamal scheme offers enhanced security compared to both the classical RSA and ElGamal schemes. In summary, the analysis showcases the advantage of the combined RSA-ElGamal scheme, highlighting its resistance to attacks as the data size grows larger. The substantial increase in attacking time for the combined RSA-ElGamal scheme, compared to the classical RSA and ElGamal schemes, suggests its heightened level of security and reinforces its suitability for cryptographic applications.

5.3 Complexity Analysis

In the following sub-sections, we provide a comparative analysis of the complexity of the RSA, ElGamal and the proposed combined RSA-ElGamal scheme algorithms. Time complexity of an algorithm is commonly expressed using the asymptotic notation of O(n), which is determined by counting the number of basic operations performed during the algorithm's execution, such as addition, subtraction, multiplication and division. The space complexity of a cryptographic algorithm refers to the amount of memory required for the algorithm to run, relative to the length of its input. Space complexity depends on the size of the input. When considering the maximum complexity for a given input size, it is referred to as worst-case complexity. Conversely, when considering the average complexity across all inputs of a given size, it is known as the expected complexity.

5.3.1 Complexity of the RSA Scheme

Complexity of key generation: To generate the key, the complexity of selecting random primes p and q and computing their product $n = pq$ is either $O(log_2^2 p)$ using the Fermat's primality test or $O(log_2^3 p)$ using the Miller-Rabin test. Computing $n=pq$ in the domain of natural integers, Z, has a complexity of $O(\log_2 p \log_2 q) \approx O(log_2^2 p)$ since $p < q$. Computing Euler's totient function $\phi(n) = (p-1)(q-1)$ has a complexity of $O(\log_2 p \log_2 q)) \approx O(log_2^2 p)$, since $p < q$. The complexity of selecting a random number e such that $0<e<\phi(n)$ with $(e, \phi(n)) = 1$ using Euclidean division is $O(log_2^3 \phi(n)) \approx O(log_2^3 pq) \approx O(log_2^3 n)$. Thus, the overall time complexity of the key-generation process is $O(log_2^3 n)$.

Complexity of the encryption process: The complexity of computing $c=m^e \pmod n$ is $O(log_2^3 n)$, since the size of e is proportional to that of n.

Complexity of the encryption process: The complexity of computing $m=c^d \pmod n$ is $O(log_2^3 n)$, since the size of e is proportional to that of n.

5.3.2 Complexity of ElGamal Scheme

Complexity of key generation: Selecting a random prime number p has a complexity of either $O(log_2^2 p)$ if the Fermat's primality test is used or $O(log_2^3 p)$ if the Miller-Rabin test is used. Selecting a value k between 2 and $p-2$ has a complexity of $O(\log_2 p)$. Finding a generator α of the multiplicative group Z_p^*has a complexity of $O(log_2^2 p)$. Computing $\alpha^a (\text{mod } p)$ has a complexity of $O(log_2^3 p)$. The overall complexity of key generation is determined to be $O(log_2^3 p)$.
Complexity of the encryption process: Computing $\gamma \equiv \alpha^k (mod\, p)$ has a complexity of $O(log_2^3 p)$. Computing $d \equiv (\alpha^k m)(mod\, p)$ has a complexity of $O(log_2^2 p)$. The overall complexity of the encryption process is $O(log_2^3 p)$.
Complexity of the decryption process: Computing $\gamma^{p-a-1} \equiv \alpha^{-ak} (mod\, p)$ using the extended Euclidean algorithm has a complexity of $O(log_2^3 p)$. Computing $m \equiv \alpha^{-ak}\delta (mod\, p)$ has a complexity of $O(log_2^2 p)$. The overall complexity of the decryption process is $O(log_2^3 p)$.

5.3.3 Complexity of the Combined RSA-ElGamal Scheme

Key-generation Complexity: Picking random primes p, q and r has a complexity of $O(log_2^2 p)$ if the Fermat's primality test is used or $O(log_2^3 p)$ if the Miller-Rabin test is used. Selecting a value k such that $2 \leq k \leq p^2 - 1$ has a complexity of $O(2log_2 p)$. Finding a generator α of the group G_p^* has a complexity of $O(log_2^2 p^2)$ using a specific algorithm. Computing α^a (mod p) has a complexity of $O(log_2^3 p)$. Computing $\phi(\eta) = (q^2 - 1)(r^2 - 1)$ has a complexity of $O(log_2^4 p)$. Selecting a value e such that $1 \leq e \leq \phi(\eta)$ and $\gcd(e, \phi(\eta)) = 1$ has a complexity of $O(log_2^3 \eta)$. Computing $d \equiv e^{-1}$ (mod $\phi(\eta)$) has a complexity of $O(log_2^3 \phi(\eta)) = O(log_2^3 \eta^2)$ using the extended Euclidean algorithm. Hence, the overall complexity of key generation is $O(log_2^4 q)$.

Encryption-process Complexity: Computing $\gamma \equiv \alpha^k (\text{mod } p)$ has a complexity of $O(log_2^3 p)$. Computing $\delta \equiv (\alpha^k m) (\text{mod } p)$ has a complexity of $O(log_2^3 p)$. Computing $c \equiv (\gamma + \delta i)^e (\text{mod } \eta)$ has a complexity of $O(log_2^3 \eta)$. Hence, the overall complexity of the message-encryption process is $O(log_2^3 p)$.

Decryption-process Complexity: Computing $M \equiv c^d (\text{mod } \eta)$ has a complexity of $O(log_2^3 \eta)$. Computing $f \equiv Re(M)^{p^2 - a - 1} (\text{mod } p)$ has a complexity of $O(log_2^2 p)$. Computing $h \equiv \text{Im}(M) (\text{mod } p)$ has a complexity of $O(\log_2 p)$. Computing $t \equiv fh (\text{mod } \eta)$ has a complexity of $O(log_2^2 \eta)\eta)$. Hence, the overall complexity of the message-decryption process is $O(log_2^3 \eta)$.

In summary, our analysis has thoroughly examined the computational complexities inherent in three prominent cryptographic schemes: RSA, ElGamal and the combined RSA-ElGamal schemes. We assessed these complexities in terms of key generation, encryption and decryption operations. For RSA, the key generation, encryption and decryption exhibit a time complexity of approximately $O(log_2^3 n)$, leveraging efficient key generation, but demanding multiple modular exponentiations for encryption and decryption. Conversely, ElGamal scheme demonstrates a similar time complexity of $O(log_2^3 p)$ for key generation, encryption and decryption, excelling in key generation and encryption processes while requiring an additional modular exponentiation during decryption. The combined RSA-ElGamal scheme combines the strengths of both RSA and ElGamal schemes, featuring key generation complexity of $O(log_2^4 p)$and encryption complexity of $O(log_2^3 p)$, akin to ElGamal scheme. However, decryption complexity increases to $O(log_2^3 \eta)$, reflecting a slight trade-off for the inclusion of RSA's capabilities. In conclusion, the combined RSA-ElGamal scheme offers a balanced approach, leveraging RSA's advantages in key management alongside ElGamal's encryption efficiency, with the choice of scheme dependent on specific security needs and computational considerations.

5.4 Real-world Applicability

5.4.1 Impact of Key Generation on Overall Performance

The combined RSA-ElGamal scheme presents a balance between enhanced security and computational efficiency. As discussed in our analysis, the key-generation process for this novel scheme is inherently more complex and time-intensive compared to the individual RSA or ElGamal

schemes. This complexity arises from the need to generate and manage three distinct prime numbers and perform additional arithmetic operations within the domain of Gaussian integers.

In practical applications, the extended key-generation time can affect the performance of systems that rely on frequent key rotations or require rapid key creation. For instance, in scenarios such as secure communications or real-time applications where keys are generated dynamically, the increased key-generation time might impact the system responsiveness. To address this issue, future research could explore optimizing key-generation processes or leveraging parallel-computing techniques to reduce the required time.

5.4.2 Integration into Existing Security Frameworks

The combined RSA-ElGamal scheme can be integrated into existing security frameworks with minimal disruption. Its dual-layer approach enhances security while requiring only minor adjustments to the existing cryptographic infrastructure. Key points include:

1. **Backward Compatibility:** The combined RSA-ElGamal scheme can be deployed alongside existing RSA or ElGamal systems, facilitating gradual adoption. This allows organizations to apply the new approach to new applications or incrementally transition from older systems.
2. **Modular Integration:** The architecture of the combined RSA-ElGamal scheme supports modular integration into existing security frameworks. It can be incorporated into established protocols, such as TLS or VPNs, either as a replacement or a complement to existing encryption algorithms, thereby enhancing security without necessitating an overhaul of the entire system.
3. **Adaptability for Specific Use Cases:** The flexibility in the combined RSA-ElGamal scheme's parameter choices enables customization for specific security needs. For example, in environments with stringent security requirements, the scheme's enhanced resistance to attacks can be particularly advantageous.
4. **Compatibility with Modern Hardware:** Given that the computational demands of the new scheme are manageable with current hardware, it can be effectively utilized in both software-based and hardware-accelerated cryptographic systems.

In summary, while the combined RSA-ElGamal scheme introduces additional computational overhead, it offers significant security benefits that can be applied to various real-world scenarios. By carefully considering the impact of key generation and thoughtfully integrating the scheme into existing frameworks, its advantages can be maximized and potential performance challenges can be mitigated.

5.5 Practical Limitations

5.5.1 Increased Computation Time

While the theoretical analysis of computational complexities provides a foundation, practical implementations often encounter additional challenges. The combined RSA-ElGamal scheme, due to its combined use of RSA and ElGamal schemes, involves complex operations that can impact performance:

1. **Key Generation:** The key-generation process for the combined RSA-ElGamal scheme requires generating three primes and performing additional arithmetic operations within the domain of Gaussian integers. This complexity can lead to significantly longer key generation times compared to RSA and ElGamal schemes individually. This extended time might affect systems that require frequent key updates or rapid key generation, such as secure-communication systems and real-time applications.
2. **Encryption and Decryption:** Although the encryption process for the combined RSA-ElGamal scheme shows comparable time requirements to traditional methods, the combined computational steps from both RSA and ElGamal schemes can lead to longer processing times in practical scenarios. For instance, the combined RSA-ElGamal scheme involves modular exponentiations and additional arithmetic operations that could contribute to a slower overall encryption and decryption process.

5.5.2 Memory Requirements

The combined RSA-ElGamal scheme's increased complexity also impacts memory usage:

1. **Storage for Intermediate Results:** The computations involved in key generation and encrypttion/decryption require storing intermediate results, which can increase memory usage. For instance, handling large integers and matrices in Gaussian-integer arithmetics can lead to higher memory demands compared to simpler encryption schemes.
2. **Ciphertext Size:** As discussed, the new scheme may result in ciphertexts that are larger due to the inclusion of both real and imaginary components. This increase in ciphertext size can impact storage requirements and bandwidth, particularly in systems with limited resources.

5.5.3 Mitigation Strategies

To address these practical limitations, several strategies can be considered:

1. **Algorithm Optimization:** Future research could focus on optimizing the combined RSA-ElGamal scheme's algorithm to reduce computation time and memory usage.
2. **Hardware Acceleration:** Implementing hardware acceleration for cryptographic operations could help manage the increased computational load and memory requirements, making the scheme more feasible for practical applications.
3. **Efficient Storage Solutions:** Exploring efficient storage and management solutions for intermediate results and ciphertexts could help mitigate memory overhead.

5.5.4 Advantages and Disadvantages

Advantages: The combined RSA-ElGamal scheme offers several significant advantages:

1. **Integration of RSA and ElGamal Schemes:** By combining RSA and ElGamal encryption schemes, the new approach delivers a dual-layered security solution. The first layer leverages the discrete-logarithm problem, while the second layer relies on the integer-factorization problem. This combination creates a robust encryption framework similar to a double onion routing shield.
2. **Smooth Implementation:** Implementing the combined RSA-ElGamal scheme requires no additional effort beyond what is expected with traditional encryption schemes. The transition to the new approach can be smoothly achieved without introducing added complexities or burdensome requirements.
3. **Expanded Parameter Range:** With computational efforts comparable to traditional encryption methods, the combined RSA-ElGamal scheme allows for a broader selection of plaintexts and private keys. In fact, the number of available options exceeds the square of those in classical encryption settings, offering increased flexibility for customized-encryption processes.
4. **Efficient Encryption and Decryption:** The combined RSA-ElGamal scheme maintains comparable time requirements for encryption and decryption processes relative to traditional encryption methods. No extra time is required for these operations, ensuring that the new approach remains efficient and practical.
5. **Enhanced Security:** The combined RSA-ElGamal scheme significantly increases resistance to attacks compared to traditional encryption methods. The time required for an attacker to compromise the combined RSA-ElGamal scheme is substantially greater than or equal to the combined time needed to break both underlying classical encryption schemes.

Disadvantages: Despite its strengths, the combined RSA-ElGamal scheme has a few drawbacks that must be considered:

1. **Increased Ciphertext Length:** In certain cases, the ciphertext generated by the combined RSA-ElGamal scheme may be twice the length of the original message. This occurs when the plaintext is real, resulting in a complex-number ciphertext that includes both real and imaginary components. The increased ciphertext length may impact storage requirements or communication bandwidth, which is an important consideration in resource-constrained

environments.

2. **Computational Overhead:** The combined RSA-ElGamal scheme involves extensive computations, particularly when dealing with large logarithmic calculations. This computational burden can result in longer processing times, potentially affecting the feasibility of encryption processes, especially when handling large datasets. The algorithm may perform sluggishly when encrypting extensive data on a single machine, necessitating optimization or the use of distributed-computing strategies.
3. **Slower Key Generation:** Key generation in the combined RSA-ElGamal scheme is slower compared to RSA and ElGamal schemes. This is because the combined RSA-ElGamal scheme requires the generation of three primes (as opposed to one in ElGamal or two in RSA) and additional computations. Moreover, arithmetic operations within the domain of Gaussian integers add further computational requirements, depending on the specific forms of the selected Gaussian primes. As a result, key generation in the combined RSA-ElGamal scheme demands more time and more computational resources.

Therefore, the proposed combined RSA-ElGamal encryption scheme combines the strengths of RSA and ElGamal schemes while avoiding excessive implementation complexities. However, careful consideration of the potential expansion of ciphertext length, computational overhead and slower key-generation processes is essential. These factors should be evaluated to determine the suitability of the combined RSA-ElGamal scheme for various cryptographic applications, taking into account the specific requirements and constraints of the intended use cases.

6. CONCLUSION

In summary, our investigation into the combined RSA-ElGamal scheme highlights its effectiveness in striking a balance between enhanced security and computational efficiency. Despite the inherent complexity in key generation compared to standalone RSA or ElGamal schemes, the new approach delivers significant security benefits that warrant this added complexity. When compared to other advanced cryptographic methods, such as elliptic curve cryptography (ECC) and lattice-based cryptography, the combined RSA- ElGamal scheme presents a distinctive combination of security advantages and practical utility.

Our study emphasizes the theoretical robustness of this new approach. Future work will focus on several key areas to further enhance and validate the scheme. We plan to provide a comprehensive analysis of the cryptanalysis methods used, including specific algorithms and their implementations. We will also include comparative studies with recent cryptographic techniques to contextualize our results and demonstrate the scheme's relative efficacy. Additionally, optimizing key generation processes, evaluating performance in various real-world scenarios and integrating the combined RSA-ElGamal scheme with existing security frameworks will be pivotal. Exploring variants and extensions of the combined RSA-ElGamal scheme will offer deeper insights into its practical advantages and limitations, guiding its future development and application.

ACKNOWLEDGEMENTS

We extend our profound gratitude to the editor and the referees for their meticulous review and invaluable feedback. Their expert insights and constructive comments have been instrumental in refining the manuscript and enhancing its quality. We deeply appreciate the time and effort dedicated by them to review our work, which has significantly enriched the depth and impact of our research. Thank you for your exceptional contributions and support.

REFERENCES

[1] T. Adamski and W. Nowakowski, "The Average Time Complexity of Probabilistic Algorithms for Finding Generators in Finite Cyclic Groups," Bulletin of the Polish Academy of Sciences: Technical Sciences, vol. 63, no. 4, pp. 989-996, 2015.

[2] E. A. Adeniyi et al., "Secure Sensitive Data Sharing Using RSA and ElGamal Cryptographic Algorithms with Hash Functions," Information, vol. 13, no. 10, p. 442, 2022.

[3] J. M. Ahmed and Z. M. Ali, "The Enhancement of Computation Technique by Combining RSA and ElGamal Cryptosystems," Proc. of the 2011 Int. Conf. on Electrical Engineering and Informatics, pp. 1-5, DOI: 10.1109/ICEEI.2011.6021779, Bandung, Indonesia, 2011.

[4] A. Aman and R. K. Aggarwal, "A Survey: Analysis of Existing Hybrid Cryptographic Techniques," Proc. of the Int. Conf. on Recent Developments in Cyber Security, Cyber Security and Digital Forensics (REDCYSEC 2023), Part of the Book: Lecture Notes in Networks and Systems, vol. 896, pp. 259-269, Springer, 2023.

[5] S. Anjana et al., "Security-enhanced Optical Nonlinear Cryptosystem Based on Phase-truncated Fourier Transform," Optical and Quantum Electronics, vol. 55, no. 12, p. 1099, 2023.

[6] Y. Awad et al., "Comparative Study between a Novel Deterministic Test for Mersenne Primes and Well-known Primality Tests," Baghdad Science Journal, vol. 20, no. 5 (Suppl.), 2023.

[7] Y. Awad, A. N. El-Kassar and T. Kadri, "Rabin Public-key Cryptosystem in the Domain of Gaussian Integers," Proc. of the 2018 Int. Conf. on Computer and Applications (ICCA), pp. 1-340, DOI: 10.1109/COMAPP.2018.8460338, Beirut, Lebanon, 2018.

[8] M. Bunder, A. Nitaj, W. Susilo and J. Tonien, "A Generalized Attack on RSA Type Cryptosystems," Theoretical Computer Science, vol. 704, pp. 74-81, 2017.

[9] J. J. Cogswell, The Theory of Indices Modulo n, Ph.D. Dissertation, Emporia State Univ., Emporia, KS, 2012.

[10] J. T. Cross, "The Euler ϕ-function in the Gaussian Integers," American Mathematical Monthly, vol. 90, no. 8, pp. 518-528, 1983.

[11] A. A. El-Douh, S. F. Lu, A. Elkony and A. S. Amein, "A Systematic Literature Review: The Taxonomy of Hybrid Cryptography Models," Proc. of Future of Information and Communication Conf., Advances in Information and Communication (FICC 2022), Part of the Book Series: Lecture Notes in Networks and Systems, vol. 439, pp. 714-721, Springer International Publishing, 2022.

[12] T. ElGamal, "A Public-key Cryptosystem and a Signature Scheme Based on Discrete Logarithms," IEEE Transactions on Information Theory, vol. 31, no. 4, pp. 469-472, 1985.

[13] A. N. El-Kassar, R. A. Haraty, Y. A. Awad and N. C. Debnath, "Modified RSA in the Domains of Gaussian Integers and Polynomials over Finite Fields," Proc. ISCA 18th Int. Conf. Comput. Appl. Ind. Eng. (CAINE), pp. 298-303, Honolulu, USA, 2005.

[14] A. N. El-Kassar, R. Haraty and Y. Awad, "Rabin Public-key Cryptosystem in Rings of Polynomials over Finite Fields," Proc. of the Int. Conf. on Computer Science, Software Engineering, Information Technology, e-Business and Applications (CSITeA'04), Cairo, Egypt, 2004.

[15] A. N. El-Kassar and S. Habre, "Greatest Common Divisor and Least Common Multiple Matrices on Factor Closed Sets in a Principal Ideal Domain," J. of Mathematics and Statistics, vol. 5, no. 4, pp. 342-347, 2009.

[16] A. A. Emmanuel, A. E. Okeyinka, M. O. Adebiyi and E. O. Asani, "A Note on Time and Space Complexity of RSA and ElGamal Cryptographic Algorithms," Int. J. of Advanced Computer Science and Applications, vol. 12, no. 7, pp. 143-147, 2021.

[17] S. M. Hardi, J. T. Tarigan and N. Safrina, "Hybrid Cryptosystem for Image File Using ElGamal and Double Playfair Cipher Algorithm," Journal of Physics: Conference Series, IOP Publishing, vol. 978, no. 1, p. 012068, 2018.

[18] J. Hoffstein, "Integer Factorization and RSA," Proc. of An Introduction to Mathematical Cryptography, Part of the Book: Undergraduate Texts in Mathematics (UTM), pp. 1-75, New York, USA, 2008.

[19] K. S. Gaur et al., "Cryptanalysis of the Optical Cryptosystem Titled 'An Asymmetric Image Encryption Based on Phase Truncated Hybrid Transform'," Journal of Optics, vol. 53, pp. 605-609, 2023.

[20] C. F. Gauss, "The Arithmetic of the Gaussian Integers," [Online], Available: https://personal.math.ubc.ca/~anstee/math444/GaussianIntegersfinal.pdf, 2020.

[21] R. Imam et al., "Systematic and Critical Review of RSA Based Public- key Cryptographic Schemes: Past and Present Status," IEEE Access, vol. 9, pp. 155949-155976, 2021.

[22] N. M. S. Iswari, "Key Generation Algorithm Design Combination of RSA and ElGamal Algorithm," Proc. of the 2016 8th IEEE Int. Conf. on Information Technology and Electrical Engineering (ICITEE), pp. 1-5, Yogyakarta, Indonesia, 2016.

[23] M. Iavich et al., "Hybrid Encryption Model of AES and ElGamal Cryptosystems for Flight Control Systems," Proc. of the 5th IEEE Int. Conf. on Methods and Systems of Navigation and Motion Control (MSNMC), pp. 229-233, Kiev, Ukraine, 2018.

[24] P. Kuppuswamy and S. Q. Al-Khalidi, "Hybrid Encryption/Decryption Technique Using New Public-key and Symmetric Key Algorithm," Int. J. of Information and Computer Security, vol. 6, no. 4, pp. 372-382, 2014.

[25] A. J. Menezes et al., Handbook of Applied Cryptography, CRC Press, 2018.

[26] M. Mohammadi, A. Zolghadr and M. Pourmina, "Comparison of Two Public-key Cryptosystems," Journal of Optoelectronical Nanostructures, vol. 3, no. 3, pp. 47-58, 2018.

[27] B. Molelekeng, Arithmetic in the Ring of Gaussian Integers, Ph.D. Dissertation, University of the Witwatersrand, Johannesburg, 2022.

[28] P. K. Panda and S. Chattopadhyay, "A Hybrid Security Algorithm for RSA Cryptosystem," Proc. of the

2017 4th IEEE Int. Conf. on Advanced Computing and Communication Systems (ICACCS), pp. 1-6, Coimbatore, India, 2017.

[29] J. M. Parenreng and A. Wahid, "The E-mail Security System Using El-Gamal Hybrid Algorithm and AES (Advanced Encryption Standard) Algorithm," Internet of Things and Artificial Intelligence J., vol. 2, no. 1, pp. 1-9, 2022.

[30] K. S. Patil, I. Mandal and C. Rangaswamy, "Hybrid and Adaptive Cryptographic-based Secure Authentication Approach in IoT Based Applications Using Hybrid Encryption," Pervasive and Mobile Computing, vol. 82, p. 101552, DOI: 10.1016/j.pmcj.2022.101552, 2022.

[31] I. S. Permana, T. Hidayat and R. Mahardiko, "Raw Data Security by Using ElGamal and SHA 256 Public-key Algorithm," Teknokom, vol. 4, no. 1, pp. 1-6, 2021.

[32] R. L. Rivest, A. Shamir and L. Adleman, "A Method for Obtaining Digital Signatures and Public-key Cryptosystems," Communications of the ACM, vol. 21, no. 2, pp. 120-126, 1978.

[33] L. B. Rivera et al., "Hybrid Cryptosystem Using RSA, DSA, ElGamal and AES," Int. Journal of Scientific & Technology Research, vol. 8, no. 10, pp. 1777-1781, 2019.

[34] A. P. U. Siahaan, B. O. Elviwani and B. Oktaviana, "Comparative Analysis of RSA and ElGamal Cryptographic Public-key Algorithms," Proc. of the Joint Workshop KO2PI and the 1st Int. Conf. on Advance & Scientific Innovation (ICASI'18), pp. 163-172, DOI: 10.4108/eai.23-4-2018.2277584, 2018.

[35] H. Singh, R. Girija and M. Kumar, "A Cryptoanalysis of Elliptic Curve Cryptography Based on Phase Truncation in the Domain of Hybrid Gyrator Hartley Transform," Optical and Quantum Electronics, vol. 55, no. 6, p. 487, 2023.

[36] N. Tahat et al., "A New RSA Public-key Encryption Scheme with Chaotic Maps," Int. Journal of Electrical and Computer Engineering (IJECE), vol. 10, no. 2, pp. 1430-1437, 2020.

[37] Q. Zhang, "An Overview and Analysis of Hybrid Encryption: The Combination of Symmetric Encryption and Asymmetric Encryption (CDS)," Proc. of the 2nd IEEE Int. Conf. on Computing and Data Science (CDS), pp. 616-622, DOI: 10.1109/CDS52072.2021.00111, Stanford, USA, Jan. 2021.

ملخص البحث:

تُقدّم هذه الورقة منهجيةً مبتكرةً تجمع بين خوارزمية RSA وخوارزمية الجَمَل لتحسين أمان وفاعلية أنظمة التّشفير ذات المفاتيح العامّة. ومن خلال الجمْع بين الخوارزميتين المذكورتين، فإنّ طريقتنا تجمع بين الاستفادة من نقاط القوّة والحدّ من نقاط الضّعْف في الأنظمة التقليدية، وبخاصّةٍ فيما يتعلق بمشكلة عَومَلة الأعداد الصّحيحة ومشكلة اللّوغاريثم المجرّد. ويعمل استخدام أعداد غاوْسْ الصّحيحة على تعزيز متانة عمليات التّشفير والتّوقيع الرّقمي، موفّراً إطارَ عملٍ أكثر أماناً. وتتضمّن دراستنا تحليلاً شاملاً للخوارزميتين اللّتين يتمّ الجمع بينهما، إلى جانب تطبيقاتٍ عمليةٍ وتقييمات لأنظمة التّشفير الّتي تُصمَّم بالمنهجية المبتكرة في هذه الدّراسة مع التركيز على التّحدّيات المتمثّلة في عَومَلة الأعداد الصّحيحة ومشكلة اللّوغاريثم المجرّد. كذلك تمّ إجراء تقييماتٍ للوقوف على جودة النّظام المقترح وفاعليته الحسابية.

فبينما كان توليد المفاتيح أبطأ مقارنةً باستخدام خوارزمية RSA أو خوارزمية الجَمَل على انفراد، فإنّ المنهجية المقترحة تنِمُّ عن أداءٍ جيّدٍ في التّشفير وإزالة التّشفير. وعلى العكس من الأنظمة المشابهة في دراسات أخرى الّتي تركّز على معالجة الصُّوَر الضوئية، فإنّ دراستنا تعملُ على توسيع نطاق أنظمة التّشفير والتّوقيع الرّقمي إلى تطبيقات أكثر، ساعيةً بذلك إلى تحسين الجانب النّظري المتعلّق بموضوع الدّراسة وتطوير التّطبيقات العملية لأنظمة التّشفير والتّوقيع الرّقمي. والجدير بالذّكر أنّ البحث المستقبلي حَوْل موضوع الدراسة سيركّز على توليد المفاتيح بطرقٍ أكثر مثاليةً، واستكشاف دمج النّظام المقترح في أُطُر العمل القائمة المتعلّقة بأمان الأنظمة، وتقييم الأداء في ظلّ سيناريوهاتٍ من العالم الحقيقي؛ وذلك من أجْل تحسين أداء النّظام المقترح والتّحقُّق من نجاعته.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

OPTIMIZATION OF FALSE ALARM RATE AND MISDETECTION RATE FOR A DESIRED THRESHOLD VOLTAGE IN COOPERATIVE COMMUNICATION

Satish Kumar Gannamaneni and Jibendu Sekhar Roy

(Received: 16-Jul.-2024, Revised: 24-Sep.-2024, Accepted: 17-Oct.-2024)

ABSTRACT

Cooperative communication systems (CCSs) involves collaboration among sensor nodes to transmit data more effectively, especially in scenarios with limited resources or challenging environmental conditions. Optimizing the total error rate (TER) for cooperative communication in wireless sensor networks (WSNs) is a critical task to enhance the reliability and efficiency of data transmission. The link quality of a WSN can be improved by cooperative relaying with a relatively low TER. In this paper, real-coded genetic algorithm (RGA) and particle swarm optimization (PSO) are used in WSNs to reduce TER. The number of nodes is varied from 1 to 16, SNR is varied from 0 dB to 20 dB, threshold is varied from 25 mV to 35 mV and mutation rate is 0.1. Minimum TER is obtained for a threshold of 25 mV to 35 mV compared to TER obtained without optimization. The optimization method provides significant improvements to achieve the desired threshold voltage with minimum false alarm rate and misdetection rate, which enhances the overall performance of the CCS in WSNs.

KEYWORDS

Cooperative communication, Cognitive radio, Optimization, RGA, PSO, Error rate.

1. INTRODUCTION

In CCS, the source node (SN) apart from sending the signal to the destination node (DN), also relays it to an intermediate node. Subsequently, this relay node (RN) forwards the signal either to another relay node or directly to the destination node using various relaying protocols [1].This technology utilizes mobile relay nodes to augment the capacity of specific users. This principle entails categorizing the system into three types of nodes: the SN, the DN and the RN [2]. In this setup, all relay nodes function as both receiving and sending antennas for particular user nodes. Consequently, the network can be conceptualized similarly to a multiple input and multiple output (MIMO) system [3]-[4]. The CCS involves 2 phases of transmission [5]. Phase I: The users distribute the source data with control information among the remaining users. Phase II: The users collectively re-transmit the data to the desired destination.

In a relay system, one user is the SN, while the other user is the RN and both the users can interchange their roles at various time intervals [6]. As outlined earlier, during Phase I, the SN user will broadcast information to both the RN and the DN. Subsequently, in Phase II, the RN can independently forward data or collaborate with the SN to enhance reception at the DN. Coordination is vital in a CCS, especially since antennas are distributed across various terminal devices, unlike centralized MIMO systems [7]. Excessive coordination may lead to a reduction in system bandwidth, but this cost is consistently offset by the substantial diversity gain achieved under high signal-to-noise ratio (SNR) conditions [5]. The expense of coordination may escalate with the number of cooperating users. Therefore, designing an efficient user-to-user communication method is necessary for a successful cooperation [5], [8]. In order to achieve energy-efficient transmission, a cooperative spectrum sensing (CSS) technique based on PSO is developed for cognitive WSNs [9]-[10]. The enhanced performance and flexibility of a re-configurable unmanned aerial vehicle relay-communication system is proposed in [11]. The optimization of energy efficiency in wireless networks beyond 5G through the integration of visible light and RF bands, employing non-orthogonal multiple access schemes for downlink using visible light to enhance data rates for cell-edge users *via* cooperative communications strategies is

S.K.Gannamaneni and J. S. Roy are with School of Electronics Engineering, KIIT Deemed to be University, Bhubaneswar, Odisha, India.
Emails: gskfet@kiit.ac.in and drjsroy@kiit.ac.in

explored in [12]. The challenges of energy constraints and security concerns in WSNs are addressed in [13]. A study on maximization of energy efficiency for the cooperative-communication system achieved by optimizing both the packet size and the modulation level is reported in [14]. To extend communication coverage, the use of energy harvesting-based combined information and power transfer is reported in [15]-[16]. The performance of the energy-detection mechanism can be evaluated based on the misdetection rate and the false-alarm rate [17]. The investigation of spectral monitoring over Rayliegh fading channels is presented in [18]. The impact of cooperative-spectrum sensing and dynamic-threshold selection on the performance of a cognitive radio network in fading environments, comparing the receiver operating characteristics for cooperative and non-cooperative scenarios and demonstrating improvements in detection probability, error probability with dynamic-threshold selection is discussed in [19].The spectrum sensing in cognitive radio using energy detection over various wireless communication channels is presented in [20]. A comprehensive review of optimization algorithms, including evolutionary, swarm intelligence and metaheuristic approaches, for enhancing wireless sensor network (WSN) node localization, evaluating their accuracy, scalability, computational complexity and robustness across different deployment scenarios is presented in [21]. In [22], authors presented a VLSI-based power-optimization model for wireless sensor networks using FPGA technology, designed to lower energy consumption by employing a collaborative unit with parallel processing and a smart power component, leading to improved efficiency compared to processor-based WSN implementations. In [23], authors proposed a PSO-based scheduling approach to maximize the lifetime of wireless sensor networks by addressing the non-disjoint sets cover problem, demonstrating competitive performance and optimal solutions compared to state-of-the-art algorithms. The research gap is to enhance the data rate in cooperative-communication system, by minimizing the TER for a desired threshold voltage.

2. METHODOLOGY

2.1 Aim of the Research Work

The current work aims to enhance the efficiency and reliability of cooperative-communication systems (CCSs) in wireless sensor networks (WSNs) by optimizing the total error rate (TER). It focuses on using particle-swarm optimization (PSO) and real-coded genetic algorithms (RGAs) to achieve these improvements.

2.2 System Model

A cooperative-communication network comprising K cognitive radios (CRs) with a common receiver is examined and illustrated in Figure 1. Here, spectrum sensing is performed by each CR independently, followed by transmitting local decisions to the shared receiver.

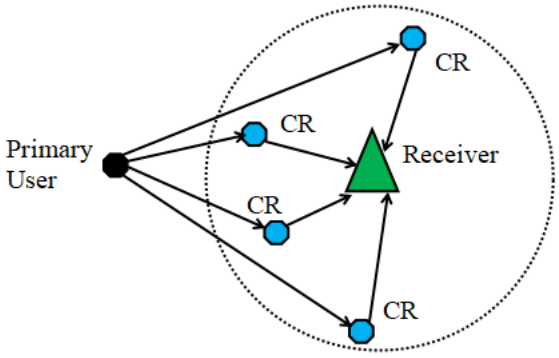


Figure 1. Spectrum-sensing model in a cooperative CR.

The receiver then consolidates all available information regarding decisions to deduce the presence or absence of the primary user (PU). Spectrum sensing essentially involves a binary hypothesis-testing problem:

H_0 : PU is not present.
H_1 : PU is present.

Considering sensing of spectrum only at ith CR, the sensing method may choose any of the following 2 hypotheses.

$$x_i(t)=\begin{cases} w_i(t), & H_o \\ h_i(t)s(t)+w_i(t), & H_1 \end{cases} \tag{1}$$

where, $x_i(t)$ is the received signal in time slot t, at the ith CR , s(t) is PU signal , $w_i(t)$ is additive white Gaussian noise (AWGN) and the complex sensing-channel gain between the ith CR and the PU is $h_i(t)$. Here, it is considered that the sensing time is lower compared to the channel's coherence time. In the process of sensing, sensing channel $h_i(t)$ is time-invariant; that is, $h_i(t) = h_i$. Again, it is assumed that during spectrum sensing, the PU's state remains constant. The energy-detection approach is the best choice for identifying zero-mean constellation signals for an unknown previous information of the PU signal [17]. The probabilities of average false alarm, average detection and average misdetection over AWGN channels for the ith CR with the energy detector are provided, accordingly, by [24].

$$P_{f,i}=\frac{\Gamma\left(u,\frac{\lambda_i}{2}\right)}{\Gamma(u)} \tag{2}$$

$$P_{d,i}=Q_u\left(\sqrt{2\gamma_i},\sqrt{\lambda_i}\right) \tag{3}$$

$$P_{m,i}=1-P_{d,i} \tag{4}$$

Here, 'u' is the energy detector's time-bandwidth product and λ_i and γ_i stand for the energy-detection threshold (mV) and instantaneous SNR at the ith CR, respectively. Q_u (n,x) is the generalized Marcum Q-function and Γ(n, x) is the incomplete gamma function, given by:

$$\Gamma(n,x)=\int_x^{\infty} t^{n-1}e^{-t}dt \tag{5}$$

$$Q_u(n,x)=\frac{1}{n^{u-1}}\int_x^{\infty} t^u e^{-\frac{t^2+n^2}{2}} I_{u-1}(nt)dt \tag{6}$$

I_{u-1} represents the first-kind modified Bessel function of u-1 order.

In CSS, each node takes a binary decision according to its local observations and then sends one bit of the decision D_i (1 representing PU presence, 0 for PU absence) to the common receiver *via* an error-free channel. All 1-bit decisions are combined at the common receiver using a logic rule:

$$Y=\sum_{i=1}^{K} D_i \begin{cases} \geq n, & Hd_1 \\ < n, & Hd_o \end{cases} \tag{7}$$

where Hd_1 and Hd_0 are the inferences of the common receiver that the PU signal is transmitted or not, respectively. The threshold integer 'n' represents the "n-out-of-K" voting rule. The AND rule is applicable to the case of n=K and the OR rule for n=1.

For an AWGN environment, one can assume that $\gamma_1= \ldots=\gamma_K=\gamma$. Again, it can be assumed that all CRs use the same threshold λ, that is, $\lambda_1=\ldots=\lambda_K=\lambda$. For an AWGN channel, $P_{d,i}$ (denoted as P_d) is independent of i. For a Rayleigh fading channel, P_d represents the average $P_{d,i}$ over the statistics of γ_i . For both types of channels, we have $P_m=1-P_d$. So, the probability of false alarm and probability of misdetection are given by [24]:

$$Q_f=P\left(\frac{Hd_1}{H_0}\right)=\sum_{l=n}^{K}\binom{K}{l}P_f^l(1-P_f)^{K-l} \tag{8}$$

$$Q_m=P\left(\frac{Hd_0}{H_1}\right)=1-\sum_{l=n}^{K}\binom{K}{l}P_d^l(1-P_d)^{K-l} \tag{9}$$

The total error rate (TER) is given by: $$TER = Q_f + Q_m \quad (10)$$

Figure 2 and Figure 3 show that the optimal voting rule across all of the simulated detection threshold ranges is n = 5 for AWGN channel and n=2 for Rayleigh channel. Though, for very small fixed thresholds, the AND rule is the optimal rule; i.e., n = 10. For fixed extremely large thresholds, it is the OR rule; i.e., n = 1, that is found to be optimal. The comparison of n values to obtain min TER w.r.t threshold values at various SNRs ranging from 0-15 dB in both AWGN channel and Rayleigh channel is depicted in Table 1.

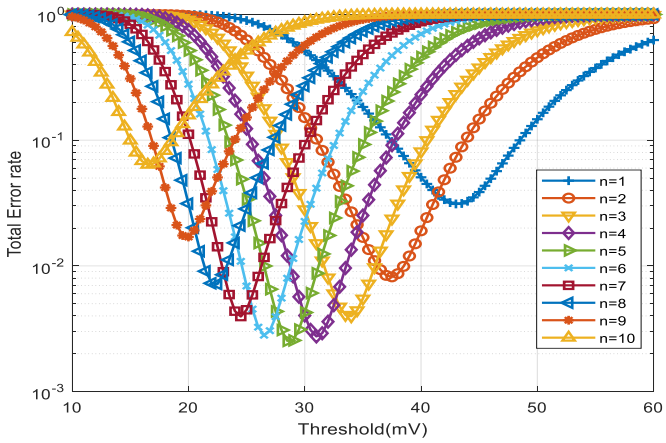


Figure 2. TER of CSS in 10-dB AWGN channel; n is voting rule, K = 10.

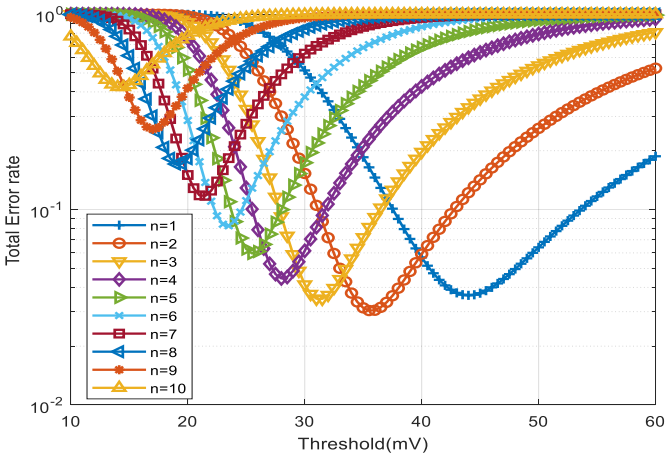


Figure 3. TER of CSS in 10-dB Rayleigh channel; n is voting rule, K = 10.

Table 1. Comparison of n values to get minimum TER w.r.t threshold values at various SNRs

SNR (dB)	Parameter	AWGN Channel	Rayleigh Channel
0	Threshold (mV)	23	25
	Minimum TER	0.696261	0.702793
	n	4	3
5	Threshold (mV)	23.5	30
	Minimum TER	0.2575	0.306736
	n	5	2
10	Threshold (mV)	29	36
	Minimum TER	0.00255473	0.030527
	n	5	2
15	Threshold (mV)	44	44
	Minimum TER	$3.19*10^{-12}$	0.00039336
	n	5	2

Optimal Voting Rule: To obtain the optimal number of CR, TER should be minimized based on [25].

$$n_{optimal} = \min\left(K, \left\lfloor \frac{K}{1+\alpha} \right\rfloor \right) \tag{11}$$

where $\alpha = \dfrac{\ln\left(\frac{P_f}{1-P_m}\right)}{\ln\left(\frac{P_m}{1-P_f}\right)}$ and [.] is the ceiling function. (12)

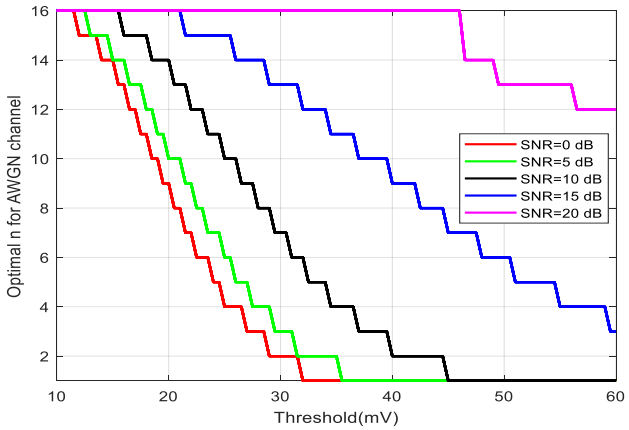


Figure 4. Optimal n *vs.* threshold of CSS for K=16 in AWGN channel.

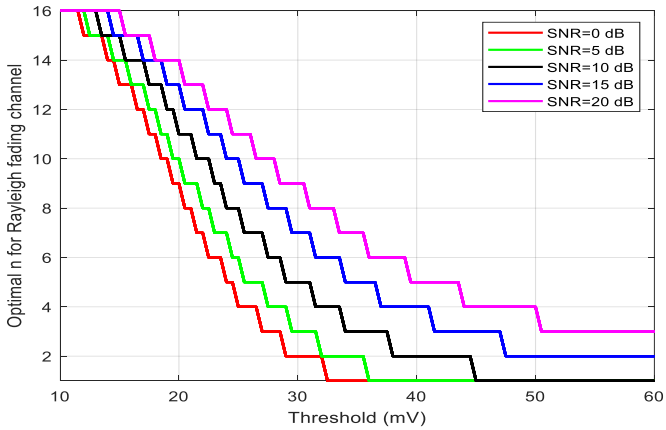


Figure 5. Optimal n *vs.* threshold of CSS for K=16 in Rayleigh channel.

The plot of n *versus* detection threshold determined by Equation 11 is displayed in Figure 4 and Figure 5, for various SNRs varying from 0-20 dB for K=16.

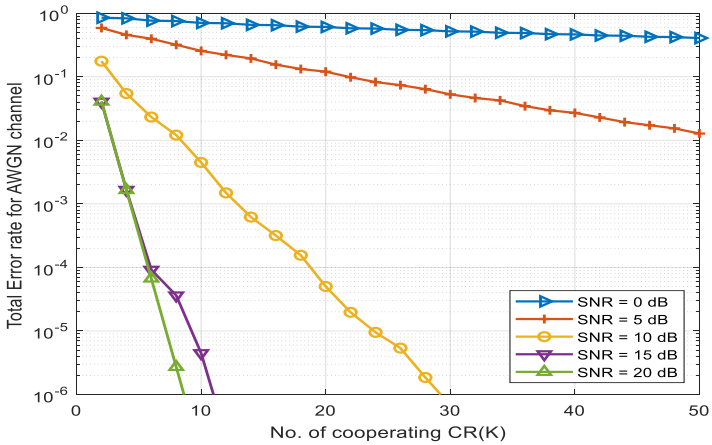


Figure 6. TER *vs.* number of cooperating CRs with λ = 25 mV in AWGN channel.

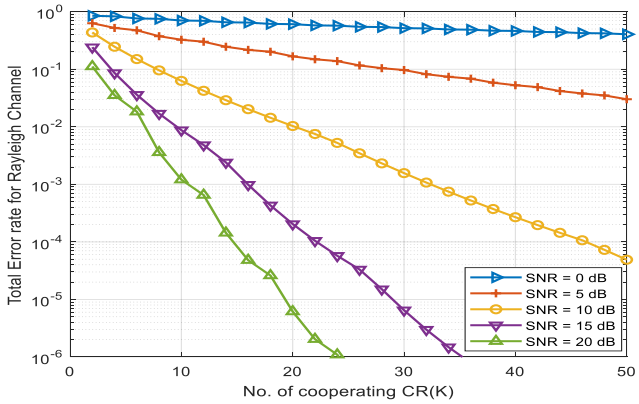


Figure 7. TER *vs.* number of cooperating CRs with λ = 25 mV in Rayleigh channel.

According to Figure 6, in an AWGN channel for a threshold of 25 mV at SNR= 20 dB, 15 dB and 10 dB, respectively, the least numbers of CRs needed to achieve the error-rate objective of 0.001 are 4, 4 and 13. According to Figure 7, in a Rayleigh channel for a threshold of 25 mV at SNR= 20 dB, 15 dB and 10 dB, respectively, the least numbers of CRs needed to achieve the error rate objective of 0.001 are 10, 16 and 33.

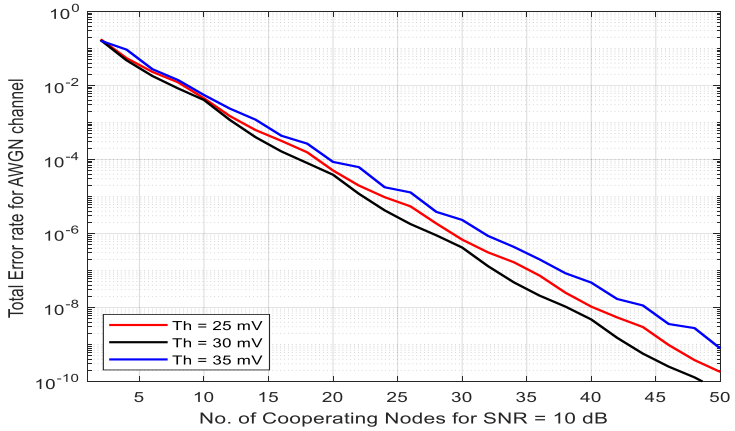


Figure 8. TER vs. number of cooperating CRs in AWGN channel.

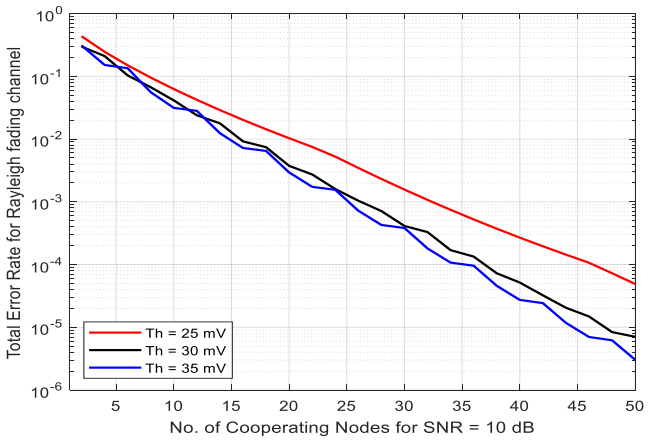


Figure 9. TER *vs.* number of cooperating CRs in Rayleigh channel.

According to Figure 8, in an AWGN channel for an SNR of 10 dB at threshold values of 25, 30 and 35 mV, respectively, the least numbers of CRs needed to achieve the error-rate objective of 0.001 are 13, 12 and 15. According to Figure 9, in a Rayleigh channel for an SNR of 10 dB at threshold values of 25, 30 and 35 mV, respectively, the least numbers of CRs needed to achieve the error-rate objective of 0.001 are 32, 27 and 25.

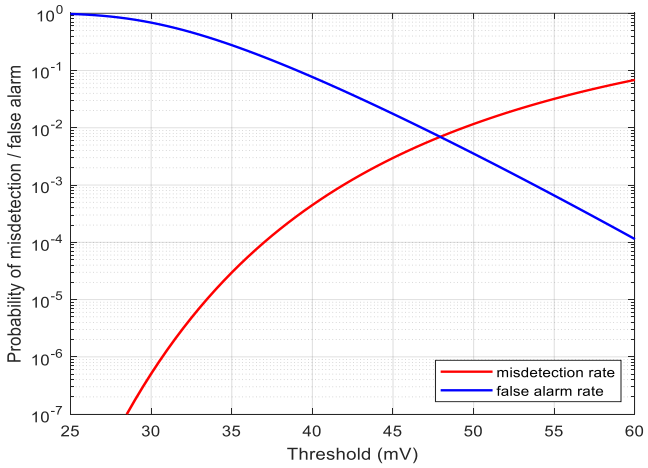


Figure 10. Trade-off between misdetection rate and false-alarm rate.

According to Figure 10, the lowering of the threshold voltage reduces the misdetection rate and increases the false-alarm rate as per Equation 8. Raising the threshold voltage reduces the false-alarm rate and increases the misdetection rate as per Equation 9. The aim is to minimize TER as given in Equation 10, which is the sum of false-alarm rate and misdetection rate. In Equation 11, optimal-voting rule has been implemented to optimize this balance to minimize TER.

2.3 Optimization of TER in Cooperative Spectrum Sensing

Here, two optimization techniques are used to optimize TER: PSO and RGA, where the variable parameters are optimal number of nodes, SNR and threshold.

The GA is a stochastic approach rooted in evolutionary-optimization principles, endeavoring to locate global minima through the emulation of genetics and natural selection. The RGA operates on continuous variables cost function optimization with natural selection and genetic recombination [26], [28]. Adjustments to the RGA involve modifying operator or parameter values. Chromosomes are populated by a group of genes with random values ranging from 0 to 1, constituting the initial population [27], [29]. The population size should balance the need for diversity with computational feasibility. A typical range is 20 to 100 individuals, but larger sizes may be used in complex problems or where more resources are available. In this work, the population size is 20.

Each chromosome's cost within this population is assessed and those with the most favorable values are chosen for the natural-selection process, while the others are discarded. Offspring are generated from these chosen-parent chromosomes. The weight 'z' is determined through the utilization of a random number 'r' and the cross-over operator 'μ' as [26]-[27]. The cross-over rate is typically set between 0.7 and 0.9, with values closer to 1 being preferred for problems where exploration is more important. In this work, the cross-over rate is 0.9.

$$z = \begin{cases} (2r)^{\frac{1}{1+\mu}} & \text{if } r > 0.5 \\ \left(\frac{1}{2(1-r)}\right)^{\frac{1}{1+\mu}} & \text{otherwise} \end{cases} \tag{13}$$

New offspring are:

$$Offspring\,1 = \frac{(1+z)\,parent_1 + (1-z)\,parent_2}{2}$$
$$Offspring\,2 = \frac{(1-z)\,parent_1 + (1+z)\,parent_2}{2} \tag{14}$$

A sub-set of randomly chosen chromosomes undergoes mutation using the mutation operator 'η' and mutation weight 'm', The mutation rate is typically kept between 0.01 and 0.1, depending on the complexity of the problem. In this work, the mutation rate is 0.1. For large or complex search spaces, a higher mutation rate might be appropriate, where,

$$m = \begin{cases} (2r)^{\frac{1}{1+\eta}} - 1 & \text{if } r \le 0.5 \\ 1 - [2(1-r)]^{\frac{1}{1+\eta}} & \text{otherwise} \end{cases} \quad (15)$$

The PSO algorithm, another variant of evolutionary algorithms, is employed to discover optimal settings or parameters necessary for obtaining a desired objective [30]-[31]. Each distinct solution within the search space of the objective function is denoted as a particle, with the initial collection of random particles forming the starting position of the swarm. These particles have the capability to assess their current fitness or positions through optimization functions. Swarms, composed of randomly generated solutions, iteratively explore the design space towards finding the optimal solution. Each particle's velocity is modified according to solution of individual best position, termed as particle best (pbest) and the best value encountered so far by any particle in the particle swarm optimizer, termed as global best (gbest). Every particle has a velocity vector $v_j(l)$ and a position vector $y_j(l)$. The equation for updating velocity is given by [30]-[31]:

$$v_j(l+1) = w \times v_j(l) + c_1 \times rand \times [pbest - y_j(l)] + c_2 \times rand \times [gbest - y_j(l)] \quad (16)$$

and the position update equation is:

$$y_j(l+1) = y_j(l) + v_j(l+1) \quad (17)$$

where, v_j - particle velocity of j^{th} iteration, w - inertia weight factor, which is a random number between (0,1), y_j - current particle position, c_1 ,c_2 are cognitive parameter and social parameter, generally, $c_1 + c_2 = 4$. Typically, $c_1 = c_2 = 2$ is a common choice, though in some cases, a slight imbalance (e.g., c_1=1.5, c_2=2.5) can promote better exploration early on, followed by stronger exploitation later. In this work, c_1=1.65, c_2 =2.35. At each iteration, the particle adjusts its position and velocity according to the procedures mentioned above in order to achieve the best solution.

2.4 Algorithm for the Implementation of PSO to Optimize TER

Step-1: Initialize parameters, like size of population, iteration number, inertia weight, personal and global learning coefficients and limit of velocity.

Step-2: Compute fitness by applying the cost function of TER (Equation 10) to each particle for both pbest and gbest solutions.

Step-3: Continuously update the velocity and position of every particle using Equation 16 and Equation 17 and repeat steps 3 and 4 until the convergence of population.

Step-4: Choose the gbest solution based on the minimum value of the cost function TER.

2.5 Algorithm for the Implementation of RGA to Optimize TER

Step-1: Initialize variables, like optimal node number, SNR and threshold. Establish upper and lower bounds of parameters along with defining the population size, mutation rate and number of generations.

Step-2: Compute fitness by applying the cost function of TER (Equation 10).

Step-3: The process involves selection, arithmetic crossover, mutation and computation of temporary fitness.

Step-4: Repeat step 3 until the convergence of population.

Step 5: Choose the threshold value with the best fitness where the TER is minimized.

3. RESULTS

The cost-function plot for RGA and PSO optimization are depicted in Figure 11.

The cost function is the TER of Equation 10, for RGA and PSO optimization. The aim is to minimize TER. The population size is 20, the iteration number is 100, the optimal number of nodes is varied from 1 to 16, SNR is varied from 0 dB to 20 dB, the threshold is varied from 25 mV to 35 mV and 0.1 is the mutation rate. The convergence of TER is shown in Figure 12 (a) and 12 (b). PSO generally

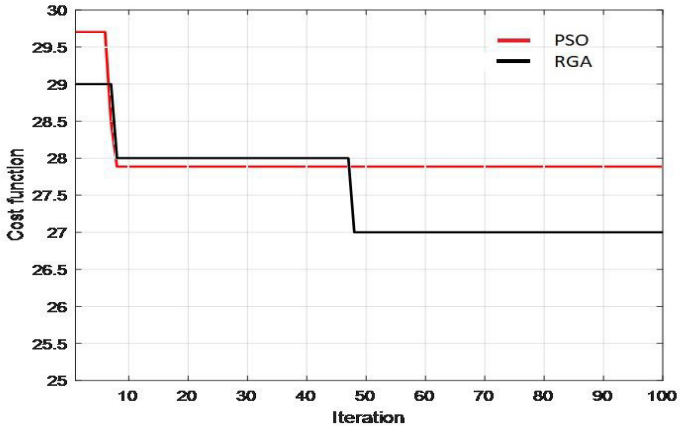
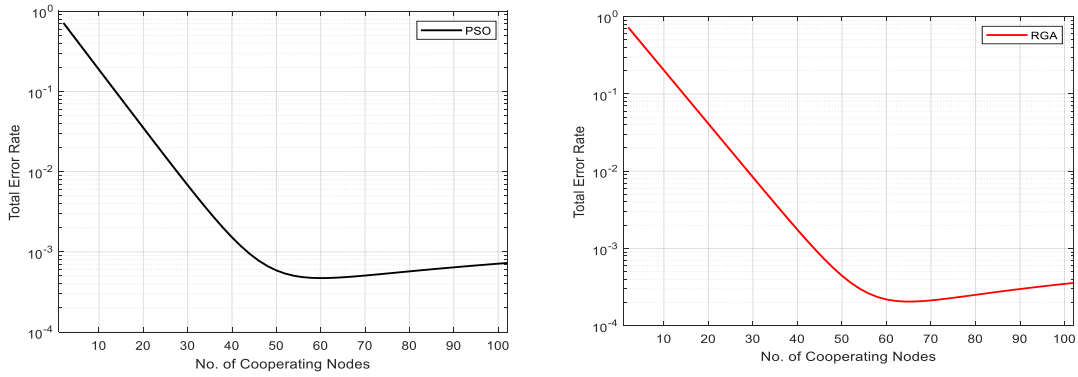


Figure 11. Cost function *vs.* iteration.



(a) (b)

Figure 12. Convergence of TER (a) PSO (b) RGA.

converges faster, but is more prone to local optima, whereas RGA provides better exploration at the cost of slower convergence. The stopping criterion is minimal value of cost function given in Equation 10 for both PSO and RGA algorithms, mentioned in step 4 of sub-section 2.4 for PSO algorithm and mentioned in step 4 of sub-section 2.5 for RGA algorithm.

The RGA and PSO optimized results of TER are compared with the results obtained without optimization [25], [32]-[34] in Figure 13.

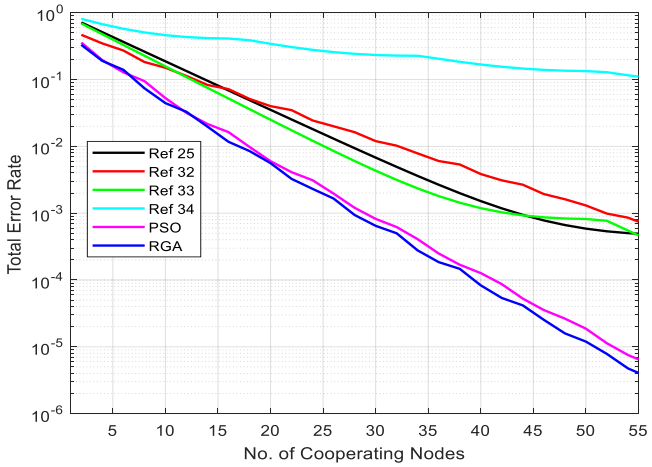


Figure 13. Comparison of TER with published results.

According to Figure 13, in a Rayleigh channel for SNR value of 10 dB, the least numbers of CRs needed to achieve the error-rate objective of 0.001 are 52, 44, 42, 29 and 27.

4. DISCUSSION

The research presented demonstrates the significant impact of PSO and RGA optimization techniques on the performance of cooperative communication systems in wireless sensor networks. By focusing on the optimization of TER, the study addresses a critical aspect of data-transmission reliability and efficiency in WSNs, particularly under constraints, such as limited resources and challenging environmental conditions.The study varies the number of nodes from 1 to 16 to analyze how the network density affects TER. The findings suggest that as the number of nodes increases, the cooperative relaying becomes more effective, leading to a reduction in TER. This can be attributed to the enhanced diversity and redundancy provided by additional nodes, which improves the overall link quality. The SNR is varied from 0 dB to 20 dB to observe its influence on TER. Higher SNR values typically result in better signal quality and lower error rates. The optimization techniques help in maintaining low TER even at lower SNR values, indicating the robustness of the CCS under different noise conditions. The threshold-voltage range of 25 mV to 35 mV is identified as optimal for minimizing TER. This range strikes a balance between sensitivity and noise immunity, ensuring that the nodes can effectively discriminate between signal and noise. The optimized threshold voltage significantly reduces false alarm and misdetection rates, which are crucial for reliable data transmission. The mutation rate of 0.1 used in the genetic algorithm ensures a good exploration of the solution space without causing excessive disruption to the convergence process. This rate is found to be effective in maintaining genetic diversity and preventing premature convergence to sub-optimal solutions. The computational complexity of RGA can be expressed as $O(G * P * (C_f + \log P))$, where O() is the Big O order, P is the population size, G is the number of generations and C_f is the cost of evaluating the fitness function. In this work, P=20 and G=100. The computational complexity of PSO can be expressed as $O(G * P * (C_f + 1))$. In this work, P=20, G=100 and PSO generally has lower complexity per iteration compared to RGA since it avoids complex selection, crossover and mutation operations. Implementing RGA and PSO in real-world WSNs requires addressing hardware limitations by using lightweight versions of the algorithms that reduce computational and memory demands. Energy consumption is critical, so adaptive techniques should minimize computation time to preserve the battery life. Scalability can be achieved by using clustering or hierarchical approaches, ensuring that the algorithms perform efficiently even in large networks. The dynamic nature of WSNs necessitates fault-tolerant and adaptive versions of RGA and PSO to handle node failures and topology changes. Finally, multi-objective optimization and hybrid approaches can improve performance by balancing trade-offs between energy, coverage and network lifetime.

5. CONCLUSION

This paper has investigated the optimization of error rates, specifically focusing on false alarm and misdetection, within cooperative communication frameworks to meet a specified threshold voltage. The change of TER for a 50-node cooperative system in a WSN is presented here for both AWGN and Rayleigh channels. The effects of optimal number of nodes, SNR and threshold on TER are simulated. Furthermore, an efficient optimization approach that satisfies the given bound error has been assessed; it requires fewer cognitive radios than the total number used in cooperative spectrum sensing. Finally, the minimum TER is achieved using RGA optimization, as shown in Figure 13. This research contributes to the broader goal of improving communication efficiency and robustness in resource-constrained and challenging environments, paving the way for enhanced applications in various fields, such as environmental monitoring, healthcare and industrial automation.

REFERENCES

[1] M. A. Hossain, R. Md Noor, K. -L. A. Yau, I. Ahmedy and S. S. Anjum, "A Survey on Simultaneous Wireless Information and Power Transfer with Cooperative Relay and Future Challenges," IEEE Access, vol. 7, pp. 19166-19198, DOI: 10.1109/ACCESS.2019.2895645, 2019.

[2] M. Peng, Y. Liu, D. Wei, W. Wang and H.-H. Chen, "Hierarchical Cooperative Relay Based Heterogeneous Networks," IEEE Wireless Communications, vol. 18, no. 3, pp. 48-56, June 2011.

[3] M. Kothari et al., "Massive MIMO Pre-coders for Cognitive Radio Network Performance Improvement: A Technological Survey," Proc. of Machine Learning, Deep Learning and Computational Intelli. for Wireless Comm., in: Lecture Notes in Electrical Eng., vol. 749, Singapore, 2021.

[4] W. Guo, N. M. F. Qureshi et al., "Cooperative Communication Resource Allocation Strategies for 5G and Beyond Networks: A Review of Architecture, Challenges and Opportunities," Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 10A, pp. 8054-8078, 2022.

[5] Y-W. Peter Hong et al., Cooperative Communications and Networking: Technologies and System Design, ISBN: 978-1-4419-7194-4, Springer Science & Business Media, 2010.

[6] A. S. Ibrahim, A. K. Sadek, W. Su and K. J. R. Liu, "Cooperative Communications with Relay-selection: When to Cooperate and Whom to Cooperate with?," IEEE Transactions on Wireless Communications, vol. 7, no. 7, pp. 2814-2827, DOI: 10.1109/TWC.2008.070176, July 2008.

[7] X. -H. You, D. -M. Wang, B. Sheng, X. -Q. Gao, X. -S. Zhao and M. Chen, "Cooperative Distributed Antenna Systems for Mobile Communications [Coordinated and Distributed MIMO]," IEEE Wireless Communications, vol. 17, no. 3, pp. 35-43, DOI: 10.1109/MWC.2010.5490977, June 2010.

[8] A. Khan, S. Rehman, M. Abbas et al., "On the Mutual Information of Relaying Protocols," Physical Communication, vol. 30, pp. 33–42, October 2018.

[9] Y. Cao and H. Pan, "Energy-efficient Cooperative Spectrum Sensing Strategy for Cognitive Wireless Sensor Networks Based on Particle Swarm Optimization," IEEE Access, vol. 8, pp. 214707-214715, DOI: 10.1109/ACCESS.2020.3037707, 2020.

[10] S. K. Gannamaneni and J. S. Roy, "Performance of Optimization Methods for Energy Efficiency in Cooperative Communication," FACTA Universitatis Series: Electronics and Energetics Journal, vol. 36, no. 3, pp. 329-341, DOI: 10.2298/FUEE2303329G, Sep. 2023.

[11] X. Liu et al., "Throughput Maximization for RIS-UAV Relaying Communications," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 10, pp. 19569-19574, Oct. 2022.

[12] K. G. Rallis et al, "Energy Efficient Cooperative Communications in Aggregated VLC/RF Networks with NOMA," IEEE Transactions on Communications, vol. 71, no. 9, pp. 5408-5419, Sept. 2023.

[13] M. Bargavi, A. P. Singh and C. P. Lora, "Secure Energy-efficient Resource Allocation and Relay Selection for Cooperative Communications in Wireless Sensor Networks," Proc. of the Int. Conf. on Optimization Computing and Wireless Comm. (ICOCWC), pp. 1-6, Debre Tabor, Ethiopia, 2024.

[14] S. Wang and J. Nie, "Energy Efficiency Optimization of Cooperative Communication in Wireless Sensor Networks," EURASIP Journal on Wireless Communications and Networking, vol. 2010, Article ID: 162326, pp. 1-8, DOI: 10.1155/2010/162326, May 2010.

[15] Y. Zheng, J. Hu and K. Yang, "SWIPT Aided Cooperative Communications with Energy Harvesting-based Selective-decode-and-forward Protocol: Benefiting from Channel Aging Effect," IEEE Transactions on Green Communications and Networking, vol. 7, no. 3, pp. 1192-1204, Sept. 2023.

[16] B. P. Chaudhary and R. K. Mishra, "Performance Analysis of SWIPT Cooperative-NOMA over Rayleigh Fading Channel," Proc. of the 15th Int. Conf. on Computer and Automation Engineering (ICCAE), pp. 541-545, DOI: 10.1109/ICCAE56788.2023.10111338, Sydney, Australia, 2023.

[17] G. Mahendru, A. Shukla and P. Banerjee, "A Novel Mathematical Model for Energy Detection Based Spectrum Sensing in Cognitive Radio Networks," Wireless Personal Communication, vol. 110, pp. 1237–1249, DOI: 10.1007/s11277-019-06783-3, 2020.

[18] E. Soltanmohammadi, M. Orooji and M. Naraghi-Pour, "Improving the Sensing–Throughput Tradeoff for Cognitive Radios in Rayleigh Fading Channels," IEEE Transactions on Vehicular Technology, vol. 62, no. 5, pp. 2118-2130, DOI: 10.1109/TVT.2012.2236116, Jun. 2013.

[19] A. Kumar, P. Thakur, S. Pandit and G. Singh, "Threshold Selection and Cooperation in Fading Environment of Cognitive Radio Network: Consequences on Spectrum Sensing and Throughput," AEU-International Journal of Electronics and Communications, vol. 117, p. 153101, 2020.

[20] V. M. Patil et al., "Signal Detection in Cognitive Radio Networks over AWGN and Fading Channels," Int. J. of Wireless Information Networks, vol. 25, no. 1, pp. 79-86, 2018.

[21] R. Ahmad, W. Alhasan, R. Wazirali and N. Aleisa, "Optimization Algorithms for Wireless Sensor Networks Node Localization: An Overview," IEEE Access, vol. 12, pp. 50459-50488, 2024.

[22] S. Leelakrishnan and A. Chakrapani, "Power Optimization in Wireless Sensor Network Using VLSI Technique on FPGA Platform," Neural Processing Letters, vol. 56, no. 2, sp. 125, Mar. 2024.

[23] S. Kamel, A. Al Qahtani and A. S. M. Al-Shahrani, "Particle Swarm Optimization for Wireless Sensor Network Lifespan Maximization," Engineering, Technology & Applied Science Research, vol. 14, no. 2, pp. 13665–13670, DOI: 10.48084/etasr.6752, Apr. 2024.

[24] F. F. Digham et al., "On the Energy Detection of Unknown Signals over Fading Channels," Proc. of the IEEE Int. Conf. on Communications (ICC '03), vol.5, pp. 3575-3579, Anchorage, AK, USA, 2003.

[25] W. Zhang, R. K. Mallik and K. B. Letaief, "Optimization of Cooperative Spectrum Sensing with Energy Detection in Cognitive Radio Networks," IEEE Transactions on Wireless Communications, vol. 8, no. 12, pp. 5761-5766, DOI: 10.1109/TWC.2009.12.081710, Dec. 2009.

[26] K. Deb and A. Kumar, "Real-coded Genetic Algorithms with Simulated Binary Crossover: Studies on Multi-modal and Multi-objective problems," Complex Systems, vol. 9, no. 6, pp. 431-454, 1995.

[27] R. L. Haupt and S. E. Haupt, Practical Genetic Algorithms, 2nd Edn., New York: Wiley, 2004.

[28] P. Nandi and J. S. Roy, "Performance Comparison of Optimization Methods for Flat-top Sector

Beamforming in a Cellular Network," Journal of Telecommunication and Information Technology, vol. 2022, no. 3, pp. 39-46, DOI: 10.26636/jtit.2022.162122, Sept. 2022.

[29] P. Nandi and J. S. Roy, "Side Lobe Reduction of Phased Array Antenna Using Genetic Algorithm and Particle Swarm Optimization," Int. Journal of Microwave and Optical Technology (IJMOT), vol. 11, no. 3, pp. 211-218, May 2016.

[30] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," Proc. of the IEEE Int. Conf. on Neural Networks, pp. 1942-1948, San Francisco, USA, 1995.

[31] A. Deb, J. S. Roy and B. Gupta, "Performance Comparison of Differential Evolution, Particle Swarm Optimization and Genetic Algorithm in the Design of Circularly Polarized Microstrip Antennas," IEEE Trans. Antennas & Propagation, vol. 62, no. 8, pp. 3920-3928, Aug. 2014.

[32] S. K. Ghosh, S. R. Trankatwar and P. Bachan, "Optimal Voting Rule and Minimization of Total Error Rate in Cooperative Spectrum Sensing for Cognitive Radio Networks," Journal of Telecommunications and Information Technology, vol.1, pp. 43–50, DOI: 10.26636/jtit.2021.144420, 2021.

[33] B. S. Karumanchi and N. R. Banavathu, "Cooperative Spectrum Sensing in Cognitive Radio Network Using Selective Soft-information Fusion Scheme," Proc. of the 2023 IEEE Region 10 Conf. (TENCON 2023), pp. 1193-1197, DOI: 10.1109/TENCON58879.2023.10322321, Chiang Mai, Thailand, 2023.

[34] J. Zhang, X. Xiao, "Soft Fusion-based Cooperative Spectrum Sensing Using Particle Swarm Optimization for Cognitive Radio Networks in Cyber-physical Systems," Concurrency Computation Practice and Experience, vol. 35, no. 13, DOI: 10.1002/cpe.6295, Apr. 2021.

ملخص البحث:

تتضمّنُ أنظمة الاتّصالات التّعاونية التّنسيق بين عُقَد المجسّات لنقْل البيانات بشكلٍ أكثر فاعلية، وخصوصاً في السّيناريوهات الّتي تنْطوي على مصادر محدودةٍ وظروفٍ بيئيةٍ تحتوي على تحدِّياتٍ. وإنّ تحسين معدَّل الخطأ الكُلّيّ (TER) في الاتّصالات التّعاونية لتقريبه من القيمة المثالية في شبكات المجسّات اللاسلكية هو مهمَّة حاسمة من أجل تحسين موثوقية نقْل البيانات وفعاليته. ومن الممكن تحسين جودة الرّبط في تلك الشّبكات عن طريق نقْل البيانات تعاونياً بقيمةٍ منخفضةٍ قَدْر الإمكان لمعدَّلِ الخطأ الكُلّيّ (TER).

في هذه الورقة، نستخدم خوارزميةً جينية مرمَّزةً بشكلٍ حقيقي (RGA) إلى جانب تقنية التّحسين المرتكزة على سِرْبِ الجُزَيئات (PSO) في شبكات المجسّات اللاسلكية لخفْض قيمة مُعدّل الخطأ الكُلّيّ. ويتمُّ في هذه الدّراسة تغيير عدد العُقد بين 1 و 16، ونسبة الإشارة إلى الضّجيج بين 0dB و dB 20. أمّا جُهْد العتبة فيجري تغييره بين 25 ميلي فولت و 35 ميلي فولت، بينما كان معدّل التّحوّل 0.1. ويتمُّ الحصول على القيمة الدنيا لمعدّل الخطأ الكُلّيّ لجُهْد عتبة يتراوح بين 25 ميلي فولت و 35 ميلي فولت مُقارنةً بقيمة معدذل الخطأ الكُلّيّ الّتي يتمُّ الحصول عليها دون إجراء عملية التّحسين. وقد برهنت عملية التّحسين على إحداث تحسيناتٍ ملحوظةٍ للحصول على جُهْد العتبة المرغوب بقيمةٍ دُنيا لمعدّل الإنذارات الكاذبة ومعدّل الكشف الخاطىء؛ من أجل تحسين الأداء الإجمالي لنظام الاتّصالات التّعاوني في شبكات المجسّات اللاسلكية.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

PRIVACY-AWARE MALARIA DETECTION: U-NET MODEL WITH K-ANONYMITY FOR CONFIDENTIAL IMAGE ANALYSIS

Ghazala Hcini and Imen Jdey

(Received: 6-Aug.-2024, Revised: 1-Oct.-2024, Accepted: 24-Oct.-2024)

ABSTRACT

Malaria detection through cell-image analysis is essential for early diagnosis and effective treatment, as timely detection can significantly reduce the risk of severe health complications. However, this process raises substantial privacy concerns due to the sensitivity of medical data. This study presents a U-Net model combined with k-anonymity to enhance data security while maintaining high accuracy. The model features a custom Spatial Attention mechanism for improved segmentation performance and incorporates advanced techniques to focus on critical image features. K-Anonymity adds controlled noise to protect data privacy by obfuscating sensitive information. The model achieved a validation accuracy of 99.60%, a Dice score of 99.61%, a precision of 99.42%, a recall of 99.96% and an F1-score of 99.69% on malaria cell images. When applied to the Cactus dataset, a real dataset, in agriculture, it achieved an accuracy of 98.58%, an F1-score of 98.44%, a Dice score of 95.08%, a Precision of 98.04% and a Recall of 98.86%, demonstrating its strong generalization capability. These results highlight the effectiveness of integrating privacy-preserving techniques with advanced neural-network architectures, improving both security and performance in diverse image-analysis applications.

KEYWORDS

Deep learning, U-net architecture, Spatial-attention mechanism, K-anonymity, Privacy preservation, Cross-domain transfer.

1. INTRODUCTION

Malaria remains a major global health challenge, causing millions of death cases every year, especially in tropical and sub-tropical areas [1]-[2]. The World Health Organization estimates that there were around 249 million malaria cases worldwide in 2022 [3], resulting in over 600,000 deaths, primarily among vulnerable populations (Figure 1). The early and accurate detection of malaria is crucial for effective treatment and disease control [4]. However, traditional diagnostic methods, like microscopy, are time-consuming and require skilled personnel, often causing delays in diagnosis and treatment [5]-[6].

Image segmentation is widely acknowledged as a crucial and fundamental task in image analysis [7]-[8]. It serves as the initial step in extracting significant information from images. The main goal of image segmentation is to divide an image into distinct segments, which allows for easier representation of objects and measurement of features. The accuracy of feature measurement is greatly affected by the quality of segmentation, emphasizing its importance in various medical imaging applications. The automation of medical-image segmentation plays a vital role in disease diagnosis, pathology localization, anatomical-structure study, treatment planning and integration with computer-assisted surgical systems [9].

Machine learning (ML) and deep learning (DL) have become integral solutions across various domains [10], particularly in healthcare, where they are leveraged for tasks, such as diagnosis, treatment planning and patient monitoring. These technologies enable the analysis of vast datasets, allowing for real-time insights that can significantly enhance healthcare outcomes. In recent years, DL techniques, specifically convolutional neural networks (CNNs) [11], have revolutionized medical-image analysis [12]-[13], providing advanced capabilities for automating disease detection and improving diagnostic accuracy [14]-[15].

Among the various architectures developed for biomedical-image segmentation, U-Net has gained prominence due to its ability to effectively capture detailed information [16]-[18]. Its encoder-decoder

G. Hcini and I. Jdey are with Department of Computer Science, ReGIM-Lab. REsearch Groups in Intelligent Machines, Sfax, Tunisia. Emails: hcinighazala@fstsbz.u-kairouan.tn and imen.jdey@fstsbz.u-kairouan.tn

structure enables precise segmentation of malaria parasites in blood smears [19], thereby facilitating rapid and accurate diagnosis. However, despite its effectiveness, the U-Net model sometimes struggles to distinguish fine details in complex images, which is crucial for accurate parasite detection. To improve the segmentation performance, this paper introduces a spatial-attention mechanism into the U-Net architecture. The spatial-attention mechanism allows the model to focus on relevant regions of the image, thereby enhancing its ability to discern subtle features and improving the overall detection accuracy [20]-[21].

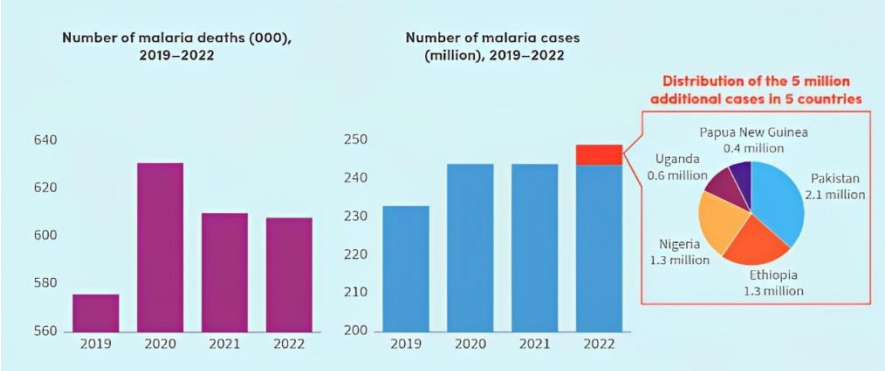


Figure 1. Malaria report: number of deaths and number of cases (2019-2022).

In addition to enhancing detection capabilities, integrating artificial intelligence in healthcare also brings up significant privacy concerns regarding handling sensitive patient data [22]. As healthcare data continues to become more digitized, it becomes crucial to maintain the confidentiality and security of patient information [23]. To address these concerns, this paper introduces a new approach that combines the enhanced U-Net model with k-anonymity techniques. This approach establishes a privacy-aware framework for malaria detection. K-anonymity is a well-established method for preserving privacy [24], as it ensures that individual data points cannot be distinguished from at least k-1 other data points [25]. This protects patient identities while allowing the model to learn from the data effectively.

K-anonymity plays a crucial role in mitigating re-identification risks, making it effective in reducing the likelihood of identifying individuals within anonymized datasets. This enhances privacy protection when sharing sensitive health information, which is increasingly demanded by researchers and regulatory bodies. The method enables the safe sharing of personal health data, facilitating advancements in medical research while striving to protect individual privacy. El Emam and Dankar [26] proposed significant improvements to traditional k-anonymity methodologies, suggesting modifications that better balance privacy protection with data utility. Their study indicates that a hypothesis-testing approach offers superior control over re- identification risks compared to baseline k-anonymity, thereby minimizing unnecessary data distortions. This advancement underscores the importance of adapting existing frameworks to meet the evolving demands of data sharing in healthcare while maintaining robust privacy safeguards.

The main contributions of our study are:

1) It developed a U-Net model combined with k-anonymity technique to enhance data privacy while maintaining high accuracy for malaria cell-image segmentation.
2) It incorporated a custom spatial-attention mechanism into the U-Net model to improve segmentation performance and focus on critical image features.
3) It demonstrated generalization capability by applying our proposed model to the Cactus dataset, a real dataset, from agriculture.

The paper is organized into four main sections: Related Works, Materials and Methods, Results and Discussion and Conclusion. The Related Works section reviews the literature on malaria detection with DL. The Materials and Methods section outlines the dataset, pre-processing steps, the architecture of the spatial attention-enhanced U-Net model and the implementation of k-anonymity. The Results and Discussion section presents experimental findings, analyzes the impact of the attention mechanism and evaluates the effectiveness of the privacy approach. Finally, the Conclusion section summarizes key findings and suggests future-research directions.

2. RELATED WORKS

Recent advancements in malaria cell-image segmentation have significantly improved diagnostic capabilities. Traditional methods laid the groundwork, but the field has evolved by adopting advanced techniques that enhance accuracy and reliability. The following review explores key recent studies that have pushed the boundaries of automated malaria detection, showcasing the progress and innovations in this critical area of medical imaging.

The proposed approach in [27] involves creating an automated system for malaria detection using a decision-tree classifier. The methodology includes pre-processing blood smear images through segmentation with the Canny edge detector and feature extraction using Hu Moments. The data is normalized to ensure consistency. The decision-tree classifier, trained and validated with 5-fold cross-validation, achieved an accuracy of 77.32%, a precision of 77.31%, a recall of 77.37% and an F1-score of 77.48%. This approach highlights the effectiveness of Hu Moments and decision-tree classification for distinguishing malaria- infected from uninfected images, offering potential improvements in diagnostic accuracy and efficiency in clinical settings. Future work should explore advanced techniques and real-time image acquisition to enhance practical applicability.

In [28], the authors proposed a hybrid model combining Capsule Neural Networks (CapsNet) with CNNs for malaria detection from blood smear microscopic images. The approach involved processing and enhancing images through rotation before feeding them into the hybrid CapsNet model. Optimized with a learning rate of 0.07 and a batch size of 20, the model demonstrated significant improvements over traditional methods. The hybrid CapsNet model achieved a detection rate of 99%, an accuracy of 99.08% and a False Acceptance Rate (FAR) of 0.97%, surpassing the DSCN-Net model, which recorded a detection rate of 98%, an accuracy of 97.2% and an FAR of 0.99%. This method underscored the enhanced efficacy of the hybrid CapsNet model in malaria detection.

The main contribution of [29] is the application of the U-Net architecture for accurate Plasmodium segmentation in thin blood smear images. The study demonstrates U-Net's effectiveness in this biomedical-imaging task and compares three loss functions—mean squared error, binary cross-entropy and Huber loss. The results reveal that Huber loss achieves the best performance metrics, with an F1-score of 92.97%, a positive predictive value (PPV) of 0.9715, a sensitivity (SE) of 89.57% and a relative segmentation accuracy (RSA) of 90.96%. These findings highlight Huber loss's ability to enhance segmentation accuracy and reliability for malaria diagnosis.

The proposed method, in [30] introduces an automated system for detecting malaria parasites in microscopic blood images. It uses bilateral filtering to enhance image quality by removing noise, followed by adaptive thresholding and morphological image processing to identify malaria parasites within individual cells. Tested on the NIH Malaria dataset, this approach achieved a detection accuracy exceeding 91%, outperforming existing methods. This algorithm provides a reliable and efficient tool for pathologists and hematologists, aiding in the accurate and timely detection of malaria.

The proposed method in [31] focuses on enhancing the analysis of malaria by automating the identification of parasitized red blood cells. It compares the performance of various models; Support Vector Machines (SVM), XG-Boost and neural networks, demonstrating that CNNs offer superior results. In experiments involving 13,750 parasitized and 13,750 non-parasitized samples, CNNs achieved an accuracy rate of 97%, outperforming SVMs (94%), XG-Boost (90%) and traditional neural networks (80%). This DL approach provides a highly accurate and robust solution for detecting malaria, significantly improving decision-making in medical diagnostics.

In [32], the proposed method involves training various object detection neural networks; YOLOv5x, Faster R-CNN, SSD and RetinaNet, on a dataset of 2,571 labeled thick blood smear images for detecting Plasmodium parasites. YOLOv5x demonstrated a high performance with a precision of 92.10%, recall of 93.50%, an F-score of 92.79% and mAP0.5 of 94.40% for detecting leukocytes, early and mature Plasmodium trophozoites. Attention modules were also tested, but showed no significant improvement over YOLOv5x. To further enhance the diagnostic process, a 3D-printed robotic system was designed for automating optical microscopy, enabling auto-focusing and slide tracking. Integrated into the iMAGING smartphone application, this system provides fully automated malaria diagnostics, including the ability to determine Plasmodium infection and parasite levels in Giemsa-stained thick blood smear samples.

In [33], the authors proposed a novel semantic segmentation neural-network architecture for rapid malaria detection. This method quickly generates classification masks that indicate the position, shape and type of detected elements in blood samples. Addressing the challenges of manual diagnosis, the approach uses light microscope imagery to classify cells into three categories: healthy, malaria-infected and background. The generated masks, which can be color-coded for better visualization, facilitate semi-automatic disease recognition while leaving the final diagnosis to specialists. The system demonstrated a high recognition accuracy of 96.65% with minimal computational demands, thus enhancing diagnostic speed and reducing misclassification rates by providing valuable additional information to healthcare providers [33].

The main contribution of [34] is the development of RBCNet, a novel pipeline for red blood-cell detection in blood smear microscopy images. RBCNet integrates a dual DL architecture: a U-Net for initial cell-cluster segmentation and a Faster R-CNN for refined detection of small cell objects. This approach, based on cell clustering rather than region proposals, enhances robustness to cell fragmentation and scalability for fine-scale structures. Tested on nearly 200,000 labeled cells from malaria smears, RBCNet achieved over 97% detection accuracy. The pipeline significantly improves detection precision and reduces false alarms, marking a crucial step toward automated malaria diagnosis.

Maqsood, A. et al. [35] proposed a custom CNN-based architecture consisting of 5 convolutional layers, 5 max-pooling layers and 2 fully connected layers. Augmentation techniques were employed to enhance the features of red blood cells before training the model. The model was evaluated using the NIH malaria benchmark dataset.

In [36], the authors modified the YOLOv4 architecture by layer pruning and backbone replacement to improve its efficiency and accuracy in detecting malaria-infected cells. By strategically removing residual blocks from layers C3 to C5 and replacing the CSP-DarkNet53 backbone with a shallower ResNet50 network, the study successfully created a lighter model that maintains a strong performance.

Chaudhry, H. et al. proposed a DL approach that not only classifies the malaria parasite type, but also identifies the life-cycle stage of the infected cell. The proposed architecture is more than twenty times lighter than the widely used Dense Convolutional Network (DenseNet) and contains less than 0.4 million parameters, making it a suitable option for mobile applications in economically disadvantaged regions for malaria detection [37].

In the context of evaluating image-segmentation models, several key metrics are commonly used to measure performance comprehensively. Accuracy is frequently used to assess the overall correctness of model predictions (Equation 1). Specificity refers to the proportion of true negative predictions among all actual negatives (Equation 2). Precision and Recall are critical for understanding the model's performance in distinguishing between classes, focusing on the accuracy of positive predictions and recall highlighting the model's ability to identify all positive instances (Equations 3, 4). The F1-score is a balanced metric by combining precision and recall, providing a single measure that considers both false positives and false negatives (Equation 5). In image segmentation, the Dice coefficient is a popular metric for evaluating the overlap between predicted and true masks, offering insights into the model's segmentation accuracy (Equation 6). These metrics collectively contribute to a robust evaluation framework, ensuring a well- rounded assessment of model performance across different tasks [38].

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{1}$$

$$Specificity = \frac{(TN)}{(TN+FN)} \tag{2}$$

$$\frac{Sensitivity}{Recall} = \frac{TP}{(TP+FN)} \tag{3}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{4}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

$$Dice = \frac{2.|A \cap B|}{|A|+|B|} \tag{6}$$

where:

- TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative).
- $|A|$ is the total number of elements in the predicted segmentation.
- $|B|$ is the total number of elements in the ground-truth segmentation.
- $|A \cap B|$is the number of elements in the intersection of the predicted and ground-truth segmentations.

Table 1. Summary of recent related works in malaria segmentation.

Study	Year	Method	Dataset	Performance
[27]	2024	Decision Tree	27,558 images	Accuracy=77.32% Precision = 77.31% Recall = 77.37% F-score = 77.48%
[28]	2024	CapsNet	27,558 images	Detection rate = 99% Accuracy = 99.08% FAR = 0.97%
[29]	2019	U-Net	30 images	F1-score = 92.97% PPV = 97.15% Sensitivity = 89.57% Accuracy= 90.9%
[30]	2020	Bilateral Filtering+Image Processing	27,558 images	Precision = 92.97% Specificity = 97.15% Recall = 89.57% Accuracy = 90.9% F1-score = 91.53%
[31]	2022	CNN	27,558 images	Accuracy = 97%
[32]	2023	YOLOv5x	2571 images	Precision = 92.10% Recall = 93.50% F-score = 92.79% mAP0.5 = 94.40%
[33]	2023	Semantic Segmentation CNN	80,000 cells	Accuracy = 99.66%
[34]	2020	RBCNet (U-Net+Faster R-CNN)	965 images	Accuracy = 97% F1-Measure = 97.76% Precision = 97.51% Recall = 98.07%
[35]	2021	CNN	27,558 images	Specificity = 97.78% Sensitivity = 96.33% Precision =96.82% Accuracy = 96.82% F1-Score=96.82%
[36]	2024	Modifed YOLOv4	A=210 images, B=472 images	mAP=90.07%
[37]	2024	DL Approach	MP-IDB = 229 images, IML-Malaria = 345 images, MD-2019 = 883 images	Accuracy =99% Accuracy =92% Accuracy=82%

3. MATERIALS AND METHODS

The research methodology (Figure 2) outlines the systematic approach taken in this study to address malaria detection through image segmentation, starting with a thorough problem analysis that identifies key challenges and establishes research objectives, inspired by related recent works. Following this, the methodology details the data collection and pre-processing steps, where images are sourced from Kaggle, a benchmark dataset, resized, normalized and prepared for training, alongside the generation of dummy masks for segmentation. The core of the methodology involves the design and implementation of a U- Net model enhanced with a spatial-attention mechanism and K-anonymity for privacy preservation, culminating in a rigorous evaluation phase that assesses the model's performance using metrics, such as accuracy, precision, recall and F1-score, ensuring its effectiveness in real-world applications.

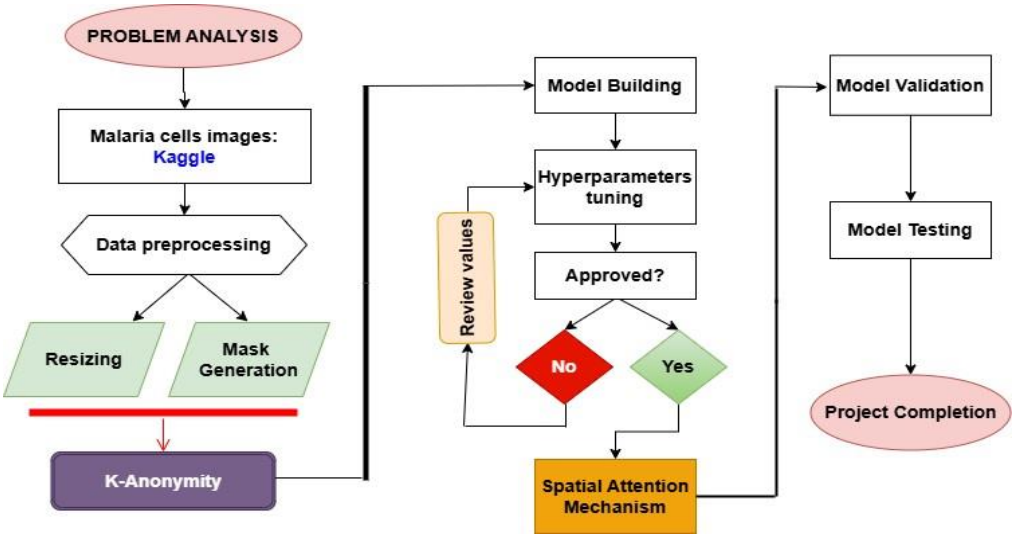


Figure 2. Research methodology overview.

3.1 Dataset

The dataset provided by the National Institutes of Health (NIH) includes publicly accessible images of peripheral blood smears from individuals, featuring both healthy subjects and those diagnosed with malaria. These images were collected at the Lister Hill National Center for Biomedical Communications, using Giemsa-stained blood samples from 150 patients infected with Plasmodium falciparum and 50 healthy individuals. The dataset comprises 27,558 images, with an equal distribution of infected and healthy cells, https://www.kaggle.com/datasets/iarunava/cell-images-for-detecting-malaria, (Figure 3).

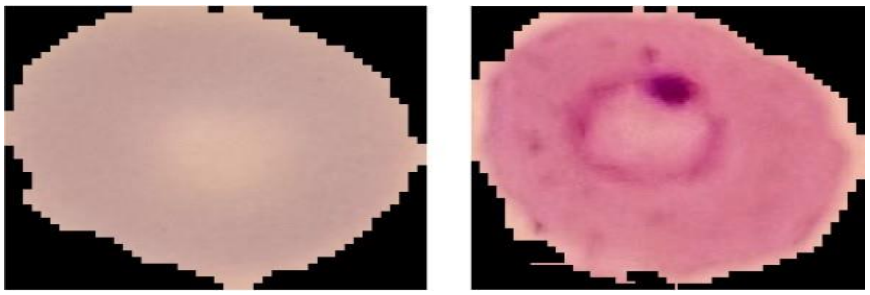


Figure 3. Dataset.

3.2 Data Pre-processing

The dataset used in our study consists of images representing both parasitized and uninfected cells. Each image is resized to a uniform dimension of 64x64 pixels and the pixel values are normalized to a range between 0 and 1, ensuring consistency and comparability in the input data. To facilitate training and evaluation, the dataset is divided into training and testing sub-sets, with 20% of the data reserved for testing to accurately assess the performance of our proposed architecture. For the segmentation task, binary masks are generated using a thresholding approach. The images are first converted into grayscale and a specific threshold is applied to create the masks, which serve as ground truth during the segmentation process.

3.3 K-Anonymity Method

Before delving into the details of the U-Net architecture, it is crucial to discuss the application of the k- anonymity method to the dataset. K-anonymity is a privacy-preserving technique used to protect sensitive information in datasets by ensuring that each record is indistinguishable from at least k-1 other records.

In this study, k-anonymity is applied to the cell images to safeguard the privacy of the individuals represented in the dataset. The function takes the dataset of cell images and a parameter k, which

determines the level of anonymity. It generates random noise with a uniform distribution between -0.1 and 0.1, with the same shape as the input images. The noise is then added to the original images and the resulting values are clipped to ensure that they remain within the valid pixel-value range of 0 to 1.

By adding random noise to the images, the k-anonymity method ensures that any identifying information related to the individuals in the dataset is obscured. This approach effectively "hides" each individual's data among at least k-1 other records, making it difficult for malicious parties to trace sensitive information back to specific individuals.

The application of k-anonymity is critical in medical and biological research, where maintaining the confidentiality of patient data is crucial. By implementing this method, researchers can conduct their analysis while adhering to ethical standards and protecting the privacy of the individuals involved in their studies. After applying k-anonymity, the noisy images are used as input to the U-Net model for the segmentation task. The model architecture is designed to maintain the utility of the data while respecting the privacy constraints imposed by the k-anonymity method.

3.4 Model Architecture

The U-Net model (Figure 4) with spatial attention is constructed to perform image-segmentation tasks effectively. It comprises three main components: the encoder, the bottleneck and the decoder. Each of these components plays a crucial role in processing the input images and generating the output masks.

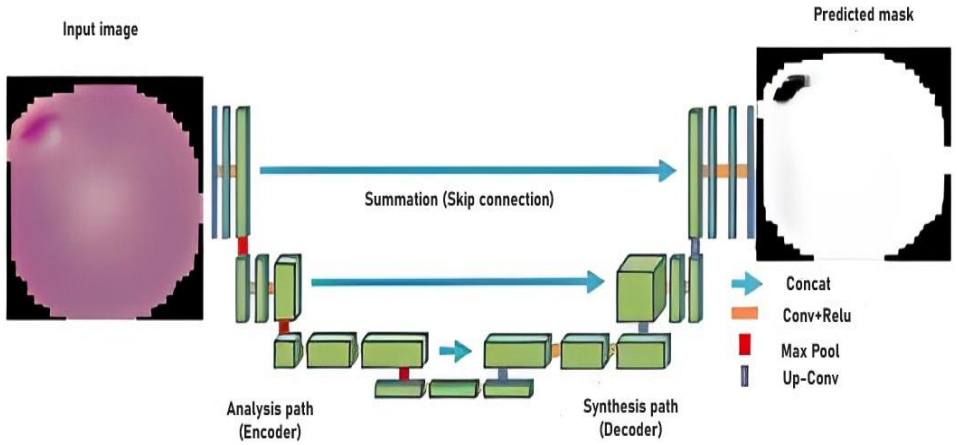


Figure 4. U-Net architecture.

3.4.1 Encoder

The encoder is responsible for capturing features from the input images through a series of convolutional and max-pooling operations. The architecture is structured as follows:

1) Convolutional Layers: The encoder begins with a series of convolutional layers that apply filters to the input images. For instance, the first block consists of two convolutional layers, each with 64 filters (3x3 kernel size) and Rectified Linear Unit (ReLU) activation. These layers help extract low-level features, such as edges and textures.
2) Max-Pooling Layers: Following the convolutional layers, a max-pooling layer (2x2 pooling size) is applied. This operation reduces the spatial dimensions of the feature maps, effectively down-sampling the input and allowing the model to focus on more abstract representations as it progresses deeper into the network.
3) Repeated Blocks: The encoder consists of multiple blocks, each progressively increasing the number of filters (128, 256 and 512) while maintaining the same structure of two convolutional layers followed by a max-pooling layer. This hierarchical structure enables the model to learn increasingly complex features at different resolutions.

3.4.2 Bottleneck

The bottleneck serves as the crucial transition between the encoder and the decoder, effectively

capturing the most salient features from the down-sampled representations. It consists of two deep convolutional layers, each equipped with 1024 filters. These layers operate on the most compressed feature maps, enabling the model to learn high-level representations that encapsulate essential information from the input images. This design allows for a more efficient encoding of the data, ensuring that vital characteristics are preserved for subsequent decoding.

3.4.3 Decoder

The decoder is designed to reconstruct the output segmentation masks from the encoded features. It mirrors the encoder's structure and includes the following components:

1) Up-sampling Layers: Each decoding block begins with an up-sampling layer that increases the spatial dimensions of the feature maps. This step effectively reverses the down-sampling performed in the encoder.
2) Skip Connections: After up-sampling, the decoder concatenates the up-sampled features with the corresponding features from the encoder. This skip connection allows the model to retain spatial information lost during down-sampling, enhancing the reconstruction accuracy.
3) Spatial Attention Layers: The primary objective of the Spatial Attention Mask (SAM) is to create an attention mask that enhances the accuracy of feature extraction from a feature map. This process involves three sequential steps. The first step is down-sampling, where the dimensions of the feature map are reduced. This is achieved by applying average-pooling and max-pooling operations along the channel axis, followed by concatenation of the results to form a compact feature descriptor. Next, this descriptor is processed through a convolutional layer with a 7×7 filter size, using padding to preserve the spatial dimensions. In the final step, the output from the convolutional layer is passed through a sigmoid activation function, which scales the values to a range between 0 and 1, resulting in the attention mask. To enhance the features further, an element-wise multiplication is performed between the original feature map and the generated attention mask, effectively emphasizing the most informative pixels while diminishing less relevant ones (Figure 5).

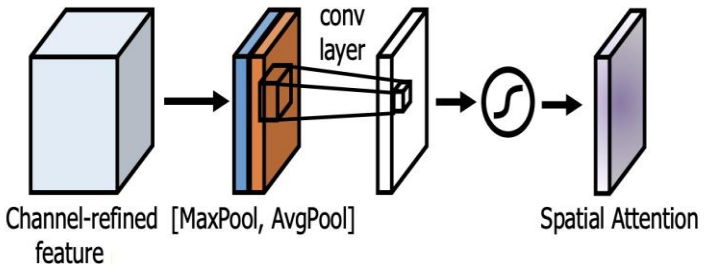


Figure 5. Spatial-attention mechanism.

The attention map highlights the most relevant spatial regions by multiplying them element-wise with the input features, effectively guiding the model to focus on important areas during the reconstruction process.

4) Repeated Blocks: The decoder continues with additional blocks, each consisting of two convolutional layers (with 512, 256 and 128 filters, respectively), followed by an up-sampling layer and a spatial-attention layer. This structure allows for a detailed reconstruction of the output mask.
5) Output Layer: Finally, a convolutional layer with a single filter and sigmoid activation is applied to the output of the last decoding block. This layer generates the final segmentation mask, providing a binary output that indicates the presence or absence of the target features.

Figure 6 illustrates the U-Net model enhanced with a spatial-attention mechanism integrated into the decoding path. The architecture includes an encoder consisting of four convolutional blocks, a central bottleneck block and a decoder with four convolutional blocks. The spatial-attention mechanism is applied after each convolutional layer in the decoder blocks to emphasize important regions of the input images. This mechanism helps in improving the segmentation accuracy by

focusing on the most relevant features in the image. The final output layer generates the segmented output based on the attention- enhanced feature maps.

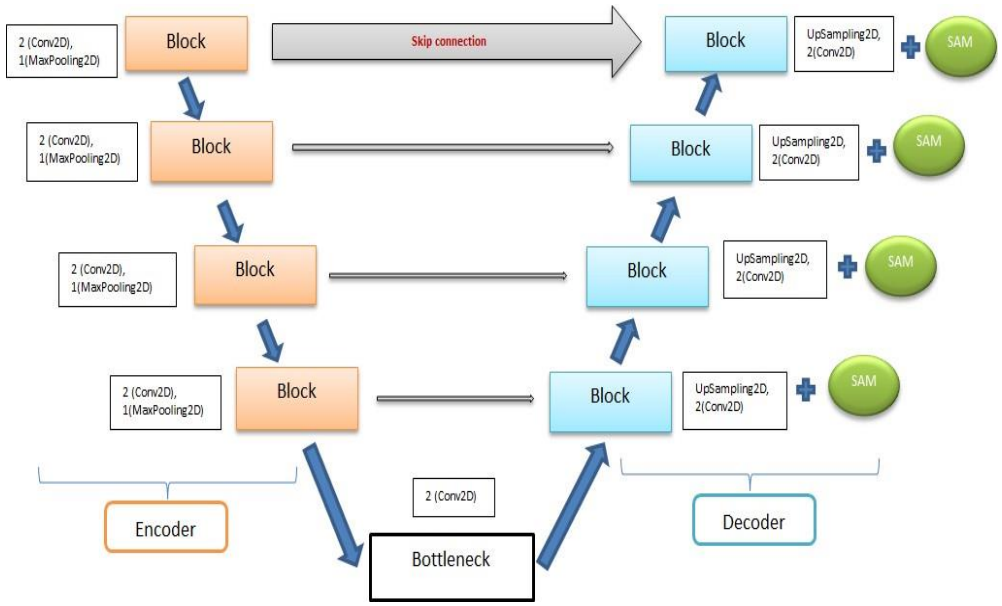


Figure 6. Proposed model: U-Net architecture with spatial-attention mechanism (SAM).

Table 2 summarizes the hyper-parameter values used in the U-Net model for malaria detection. These values were selected after extensive testing and empirical evaluation to optimize the model's performance. The careful tuning of these hyper-parameters is crucial for achieving effective segmentation and ensuring robust results across different datasets.

Table 2. Hyper-parameter summary for our proposed model.

Hyper-parameter	**Value**
Batch Size	16
Learning Rate	0.001
Number of Epochs	10
Convolutional Filter Sizes	
Layer 1	64 filters (3x3)
Layer 2	128 filters(3x3)
Layer 3	256 filters (3x3)
Layer 4	512 filters (3x3)
Bottleneck Layer	1024 filters (3x3)
Activation Function	ReLU
Pooling Dropout Rate	Max-pooling (2x2) 0.2
Spatial-attention Layer	
Convolution Kernel Size	7x7
Number of Filters	1
Activation Function	Sigmoid
K-Anonymity Parameters	
Noise Level	0.1
Number of Neighbors (k)	3

4. Results and Discussion

The segmentation results demonstrate the effectiveness of the proposed approach in accurately identifying and delineating malaria-infected cells from microscopic blood smear images (Figure 7). The input image shows a representative field of view containing both infected and uninfected red blood cells. The true mask serves as the ground truth for infected cell regions. The predicted mask, generated by the segmentation model, closely matches the true mask, indicating a high degree of accuracy in identifying the infected cells. By comparing the predicted segmentation to the true mask,

the model achieves a Dice score of 99.61%, confirming its ability to locate and segment malaria-infected cells precisely. These results suggest that the proposed approach can serve as a reliable tool for automated malaria diagnosis, potentially aiding healthcare workers in resource-limited settings, where access to expert microscopy is limited.

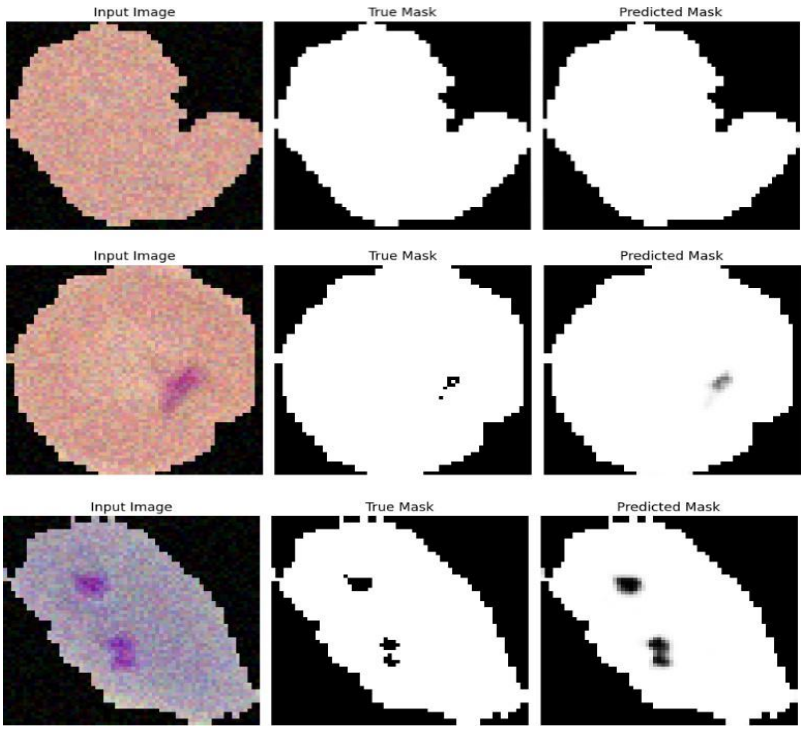


Figure 7. Segmentation results: Input image with k-anonymity, true mask and predicted mask.

The accuracy and loss curves (Figures 8, 9) demonstrate the model's performance over training epochs. The accuracy curve shows a consistent upward trend, indicating that the model is effectively learning and improving its predictions. Simultaneously, the loss curve exhibits a downward trajectory, reflecting a reduction in prediction error. Together, these curves suggest that the model is successfully optimizing its performance, achieving high accuracy while minimizing loss. This correlation between accuracy and loss indicates a well-trained model capable of generalizing well to unseen data.

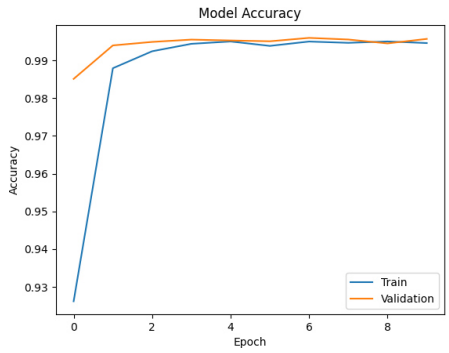


Figure 8. Accuracy curve.

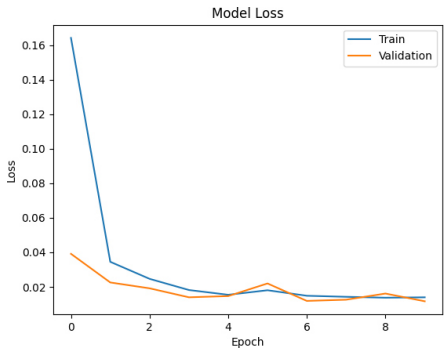


Figure 9. Loss curve.

This study introduces several significant contributions to the field of medical-image segmentation, specifically for malaria-cell detection. Our approach integrates multiple advanced techniques to enhance both the accuracy of segmentation and the privacy of sensitive data.

4.1 Enhanced U-Net Architecture with Spatial-attention Mechanism

One of the primary contributions of this study is the development of a U-Net architecture enhanced with a spatial-attention mechanism. The incorporation of this mechanism enables the model to focus on the most relevant regions of the input images, improving the extraction of critical features and

thus enhancing segmentation performance. This attention mechanism allows the network to dynamically emphasize important areas while suppressing less relevant information, leading to more precise and reliable segmentation results. The empirical performance improvements observed validate the effectiveness of this approach in addressing complex medical-imaging tasks.

4.2 Privacy Preservation through K-Anonymity

In addition to architectural improvements, we have incorporated k-anonymity to address privacy concerns associated with medical data. By adding random noise to the images, k-anonymity ensures that individual identities remain confidential, effectively protecting patient privacy while maintaining the integrity of the dataset. This technique demonstrates our commitment to ethical considerations in data handling and supports the deployment of ML models in sensitive healthcare environments without compromising data security.

4.3 Generalization Capability

In recent studies, the generalization capability of deep-learning models has been increasingly recognized as a crucial factor in their effectiveness across various domains [39]. This sub-section discusses the successful application of a model initially trained on malaria-cell images to a newly introduced real Cactus-disease dataset, which has not yet been published. The dataset comprises 343 images of healthy cacti and 285 images of diseased cacti. The results demonstrate the model's ability to generalize and perform effectively in the agricultural domain.

The model architecture originally designed for analyzing malaria cell images was adapted to address the challenges posed by the Cactus-disease dataset. The U-Net model, known for its efficacy in image-segmentation tasks, was employed to identify and classify diseases affecting cactus plants. The training process involved using the Cactus-disease dataset included a diverse set of images, depicting healthy and diseased cactus specimens.

The successful application of the model to the Cactus-disease dataset illustrates its generalization capability, showcasing how insights gained from one biological domain (malaria-cell images) can be effectively transferred to another (cactus diseases). This cross-domain application not only highlights the versatility of DL models, but also emphasizes their potential in agricultural practices, where timely and accurate disease detection is critical for crop management (Figure 10).

Figure 10. Segmentation results for the second dataset in agriculture domain.

The evaluation metrics for the Cactus-disease detection model are as follows: the F1-score is 98.44%, indicating a strong balance between precision and recall; the Dice score is 95.08%, reflecting the model's accuracy in segmenting the diseased areas; the Precision is 98.04%, demonstrating the model's effectiveness in minimizing false positives; and the Recall is 98.86%, highlighting the model's ability to identify diseased cacti correctly. These results underscore the model's high performance in accurately detecting cactus diseases (Figure 11).

4.4 Comprehensive Evaluation and Validation

The study includes a thorough evaluation of the model's performance using a range of metrics, such as accuracy, precision, recall and F1-score. This comprehensive assessment provides a robust understanding of the model's capabilities and potential for real-world applications. The positive results from these evaluations further strengthen the validity of our proposed method.

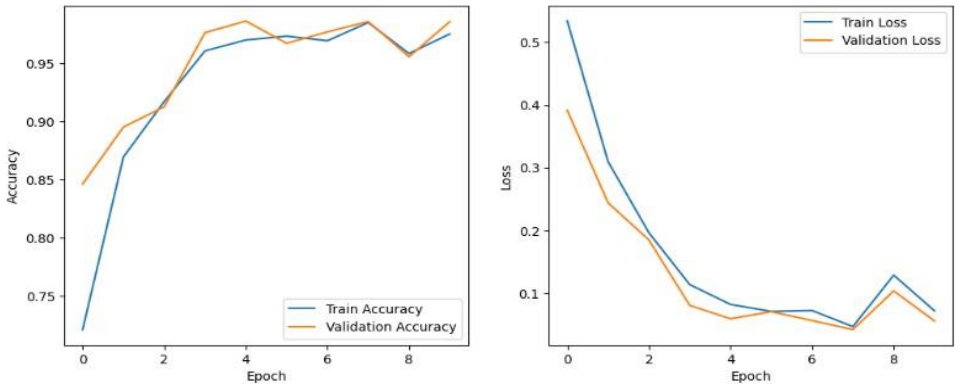


Figure 11. Accuracy and loss curves for Cactus dataset.

Table 3. Comparison of different methods based on performance, privacy and generalization.

Method	Dataset	Performance	Privacy	Generalization
Decision Tree	27,558 images	Accuracy=77.32% Precision = 77.31% Recall = 77.37% F-score = 77.48%	x	x
CapsNet	27,558 images	Detection rate = 99% Accuracy = 99.08% FAR = 0.97%	x	x
U-Net	30 images	F1-score = 92.97 PPV = 97.15% Sensitivity = 89.57% Accuracy = 90.9%	x	x
Bilateral Filtering + Image Processing	27,558 images	Precision = 92.97% Specificity = 97.15% Recall = 89.57% Accuracy = 90.9% F1-score = 91.53%	x	x
CNN	27,558 images	Accuracy = 97%	x	x
YOLOv5x	2,571 images	Precision = 92.10% Recall = 93.50% F-score = 92.79% mAP0.5 = 94.40%	x	x
Semantic Segmentation CNN	80,000 cells	Accuracy = 99.66%	x	x
U-Net + Faster R- CNN	965 images	Accuracy = 97% F1-Measure = 97.76% Precision = 97.51% Recall = 98.07%	x	x
CNN	27,558 images	Specificity = 97.78% Sensitivity = 96.33% Precision =96.82% Accuracy = 96.82% F1-Score=96.82%	x	x
Modified YOLOv4	A=210 images, B=472 images	mAP=90.07%	x	x
DL Approach	MP-IDB=229 images, IML-Malaria=345 images, MD-2019= 883 images	Accuracy =99% Accuracy =92% Accuracy=82%	x	x
Proposed Method	**27,558 images**	**Accuracy = 99.60%** **Dice Score = 99.61%** **Precision = 99.42%** **Recall = 99.96%** **F1-score = 99.69%**	√	√

4.4.1 Limitations

In our work on Kaggle, we have encountered limitations related to computational resources, particularly when using central-processing units (CPUs) for our tasks. Although Kaggle provides a robust environment with access to graphical-processing units (GPUs), reliance on CPUs can result in slower processing times, hindering our ability to efficiently test and iterate on models. The platform's current restrictions, such as the cap on GPU usage at thirty hours per week, also require us to strategically manage our computational tasks to maximize efficiency. This necessitates optimizing our code and pre-processing steps to alleviate the burden on CPU resources. Consequently, while we can leverage Kaggle's capabilities for our projects, these resource limitations demand careful planning and execution to achieve our objectives effectively.

4.4.2 Advantages and Disadvantages

The integration of our proposed approach combines advanced malaria-detection techniques, resulting in several significant advantages. First, the model achieves a high diagnostic accuracy, with a validation accuracy of 99.60% and a Dice score of 99.61%, ensuring reliable malaria detection. Furthermore, it improves data privacy by using the k-anonymity technique, which protects sensitive medical information and minimizes the risk of re-identification, thus ensuring compliance with privacy regulations. Moreover, the approach exhibits improved segmentation performance through a customized spatial-attention mechanism that focuses on critical image features, leading to better results. It also demonstrates a strong generalization ability, achieving an accuracy of 98.58% on the Cactus dataset, a real-world image dataset, indicating its adaptability beyond malaria detection. Automating diagnostic processes reduces the workload of healthcare professionals by streamlining the analysis of cellular images, enabling more efficient patient care. Finally, early and accurate diagnosis through this approach can significantly reduce the risk of serious health complications associated with malaria, contributing to improved health outcomes.

One notable disadvantage of this study is the manual selection of hyper-parameters for the U-Net model, which demands significant time and effort to identify optimal values. This process involves extensive testing and experimentation, making it time-consuming and potentially subjective. Consequently, the chosen hyper-parameters may not represent the best possible configuration for all datasets.

5. CONCLUSION

In this study, we presented a novel approach to malaria detection through cell-image segmentation using a U-Net architecture enhanced with a custom spatial-attention mechanism and K-anonymity for privacy preservation. The model demonstrated an exceptional performance on the malaria-cell image dataset, achieving a high validation accuracy, as well as high Dice score and precision. These results underscore the effectiveness of integrating advanced neural-network architectures with privacy-preserving techniques, addressing both the need for accurate disease detection and safeguarding sensitive patient data.

Furthermore, the model's generalization capability was validated through its application to a real Cactus-disease dataset, where it maintained a strong performance. This indicates the potential for cross-domain applications of the model, paving the way for its use in various agricultural and medical-imaging tasks. While the study achieved significant results, it also highlighted the challenges associated with manual hyper-parameter selection, which can be time-consuming and subjective. Future work should focus on automating this process to enhance model efficiency and adaptability.

Overall, this research contributes to the development of secure and reliable diagnostic tools in both medical and agricultural fields, promoting the integration of privacy-aware methodologies in machine-learning applications.

Future research could focus on several approaches to further enhance the effectiveness and applicability of the proposed model. For instance, employing automated hyper-parameter optimization methods, like grid search or Bayesian optimization, could refine the tuning process and boost the model performance.

REFERENCES

[1] E. O. Kolawole et al., "Malaria Endemicity in Sub-Saharan Africa: Past and Present Issues in Public Health," Microbes and Infectious Diseases, vol. 4, no. 1, pp. 242-251, 2023.

[2] J. Li et al., "Current Status of Malaria Control and Elimination in Africa: Epidemiology, Diagnosis, Treatment, Progress and Challenges," Journal of Epidemiology and Global Health, vol. 14, no. 3, pp. 561-579, DOI: 10.1007/s44197-024-00228-2, 2024.

[3] P. Venkatesan, "The 2023 WHO World Malaria Report," The Lancet Microbe, vol. 5, no. 3, p. e214, 2024.

[4] A. Mbanefo and N. Kumar, "Evaluation of Malaria Diagnostic Methods As a Key for Successful Control and Elimination Programs," Tropical Medicine and Infectious Disease, vol. 5, no. 2, p. 102, 2020.

[5] P. Gupta, "Rapid Diagnostic Tests for Malaria: Challenges and Future Prospects, a Brief Review," Challenges and Advances in Pharmaceutical Research, vol. 8, pp. 152-62, 2022.

[6] O. O. Oyegoke et al., "Malaria Diagnostic Methods with the Elimination Goal in View," Parasitology Research, vol. 121, no. 7, pp. 1867-1885, 2022.

[7] S. Minaee et al., "Image Segmentation Using Deep Learning: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3523-3542, 2021.

[8] M. Z. Khan et al., "Deep Neural Architectures for Medical Image Semantic Segmentation," IEEE Access, vol. 9, pp. 83002-83024, DOI: 10.1109/ACCESS.2021.3086530, 2021.

[9] D. D. Patil and S. G. Deore, "Medical Image Segmentation: A Review," International Journal of Computer Science and Mobile Computing, vol. 2, no. 1, pp. 22-27, 2013.

[10] I. Jdey et al., "Fuzzy Fusion System for Radar Target Recognition," International Journal of Computer Applications & Information Technology, vol. 1, no. 3, pp. 136-142, 2012.

[11] G. Hcini et al., "HSV-Net: A Custom CNN for Malaria Detection with Enhanced Color Representation," Proc. of the 22th IEEE Int. Conf. on Cyberworlds (CW), pp. 337-340, Sousse, Tunisia, 2023.

[12] G. H Hcini, I. Jdey and H. Ltifi, "Improving Malaria Detection Using L1 Regularization Neural Network," JUCS: Journal of Universal Computer Science, vol. 28, no. 10, pp. 1087-1107, 2022.

[13] M. A. Abdou, "Literature Review: Efficient Deep Neural Networks Techniques for Medical Image Analysis," Neural Computing and Applications, vol. 34, no. 8, pp. 5791-5812, 2022.

[14] Tobias Mourier et al., "The Genome of the Zoonotic Malaria Parasite Plasmodium Simium Reveals Adaptations to Host Switching," BMC Biology, vol. 19, p. 219, pp. 1-17, 2021.

[15] D. Sukumarran et al., "Machine and Deep Learning Methods in Identifying Malaria through Microscopic Blood Smear: A Systematic Review," Engineering Applications of Artificial Intelligence, vol. 133, no. E, p. 108529, 2024.

[16] Y. Lv et al., "Attention Guided U-Net with Atrous Convolution for Accurate Retinal Vessels Segmentation," IEEE Access, no. 8, pp. 32826-32839, DOI: 10.1109/ACCESS.2020.2974027, 2020.

[17] G. Du et al., "Medical Image Segmentation Based on U-Net: A Review," Journal of Imaging Science & Technology, vol. 64, Article ID: jist0710, 2020.

[18] R. Azad et al., "Medical Image Segmentation Review: The Success of U-Net," IEEE Trans. on Pattern Analysis and Machine Intelligence, Early Access, pp. 1-20, DOI: 10.1109/TPAMI.2024.3435571, 2024.

[19] A. Ben Hamida et al., "Deep Learning for Colon Cancer Histopathological Images Analysis," Computers in Biology and Medicine, vol. 136, p. 104730, 2021.

[20] Z. Cheng, A. Qu and X. He, "Contour-aware Semantic Segmentation Network with Spatial Attention Mechanism for Medical Image," The Visual Computer, vol. 38, no. 3, pp. 749-762, 2022.

[21] Z. Chen et al., "An Object Detection Network Based on YOLOv4 and Improved Spatial Attention Mechanism," Journal of Intelligent & Fuzzy Systems, vol. 42, no. 3, pp. 2359-2368, 2022.

[22] V. D. Karalis, "The Integration of Artificial Intelligence into Clinical Practice," Applied Biosciences, vol. 3, no. 1, pp. 14-44, 2024.

[23] I. Jdey, "Trusted Smart Irrigation System Based on Fuzzy IoT and Blockchain," Proc. of the Int. Conf. on Service-oriented Computing (ICSOC 2022), pp. 154–165, Sevilla, Spain, 2022.

[24] C. N. Sowmyarani et al., "Enhanced k-Anonymity Model Based on Clustering to Overcome Temporal Attack in Privacy Preserving Data Publishing," Proc. of the 2022 IEEE Int. Conf. on Electronics, Computing and Comm. Techn. (CONECCT), DOI: 10.1109/CONECCT55679.2022.9865682, Bangalore, India, 2022.

[25] A. Majeed and S. Lee, "Anonymization Techniques for Privacy Preserving Data Publishing: A Comprehensive Survey," IEEE Access, vol. 9, pp. 8512-8545, 2020.

[26] K. El Emam and F. K. Dankar, "Protecting Privacy Using K-Anonymity," Journal of the American Medical Informatics Association, vol. 15, no. 5, pp. 627-637, 2008.

[27] N. Rismayanti, "Segmentation and Feature Extraction for Malaria Detection in Blood Smears," International Journal of Artificial Intelligence in Medical Issues, vol. 2, no. 1, pp. 18-29, 2024.

[28] S. Aanjan Kumar et al., "Application of Hybrid Capsule Network Model for Malaria Parasite Detection on Microscopic Blood Smear Images," Multimedia Tools and Applications, vol. 2024, DOI: 10.1007/s11042-024-19062-6, 2024.
[29] J. B. Abraham, "Malaria Parasite Segmentation Using U-Net: Comparative Study of Loss Functions," Communications in Science and Technology, vol. 4, no. 2, pp. 57-62, 2019.
[30] F. Tehreem and M. Shahid Farid, "Automatic Detection of Plasmodium Parasites from Microscopic Blood Images," Journal of Parasitic Diseases, vol. 44, no. 1, pp. 69-78, 2020.
[31] A. H. Alharbi et al., "Detection of Peripheral Malarial Parasites in Blood Smears Using Deep Learning Models," Computational Intelligence and Neuroscience, vol. 2022, no. 1, p. 3922763, 2022.
[32] C. R Maturana et al., "iMAGING: A Novel Automated System for Malaria Diagnosis by Using Artificial Intelligence Tools and a Universal Low-cost Robotized Microscope," Frontiers in Microbiology, vol. 14, p. 1240936, DOI: 10.3389/fmicb.2023.1240936, 2023.
[33] N. Wojtas et al., "Malaria Detection Using Custom Semantic Segmentation Neural Network Architecture," Medycyna Weterynaryjna, vol. 79, no. 8, pp. 406-412, 2023.
[34] Y. M. Kassim et al., "Clustering-based Dual Deep Learning Architecture for Detecting Red Blood Cells in Malaria Diagnostic Smears," IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 5, pp. 1735-1746, 2020.
[35] A. Maqsood et al., "Deep Malaria Parasite Detection in Thin Blood Smear Microscopic Images," Applied Sciences, vol. 11, no. 5, p. 2284, 2021.
[36] D. Sukumarran et al., "An Optimised YOLOv4 Deep Learning Model for Efficient Malarial Cell Detection in Thin Blood Smear Images," Parasites & Vectors, vol. 17, no. 1, p. 188, 2024.
[37] H. A. H. Chaudhry et al., "A Lightweight Deep Learning Architecture for Malaria Parasite-type Classification and Life Cycle Stage Detection," Neural Computing and Applications, vol. 36, pp. 19795-19805, 2024.
[38] G. Hcini et al., "Investigating Deep Learning for Early Detection and Decision-making in Alzheimer's Disease: A Comprehensive Review," Neural Processing Letters, vol. 56, Article no. 153, 2024.
[39] J. Wang et al., "Generalizing to Unseen Domains: A Survey on Domain Generalization," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 8, pp. 8052-8072, 2022.

ملخص البحث:

يُعدّ الكشْف عن الملاريا عن طريق تحليل صور الخلايا أمراً أساسياً للتّشخيص المبكّر والعلاج الفعّال؛ نظراً لأنّ الكشف في الوقت المناسب يُمكنُهُ أن يخفّض من مخاطر حُدوث مُضاعفاتٍ خطرة. إلا أنّ هذه المسألة تنطوي على تحدّياتٍ ترتبط بالخصوصية بسبب حساسية المعلومات الطبية.

تقدّم هذه الدّراسة نموذجاً مبتكراً يراعي الخُصوصية ويتجنّب الكشف عن هوية المريض؛ من أجل تحسين الأمان جنباً إلى جنبٍ مع الحفاظ على دقّة عالية. ويعمل النّموذج المقترح بآليةٍ تضْمن تحسينَ أداءِ تجْزِئة الصُّوَر، ويعتمد على تقنياتٍ متقدّمةٍ للتّركيز على السِّمات الحاسمة دون سواها. أمّا عدمُ الكشف عن الهوية فيتضمّن الحفاظَ على سِرّيّة المعلومات الحسّاسَة.

ولدى تقييم النّموذج المقترح بناءً على مجموعةٍ من مؤشّرات الأداء، أبْدى النّموذج نتائج جيّدة عند تجريبه على صُوَرِ خلايا الملاريا بدقّةٍ بلغت 99.60%. وعند تطبيقِ النّموذجِ على مجموعةِ بياناتِ نَباتِ الصّبّار، بلغت دقّته 98.58%، الأمر الّذي يعني إمكانية تعميم النّموذج لِيُطبَّق في مجالاتٍ متعدّدةٍ. وتشيرُ نتائجُ التّقييم إلى أنّ النّموذجَ المقترحَ يَجمع بين تحسينِ الأمانِ وارتفاعِ مُستوى الأداءِ في تطبيقاتٍ متنوّعةٍ لتحليلِ الصُّوَر.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

BLOCKCHAIN-BASED DEVICE AUTHENTICATION IN EDGE COMPUTING USING QUANTUM APPROACH

Vinayak A. Telsang[1], Mahabaleshwar S. Kakkasageri[2] and Anil D. Devangavi[3]

(Received: 26-Aug.-2024, Revised: 25-Oct.-2024, Accepted: 28-Oct.-2024)

ABSTRACT

The Internet of things (IoT) emerged as a new technology, where everything is connected. Large amounts of data need to be stored for processing; hence, edge computing can reduce the storage of data in a distributed environment, which enhances processing speed and low usage of bandwidth. With an ever- increasing use of IoT devices, issues such as authentication of devices, privacy of data stored and integrity of data have also increased. The authentication of devices is a major concern for edge-connected IoT devices. The problem was solved by using classical cryptographic algorithms such as Elliptic Curve Cryptography (ECC), Rivest-Shamir-Adleman (RSA) and Diffie-Hellman (DH) for message encryption by using public and private keys that need to be stored. These keys need to be stored on a server for device authentication. In device authentication, storing many keys leads to more computation and storage costs and to an increase in delay. With quantum computing and quantum algorithms, such as Shor's and Grover's, it becomes easy to break the keys of cryptographic algorithms, making the system vulnerable. The proposed work Blockchain-based Device Authentication in Edge Computing Using Quantum Approach (BDAEC-QA) provides authentication for IoT devices using context information, quantum key distribution (QKD) and blockchain. The proposed scheme uses the smart contracts to store an information of the IoT devices on the server side, which is used by blockchain to provide secure authentication between the edge server and the IoT devices. The proposed scheme also provides communication between IoT devices across the network. The proposed work is compared with "Lightweight Two-factor-based User Authentication Protocol for IoT-Enabled Healthcare Ecosystem in Quantum Computing" (LTBA) and "A Blockchain-based Mutual Authentication Scheme for Collaborative Edge Computing" (BBMA) and has less registration, key generation and authentication delay, respectively. The BDAEC-QA scheme uses less computation and storage costs as compared with other existing schemes. The proposed scheme is simulated using the AVISPA tool, to provide the security proofs and analysis that indicate that the BDAEC-QA scheme is resistant to well-known attacks.

KEYWORDS

IoT, 5G, Data security, Cryptic algorithms, Blockchain.

1. INTRODUCTION

The traditional cloud-computing technology is a centralized server that allows users to access resources as and when needed [1]. But, centralized computing technology suffers from denial of service (DoS) and distributed denial-of-service (DDoS) attacks. With the growth of connected devices, cloud computing suffers from latency, quality of service (QoS) and time delays. Hence, edge computing emerged as a new alternative to compute, store and process data at the edge of the network [2]. In the era of the internet of everything (IoE) and the advent of industry 4.0, edge-computing technology has become very popular in the Industrial Internet of Things (IIoT) [3]. In IIoT devices, such as sensors, mobile phones, …etc. are connected to an edge server where computation occurs, reducing transmission time and network traffic and improving QoS [4].

With the rapid development of the IoT, the security and privacy issues of IoT devices are issues of concern [5]. Issues, such as authentication, confidentiality and integrity, need to be addressed. The authentication of connected devices is a real challenge, because if the authentication of the connected device does not occur in the network, it leads to leakage of sensitive data. To solve this problem, many authors have proposed schemes that are based on classical cryptosystems, using public and private keys. The private key is used for encrypting the data, while the public key is used for decrypting the

1. V. A. Telsang is with Biluru Gurubasava Mahaswamiji Institute of Technology, Department of Computer Science and Engineering, Mudhol-587317, Visvesvaraya Technological University, Belagavi, Karnataka, India. Email: v.telsang@gmail.com
2. M. S. Kakkasageri is with Basaveshwar Engineering College, Department of Electronics and Communication Engineering, Bagalkote587102, Visvesvaraya Technological University, Belagavi, Karnataka, India. Email: mahabalesh_sk@yahoo.co.in
3. A. D. Devangavi is with Basaveshwar Engineering College, Department of Artificial Intelligence and Machine Learning, Bagalkote587102, Visvesvaraya Technological University, Belagavi, Karnataka, India. Email: anildevangavi_s@yahoo.co.in

data and *vice versa*. Other techniques include hashing of the data by using public or private keys [6], one-time password [7], secret key mechanisms [8], biometric verification [9], third-party servers for key generation, distribution and verification [10]. All such cryptosystems are guaranteed by the hardness of the discrete logarithmic problem that they have adopted. But, if any advanced system overcomes the hardness employed by those cryptosystems, then the system will be compromised.

With the growth of quantum computing, solving the hardness of mathematical problems of traditional cryptographic algorithms will not be an issue. The security of such a cryptosystem, which is guaranteed by the keys used by classical cryptographic algorithms, may become easy to crack [11]. Since the use of quantum algorithms, such as Shor's and Grover's algorithms which can break the cryptosystems in the near future. So one should incorporate the principles of Quantum Cryptography (QC), based on photons and their quantum properties, the photons have different quantum states measured at any time, which helps in developing a secure cryptosystem [12]. The cryptosystem, which is developed using QC mechanics, is believed to be more secure and nearly impossible to break [13].

The classical algorithms use keys to be generated and stored on the server for IoT device verification during the authentication phase. The security of keys is again a major concern, because it can lead to impersonation, man in middle and eavesdropping attacks [14]. Since IoT devices are distributed in nature and due to the growth of blockchain technology as a result of its decentralized nature, cryptographic properties and improved reliability, as well as fault tolerance and unforgeability makes it suitable for providing a solution to store data along with those generated keys [15].

Blockchain technology supports peer-to-peer networking; whenever a transaction occurs, it is verified by the node before being added to the blockchain. The blockchain contains a number of blocks and each block contains a set of transactions that are structured in the merkle hash tree. Whenever a new block is added, the previous block's hash value is stored in the new block to create the structure of the entire blockchain, which ensures that data cannot be modified. The node maintains the ledger and is updated whenever transaction occurs. The ledger maintains the number of blocks that are chained together with a hash mechanism by storing information, like time stamp, hash of current block and hash of previous block. The blockchian is implemented through different consensus algorithms, such as: Proof of Work (PoW) in which the node solves mathematical calculation to add to the blockchain. Proof of Stack (PoS) uses cryptocurrency validation to select the node. Byzantine Fault Tolerance (BFT) is used in the network of nodes where nodes exchange messages and reach consensus [16].

Smart contracts are unchangeable or immutable computer codes that carry out terms according to the occurrence of a set of pre-determined events. Leveraging blockchain technology, smart contracts enable trusted transactions and agreements among anonymous entities without the help of a central authority or an additional enforcement mechanism [17].

Quantum cryptography is an area that helps develop the cryptosystem using the rules of quantum mechanics. The quantum mechanics uses the smallest unit called the qubit, which is in two quantum states: 0 or $\{|0\rangle\}$ or 1 or $\{|1\rangle\}$. Quantum cryptography is based on using photons and their qubit properties to develop unbreakable cryptosystems. The photon exists in more than one state simultaneously and the state is changed when measured [18]. Quantum key distribution is a technique used in quantum cryptography, where a stream of photons is used to transmit data. These photons have a property called a spin, which is of 3 types: Horizontal (↔), Vertical (↕) and Diagonal (↗) or (↖). Whenever a message is transferred from party A to party B, A sends the polarized bits by using the randomly chosen bases (+) (×). On receiving the polarized bits, B chooses the basis and calculates the polarized bits. The polarized bits, which are similar to parties A and B, are used as a quantum key between them [19].

1.1 Objectives

The objectives of the proposed BDAEC-QA scheme are as follows:

- To reduce the cryptographic attack by using quantum mechanics by storing context information of IoT devices.
- To enhance the security of IoT device information by storing it at the edge server using the blockchain.
- To reduce computation and storage costs at the edge server by using a quantum key (QKey).

1.2 Contributions

The authentication of IoT devices in an edge-based network is solved through quantum mechanics, which includes three phases: initialization, key generation, distribution and authentication. In addition, a security analysis of the proposed scheme is also discussed.

- The proposed quantum-based authentication scheme identifies IoT devices using their context information and QKey.
- The use of blockchain to store information of IoT devices at the edge server using smart contracts.
- Establishing communication between IoT devices within the vicinity of the edge server and outside the edge server.

The rest of the paper is organized as follows. Related research works are presented in Section 2. The proposed scheme for authentication of IoT devices is presented in Section 3. Simulation and analysis result are discussed in Section 4. The result discussion of the proposed scheme with different schemes is given in Section 5. Section 6 presents a conclusion.

2. RELATED WORKS

The authentication scheme discussed in [20] has initialization phases consisting of system registration and device registration. Each device has an ID (EID) once it registers and the system ID (SID) is provided by system admin and gets the registration token. The token is stored in the blockchain by using a smart contract with information about the SID, EID and device address (EIP) and an authpass is given to each device. Whenever the authentication of a device is requested, it sends the authpass to fog nodes by encrypting the request using its private key; decryption of the request is carried out using the public key of the device at the fog node. The blockchain enabled fog node verifies the EID present in the blockchain as well as the smart contract. If verification is successful, then authentication is successful; otherwise, the device request is rejected. A computation time of 1.06 ms and a power consumption of 7.24 mW are achieved.

The blockchain-based authentication mechanism is discussed in [21]. It uses smart contract to store the user's request. The miner nodes are used to check the smart contracts of IoT devices. The miner node generates the token for the device upon a token-generation request from the device. The token is signed with its private key and sent to the requested device. During the verification phase, the signed token is issued to the blockchain and if verified successfully, authentication is successful. The scheme achieves a communication delay of 1.6 sec and a communication overhead of 3 sec. A post-quantum fuzzy commitment scheme is provided in [22] and used for the healthcare system. Here, the user must register and authenticate herself/himself with the medical server to access the medical data. The medical data is collected and measured using a smart card and biometric data. The verification of the medical data is successful if the extracted value matches the biometric data and the smart card. The system is complex and it becomes difficult for device authentication with more parameters. The work achieves a computational cost of 20 msec.

The quantum communication authentication for drones discussed in [23] uses a database server to store pre-shared private information with both the ground station and the legitimate drones. The private information of the drone, random key and quantum states is encoded with a private key and sent to the ground station. A random key is used, which guarantees the security of the secret messages. The drone and ground stations authenticate themselves through the secret messages. The schemes provide the secure communication by solving information leakage by detecting the probability of attacks as 0.998. According to the hybrid authentication mechanism based on the vehicle-access network scheme discussed in [24], the scheme identifies information and uses a hash function. The vehicle-resource utilization is efficient, since it uses a multi-vehicle task-management model. For messages between 10-80, the scheme achieves an authentication time of 10-45 msec with a loss rate of 15% and a latency of 35 msec. The resource consumption of the scheme can be optimized by using the master node.

The static and mobile IoT devices using certificate-less cryptography provided in [25] elaborated on the key-generation procedures, lightweight key negotiation and mutual authentication for IoT devices between inter-edge and intra-edge servers. The scheme overcomes most security attacks and achieves

an authentication time and a registration time of 0-2.2 msec and 0.2-1.4 sec, respectively (for 10-100 devices) with a CPU usage time of 28%. The multi-party protocol based on lattice-based cryptography discussed in [26] generates a pair of master keys by using the security parameters by the server; i.e., master secret key and master public key. The user who wants to be part of the network has to request the server by sending her/his public key. The server generates an identity for the user by using the master secret key and the user's public key. The scheme is power efficient and secures communication by eliminating public certificates. A power consumption of 40mW and a CPU usage of 40% are achieved by this scheme.

The two-factor authentication scheme for medical server provided in [27] has a server where the user requests registration with a user ID and a password and if the user ID does not exist, the server responds with the smart card, which contains the hashed values of the user information. Whenever the user wants to communicate, he/she can use a smart card along with a user ID and a password. The scheme achieves a communication cost of 320-800 bits, with an execution time of 0.095 msec. The scheme is secure with a session key generated for each user and with the use of two-factor authentication.

The protean authentication scheme based on minimal initialization vectors provided in [28] uses an edge server to store initialization vectors (V). The gateway maintains hardware (H) information for the edge along with initialization vectors. During each authentication cycle, H and V are used by the gateway to generate a random number in each cycle as an authentication key and securely transfer that information to the edge server, making it virtually impossible to arrive at the authentication keys. The key is generated at each cycle, which makes the scheme more secure, but it is resource intensive with a voltage drainage at edge and router occurring for every 4 and 3 hours, respectively.

Lattice-based device to device authentication discussed in [29] uses edge computing and blockchain technology to reduce the computation overhead on IoT devices. The decentralized blockchain is used for public-key management which simplifies key revocation and enhances security. The scheme uses: registration phase, where the IoT device is registered by its edge server and its public keys are added to the blockchain ledger. In the authentication and key-agreement phase, registered IoT devices can authenticate each other and generate a shared session key. The distributed ledger ensures that the edge servers verify the authenticity and validity of public keys of IoT devices. The protocol uses less communication cost as compared with other lattice-based schemes and a storage cost of 1536 bits.

The lattice-based authentication for vehicular communications provided in [30] uses the registration phase, where edge nodes register with the cloud server with public keys stored in the blockchain. The blockchain uses hyper-ledger fabric with smart contract for adding edge node public key. During the authentication phase, the edge nodes mutually authenticate each other using session key. The revocation phase involves the raft consensus algorithm which ensures transaction integrity and ensures that the public keys can be modified by authorized edge nodes. A computation cost of 11,046 μsec and a storage cost of 2112 bits were obtained during the analysis of the scheme.

3. PROPOSED SCHEME

3.1 Network Architecture

The cloud servers are placed far from the IoT devices and moving data for computation requires more time. Despite the cloud server's processing power, time-intensive applications could not be dealt with, since they suffer from latency and bandwidth-consumption problems. The edge server can provide a solution to these problems when used in combination with a cloud server. The general 3-layer edge architecture is shown in Figure 1. It consists of physical devices at the device layer, edge servers at the edge layer and service providers at the cloud layer. The physical devices that are in proximity to the edge server are connected to that edge server for information exchange and computation. The edge server collects the information from IoT devices through edge controllers, analyzes it using emerging technologies implemented as generic capabilities, called Application Programming Interface (API) and provides the result. The edge server also implements algorithms, data-security techniques and machine-learning algorithms for computation, analyzing and storing the results. The edge layer is close to the device layer and is more suitable for time-intensive applications and intelligent processing. Hence, it is more secure and efficient as compared with cloud computing.

3.2 Preliminaries

- $ES = \{ES_1, ES_2, ...ES_n\}$, where ES is the set of edge servers.
- $IoT_{dev} = \{IoTdev_1, IoTdev_2, ...IoT\ dev_n\}$, where IoT_{dev} is the set of IoT devices.
- Each edge server and IoT device have their pair of public and private keys for encryption and decryption, respectively. ES has its pair of keys $\{k_{pues}, k_{pres}\}$ and IoT_{dev} has its pair of keys $\{k_{puit}, k_{prit}\}$.
- Each IoT device is assigned with unique device ID (DID) by the edge server.
- The hash value of the *input* is calculated by using one-way hash function h (*input*).
- The encryption function *Encrpt* (pu_{key}, *message)* is used to encrypt the message by using public key.
- The decryption function *Decrpt* (pr_{key}, *message)* is used to decrypt the message by using private key.
- Context information (*CI*) is the information of the device which consists of its MAC address (*M ACadd*), location information (*locinf o*) and timestamp, as shown in Equation (1).

$$CI = \{M\ ACadd, locinfo, timestamp\} \quad (1)$$

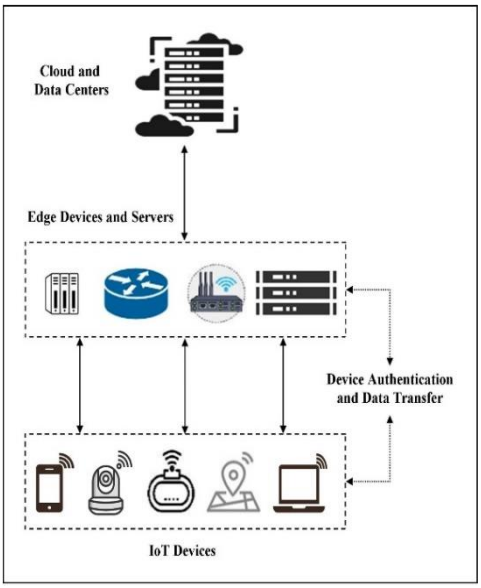


Figure 1. Edge architecture.

3.3 Notations

The notations considered in the proposed BDAEC-QA scheme are listed in Table 1.

Table 1. Notations.

Notations	Description
ES_i, ES_i, ES_m	*Edge server*
IoTdev	*IoT device*
CI	*Context information*
$k_{pues}, k_{puesi}, k_{puesi}, k_{puesm}$	*Public key of ES*
$k_{pres}, k_{presi}, k_{presi}, k_{presm}$	*Private key of ES*
kpuit	*Public key of IoTdev*
kprit	*Private key of IoTdev*
|	*Concatenation operation*
Mreg	*Registration request*
Mpinfo	*Quantum sequence information Device*
DID	*ID of IoTdev*
QKey	*Quantum key*
Sk	*Session key*
Mesv	*Verification message within edge*
Mesoev	*Verification message outside edge*

3.4 Proposed Architecture

The proposed BDAEC-QA scheme for the authentication of IoT devices is shown in Figure 2. It consists of an IoT layer, an edge layer and a blockchain layer. The IoT layer consists of IoT devices that provide context information that needs to be processed by the edge server. The edge layer consists of edge servers, which store and apply computation to generate the device ID and quantum key. It also communicates with the blockchain layer to store the IoT-device information along with the quantum key using a smart contract. The proposed authentication architecture has three phases: registration, key generation and distribution and authentication, as shown in Figure 2. In the registration phase, the edge server broadcasts its public key in the network. IoT device in the vicinity of the edge server uses the edge server's public key to send its context information to the edge server. After registration, the edge server generates the quantum sequence using quantum bits and basis information, as shown Table 2 and sends the quantum-bit information to the IoT device to begin quantum key generation. In the key generation and distribution phase, the IoT device also generates a quantum sequence using quantum bits and choosing a random basis, where the basis information of the IoT device is sent to the edge server. The edge server, upon receiving basis information, matches and extracts the matched sequence number from its quantum sequence and sends only the matched quantum-sequence number to the IoT device to generate a quantum key (QKey) between the respective IoT device and the edge server. Each edge server stores the information of the requested IoT device in the blockchain by creating the markle tree by using the information sent by the device along with the QKey. In the authentication phase, the IoT device sends an authentication request to the edge server. The edge server verifies the authentication request stored in the blockchain. Based on the verification, the IoT device is either authenticated or unauthenticated. The operations involved in the 3 phases are explained in detail below.

Table 2. Quantum-sequence generation.

Bases	1	0
+	↕	↔
×	↗	↖

3.4.1 Registration

- **IoT-device Registration:** The following steps are involved in the registration process of the IoT device to the edge server.
  - Each edge server has its pair of $\{k_{pues}, k_{pres}\}$ keys. Each edge server broadcasts its public key k_{pues} in the network, so that any IoT device can send a registration-request message M_{reg} by using k_{pues}.
  - The IoT device sends an encrypted message M_{reg} to the edge server by using its public key k_{pues}. The registration message sent from the IoT device consists of context information and its public key k_{puit}.

$$M_{reg} = Encrpt(k_{pues}, CI \mid k_{puit}) \quad (2)$$

  - The edge server decrypts the M_{reg} by using its k_{pres} as: Decrpt (k_{pres}, M_{reg}) and gets the context information of the IoT device and its public key; i.e., (CI | k_{puit}).
  - The edge server then generates a unique device ID (DID) for each IoT device and registers it along with its CI information. After registering, the ES generates the quantum sequence, as shown in Table 2 by randomly choosing the quantum bits and basis. The quantum-sequence information (M_{pinfo}) is encrypted and sent to the IoT device by using its key k_{puit} as shown in Equation 3. After sending the information, the ES initiates the quantum-key (QKey) generation and distribution phase.

$$M_{pinfo} = Encrpt(k_{puit}, sequence(\updownarrow, \leftrightarrow, \nearrow, \nwarrow)) \quad (3)$$

- **Edge-server Registration:** The edge server registers with the cloud server by using the public key of the cloud k_{pucs} and sends the registration message as $M\ es_{es}$=*Encrypt* (k_{pucs}, (ID_{es} | $\boldsymbol{n_1}$)). The cloud server, after receiving the message ($M\ es_{es}$), replays with a session key (Ski) to the edge server to confirm registration, as shown in Equation 4.

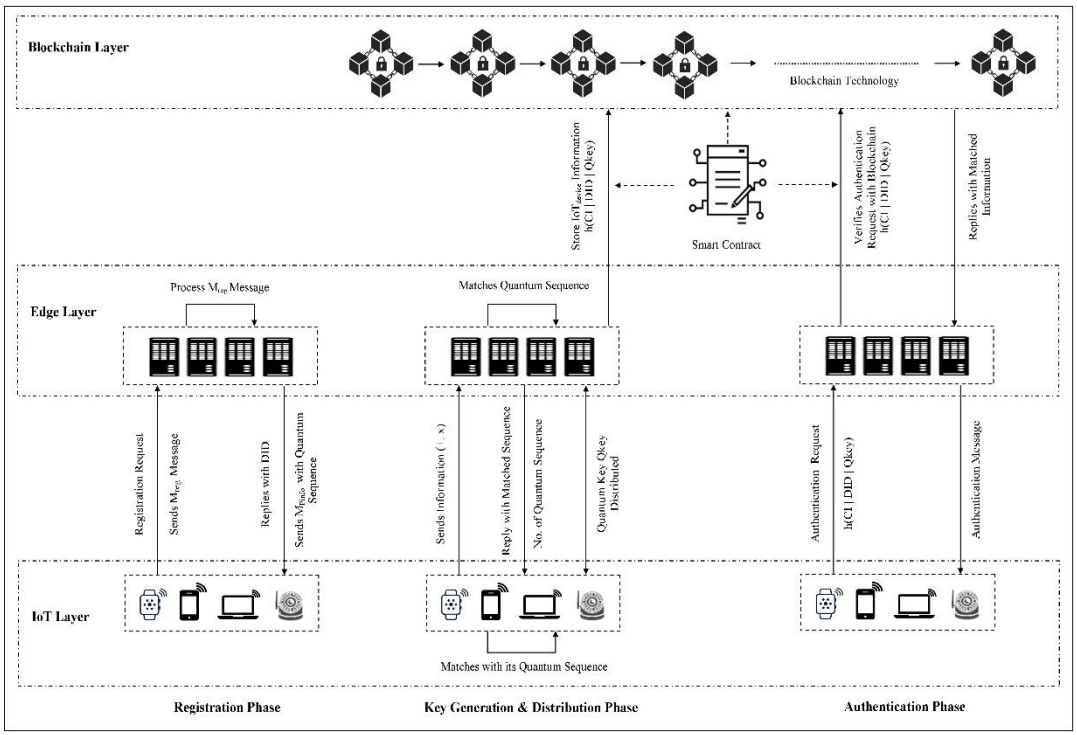


Figure 2. BDAEC-QA IoT-device authentication scheme.

$$Msg_{es} = Encrpt(k_{pues}, (S_{ki} \mid n1)) \tag{4}$$

The edge server then decrypts the message to get Ski as Decrypt (k_{pres}, $M\ sg_{es}$) to complete the registration process.

3.4.2 Key Generation and Distribution

- The IoT device decrypts the message M_{pinfo} using its private key k_{prit} to receive the quantum-bit sequence information as *Decrpt* (k_{prit}, M_{pinfo}). The IoT device uses this quantum-bit sequence and the randomly generated basis to generate the quantum sequence, as shown in Table 2. Equation 5 represents the generated quantum sequence and Equation 6 represents the four states of the qubits used to generate the QKey.

$$|\psi\rangle_{q1,q2,\dots qk} = (|000\dots00\rangle_{q1,q2,\dots qk} + |111\dots11\rangle_{q1,q2,\dots qk})/\sqrt{2} \tag{5}$$

$$\{|0\rangle, |1\rangle, |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, |-\rangle = (|0\rangle - |1\rangle)/\sqrt{2} \tag{6}$$

Algorithm 1. IoT-device Registration

```
Input: Context information(CI), k_pues, k_pres, k_puit
Output: Generating DID for IoT devices
if IoT_dev in the vicinity of ES then
    Send CI to ES as shown in Eq.(1)
end if
while true do
    if ES receives CI then
        for each CI received from IoT_dev do
            Decrypt M_reg as in Eq.(2)
            Generate DID for each IoT_dev
        end for
        for each IoT_dev with DID do
            Send M_pinfo as shown in Eq.(3)
        end for
    end if
endwhile
```

- The IoT device then sends its randomly chosen basis-sequence information to the edge server by encrypting it with using k_{pues} as: *Encrpt* (k_{pues}, *sequence* (+, ×)) for quantum-key mapping, as shown in Table 2.
- The edge server decrypts the basis sequence using *kpres* and gets the information as: *sequence* (+, ×) sent by the IoT device. The ES uses IoT_{dev} basis and matches it with its generated quantum sequence. The ES extracts the quantum-sequence number from a matched pair of ES quantum sequences and the IoT quantum sequences to generate an QKey.
- The ES stores the IoT-device information as: $h\ (CI \mid DID \mid QKey)$ in the blockchain by creating a new block. This information is stored for each requested IoT device separately in the ES; i.e., the information uses the CI, DID and the Qkey of the respective device.
- The ES sends the matched quantum-sequence number to the respective IoT_{dev} device by using its DID. Upon receiving IoT_{dev} it is matched with its quantum sequence to get the QKey.
- After the QKey is generated and distributed such that both IoT_{dev} and ES have an Qkey, which is a unique and symmetric key between each other, respectively. The same key is used for authentication between the respective IoT_{dev} and the ES.

3.4.3 Authentication

- Whenever IoT_{dev} wants to communicate, it should be authenticated. For authentication, IoT_{dev} sends an authentication request to the ES as: *Encrpt* (*QKey, CI* | *DID*).
- The edge server, upon receiving the authentication request from IoT_{dev} decrypts it using the QKey (already obtained in Phase-2) and matches the information stored in the blockchain. If the authentication request matches, then IoT_{dev} is authenticated; otherwise, it is not authenticated.

Algorithm 2. Key Generation and Distribution

Input: Quantum sequence (Qseq) and M_{pinfo}
Output: QKey distribution at both ES and IoT_{dev}
Edge server sends M_{pinfo} to IoT_{dev}
if IoT_{dev} registered **then**
    Generate Qseq and basis
    Send M_{pinfo} to IoT_{dev}
end if
IoT_{dev} decrypts M_{pinfo} and generates Qseq and basis and sends basis information to ES for QKey generation
for each IoT_{dev} registered **do**
    if Qseq of IoT_{dev} = = Qseq of ES **then**
        ES extracts the matched quantum sequence number ($Qseq_{num}$)
    end if
end for
At ES: The matched quantum sequence number is QKey (generated)
ES sends the quantum sequence number information to IoT_{dev}
The IoTdev matches the ES $Qseq_{num}$ to its generated $Qseq_{num}$
for each IoT_{dev} : ES $Qseq_{num}$ **do**
    if Qseq of IoT_{dev} = = ES $Qseq_{num}$ **then**
        IoT_{dev} extracts the matched $Qseq_{num}$
    end if
end for
At IoT device: The matched $Qseq_{num}$ is QKey (distributed)
At ES: Stores the IoT_{dev} information in blockchain using smart contract
for each IoT_{dev} : QKey generated **do**
    create a block information as: $h\ (CI \mid DID \mid QKey)$
    Store the device information in blockchain
end for

3.4.4 Communication of IoT Devices

In this phase, IoT devices want to communicate with other IoT devices within the edge network or outside the network. Edge servers interact with each other to validate the IoT devices. The edge servers also share the registered information with the cloud server for communication outside the edge network. The registered device information is shared with the cloud server by using the Ski along with the context information of the IoT device, as shown in Equation 7.

$$M\,sgreg_{IoT} = Encrpt\,(S_{ki},\, h\,(CI \mid DID \mid QKey) \mid ID_{esi}) \quad (7)$$

Algorithm 3. IoT-device Authentication

```
1: Input: Authentication request (CI | DID, QKey)
2: Output: Authentication message
3: for each Authentication request from IoT_dev do
4:     if Request matches with information stored in blockchain h (CI | DID | QKey) then
5:         Authentication Successful
6:     else
7:         Authentication Unsuccessful
8:     end if
9: end for
```

The cloud server then decrypts the information and updates the registered device information for the respective edge server as: Decrypt (S_{ki}, $M\,sgreg_{IoT}$).

- **Within the Same Edge Network**

Whenever an IoT device moves from one edge server (ES_i) to another edge server (ES_j), then the validity of IoTdev has to be checked to communicate within the network of ES_j. The IoTdev sends the its registered information to ES_j as: $M_{esv} = Encrypt\,(k_{puesj}, M1)$, where $M1 = (CI \mid DID \mid QKey) \mid kpuesi$. The ES_j then decrypts M_{esv} and sends the M1 information to edge server ES_i as: Encrypt (k_{puesi}, $M\,1$). The server ES_i matches M1 with its registered IoT_{dev} information and replies with a message as "valid" to ES_j, then IoT_{dev} can communicate within the ES_j network.

- **Outside the Edge Network**

When the IoT device IoT_{dev} moves from the edge network, the validity of the IoT_{dev} is not verified outside the edge network. When it sends the message to ES_m as: $M\,es_{oev} = Encrypt\,(k_{puesm}, (CI \mid DID \mid QKey) \mid kpuesi)$, ES_m in turn sends the message (cs_{msg}) to the cloud server, as shown in Equation 8.

$$cs_{msg} = Encrypt\,(S_{ki},\, (CI \mid DID \mid QKey) \mid k_{puesi}) \quad (8)$$

The cloud server, after receiving the cs_{msg} decrypts and sends the validity of IoT_{dev} if the information is updated by the edge servers, IoT_{dev} can communicate outside the edge network.

3.5 Case Study Discussion

The QKD-based system uses power, but provides more security while used during key exchange and encryption. Traditional QKD systems use quantum transmitters and receiver components in the network infrastructure, causing more power consumption in large-scale IoT networks. This power hungry nature of the QKD can be optimized by developing quantum hardware, where IoT devices can perform minimum cryptographic operations and quantum operations can be lifted to cloud servers. As the technologies mature and there may be development of chips that can be integrated into low power IoT devices, this makes them consume less power and provide more security with the QKD approach.

Some of the Real World Solutions Using QKD Approaches

The SwissQuantum network testbed deployed in Geneva uses the BB84 protocol for secure communication using QKD [31]. The project shows the feasibility of implementing QKD with regular telecom infrastructure by using quantum encryption. The network guarantees the secure transfer of government and financial data and shows that the QKD can be implemented to provide solutions to real-world problems. Toshiba Europe Ltd. has developed the chip-based Quantum Key Distribution (QKD) system with focus on reducing the size, weight and power consumption of QKD systems, by

integrating them into semiconductor chips [32]. These chips are more power-efficient and can be mass-produced with significantly lower cost. These chips with QKD are used to provide a robust level of security for highly-sensitive data.

The proposed scheme is based on QKD approach with blockchain to authenticate IoT devices. The BDAEC-QA scheme considers the aspect security rather than power consumption at the IoT device. Our simulation results show that the proposed scheme performs better in terms of various delays, but also resists different attacks. With research going on, quantum-based solutions, the QKD approach and PQC can be implemented to provide more security with less network resources.

4. SIMULATION MODEL

This section describes simulation settings, different performance parameters and different security threats applicable to the BDAEC-QA scheme.

4.1 Simulation Settings

The proposed BDAEC-QA scheme is simulated using the Eclipse platform with Java SDK 11. Three edge servers were created and each was registered with three IoT devices. We set the communication distance at 50 metres. We also used the public blockchain to store the device information. The metamask is used to fetch the information from the blockchain in real time during device authentication.

4.2 Performance Analysis

We have simulated BDAEC-QA scheme and compared it with LTBA [22] and BBMA [25] schemes. Different performance parameters mentioned below are analyzed to test the effectiveness of the proposed scheme.

- **Registration delay:** It is the time taken by the IoT device to register to the edge server. It is measured in milliseconds. We observe that from Figure 3, as more devices register for different edge servers randomly, there is an increase in the registration time. The BDAEC- QA scheme used 3 edge servers and devices can register with any edge server. There is liner growth, which shows that the proposed scheme is stable.
- **Key-generation delay:** It is the time taken by the edge server or the IoT device to generate the QKey and is denoted as T_{qk}. It is measured in milliseconds. The key-generation delay of the BDAEC-QA scheme w.r.t the number of devices is shown in Figure 4. The BDAEC-QA scheme key-generation delay is reduced by 14% and 15% than in LTBA and 10% and 11% than in BBMA, when the edge devices considered are 50 and 100, respectively. The BDAEC-QA scheme generates quantum keys using bases and quantum sequence information rather than complex mathematical computations and hence takes less time.
- **Encryption delay:** It is the time taken by the edge server or the IoT device to encrypt the message and is denoted as T_e. It is measured in milliseconds. Figure 5 shows the encryption delay w.r.t key size and the number of devices. The BDAEC-QA scheme encryption delay is decreased by4% and 4.9% than in LTBA and 2.8% and 5.6% than in BBMA respectively. BDAEC- QA uses the quantum key which is a symmetric key and hence the key-generation delay is lower, resulting in a lower encryption delay.
- **Decryption delay:** It is the time taken by the edge server or the IoT device to decrypt the message and is denoted as T_d. It is measured in milliseconds. The BDAEC-QA scheme decryption delay is 5.2% and 10% less than in LTBA and 2.9% and 5.8% less than in BBMA, respectively. Figure 6 shows the decryption delay w.r.t varying key size and the number of devices. Since the key-generation delay is less because BDAEC-QA uses the quantum key which is a symmetric key, this results in a lower decryption delay.
- **Authentication delay:** It is the time taken by the edge server to authenticate the registered IoT device. It is measured in milliseconds. Figure 7 shows the authentication delay of an IoT-device with key size. We observed that BDAEC-QA scheme takes 16% and 7% less authentication time than LLBA and BBMA schemes, respectively.

- **Storage cost:** It is the number of bits required to store the information at the IoT device and the edge server during the operations discussed in the proposed scheme. It is denoted as S_{cost}.
- **Computation cost:** It is the number of bits required to complete the operations discussed in the proposed scheme by the edge server and the IoT device. It is denoted as C_{cost}.

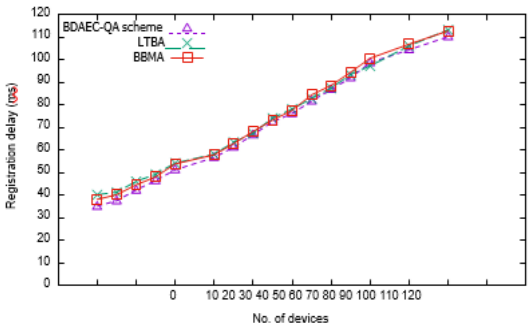


Figure 3. Registration delay.

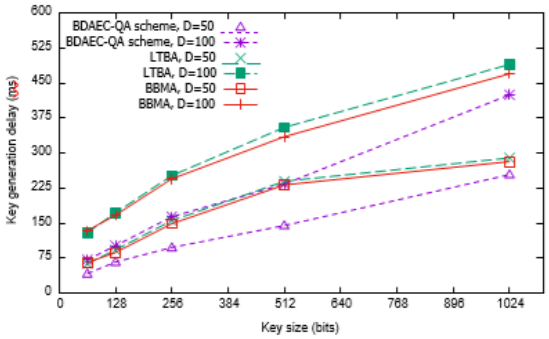


Figure 4. Key-generation delay.

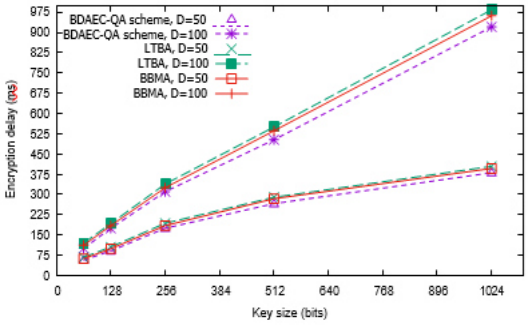


Figure 5. Encryption delay.

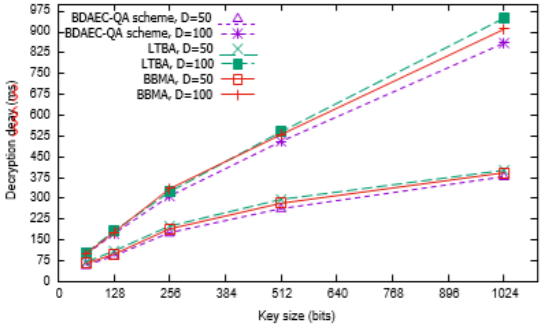


Figure 6. Decryption delay.

4.3 Adversary Model

The Cannetti-Krawczyk (CK) adversary model [33] evaluates the proposed authentication protocol. In the CK model, the adversary is provided with the information about the messages exchanged between authorized parties. The adversary uses the information and impersonates the authorized users. The goal of CK model is to determine the level of security that the protocol should provide and withstand against various attacks. Along with the CK model, additional security requirements are discussed.

- **Usual attacks:** An attacker can steal the information of a device by stealing its identity. If an attacker can impersonate a device, he/she can change the authentication process. The OFMC report provided in Fig. 8 suggest that in the session role, the keys of device and server are made available to the intruder, but still the system is "SAFE" as shown in Fig. 9. In BDAEC- QA scheme, the authentication of IoT devices is carried out by using the context information of the device, QKey, which is stored in the blockchain on the server side. These pieces of information are difficult to steal and hence, the proposed scheme resists to reply and impersonation attacks.
- **Ephemeral Secret Leakage (ESL) attack:** It refers to the preservation of identity of the IoT device privacy. In BDAEC-QA scheme, only registered devices can be authenticated. The confidentiality of the BDAEC-QA scheme is checked by disposing the partial information of the edge device, such as ID, public key and private key. During registration, the context information is passed as Encrpt (k_{pues}, CI $\mid k_{puit}$). Also, during authentication, the IoT device sends the authentication request as: h (CI $\mid DID \mid QKey$), in encrypted format and only IoT devices can manipulate the information. Thus, it guarantees the confidentiality of the proposed BDAEC-QA scheme.
- **Conditional anonymity (CA):** The CI of the IoT device and the private key of the edge server is provided to the attacker, but in the BDAEC-QA scheme, it is not possible to impersonate, since the device information is stored in the blockchain along with its Qkey. The device information is stored in the blockchain with h (CI $\mid DID \mid QKey$); hence, in the proposed scheme, the device is not revealed to any server.

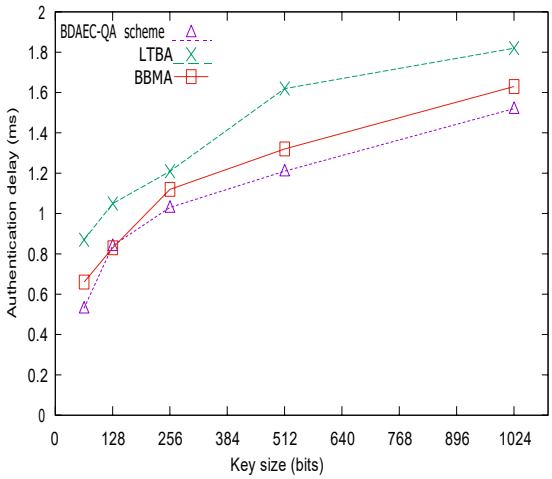


Figure 7. Authentication delay.

- **No Key Escrow:** The proposed BDAEC-QA scheme ensures that the information of the IoT devices is known by storing the hashed information of the IoT devices at the edge server.
- **Integrity:** It refers to the authentication information of an IoT device stored at the edge server, which can not be modified once it is stored unless this done by the IoT device itself. In the BDAEC-QA scheme, once the context information is accepted and the QKey is generated at the edge server, these are stored in the blockchain using a smart contract as h (CI $\mid DID \mid QKey$). Once the information is stored, it cannot be modified, thus guaranteeing the integrity of the proposed scheme.
- **Device capture or (Man-in-middle attack):** A device-capture attack happens when an intruder acquires the information of communicating devices and behaves as an authenticated device. It either steals or alters the data as required, which affects the communication between the devices. The proposed BDAEC-QA scheme prevents such attacks by using QKey and blockchain, because the context information is hashed by using QKey: h (CI $\mid DID \mid QKey$) and can not be easily compromised. Hence, it entrusts the message only to legitimate devices.
- **Resistance eavesdropping attack:** Information leakage is crucial for any authentication protocol; otherwise, an attacker can deduce the message exchanged between an IoT device and an edge server and extract information. In the BDAEC-QA scheme, the public and private keys are used during the initialization phase and QKey is used after the key-generation phase. In proposed scheme, it is not possible to extract the IoT-device information such as: context information, QKey and timestamp as easily. Even if the intruder tries to extract the quantum bits and basis information, the QKey information cannot be found due to the randomness of the QKD protocol.
- **Blockchain-data transfer:** In the BDAEC-QA scheme, we preserve the device data by storing it in the blockchain at the edge server. The blockchain ensures that data cannot be modified once it is stored, by using the previous block's hash value while creating the structure of the blockchain. Hence, the proposed scheme is more secured.
- **Quantum-attack resistance:** The BDAEC-QA scheme uses QKey which is generated by using the quantum bits and basis at the edge server and the IoT devices. It is not possible to detect the state of the quantum key. If there is any modification to the quantum bits, then a new quantum sequence is generated and hence a different QKey, which is detected by our authentication scheme, thereby preventing quantum attacks.

4.4 Security Analysis

The BDAEC-QA scheme is analyzed with fortifcation against different attacks to ensure that the proposed scheme is well protected. Table 4 provides the security properties comparison between the proposed BDAEC-QA scheme and existing schemes. The proposed work is analyzed using the Automated Validation of Internet Security Protocols and Applications (AVISPA) simulation tool in order to verify the authentication protocol against various attacks, including MIM, impersonation, replay and key secrecy [40]. AVISPA is based on High-level Protocol Specification Language

(HLPSL), which is an expressive, modular, role-based, formal language that allows for specification. HLPSL uses the temporal logic of action for specified semantics, converting the latter into operation semantics as an Intermediate Format (IF) and the output is in Output Format (OF). IF specifications are input into the 4 back-end models: On-the-Fly Model Checker (OFMC), Constraint Logic-based Attack Searcher (CL-AtSe), SAT-based Model Checker (SATMC) and Tree Automata based on automatic approximations for analysis of security protocols (TA4SP) used by AVISPA. Figure 8 shows the roles of device, edge server and session role. Figure 9 shows the OFMC report, which performs protocol falsification and bounded verification, CL-AtSe report applies constraint solving and implements redundancy elimination techniques and simplification heuristics and the SATMC report, which represents a violation of the security properties of the protocol. The OFMC and CL-AtSe models use formal verification and if the result of an authentication protocol is safe, then security requirements are met.

The comparison of storage and computation costs with various schemes is provided in Table 3. The T_{0+1}, T_{c+cs}, T_{4+9} use 128-bit computation cost. T_h generates 128-bit hashed output. T_e and T_d also generate 128-bit encrypted and decrypted outputs. T_b uses 20 bits to store the block information. T_{qk} uses 64 bits as key size. The existing schemes [34]-[35],[37]-[38] and [39] use more C_{cost} and [36] uses less C_{cost} compared with the proposed scheme. The S_{cost} is used by the scheme discussed in [34]-[36] and [39] uses more as compared with the proposed scheme.

Table 3. Comparison of computation cost and storage cost with existing schemes.

Schemes	C_{cost} **(in bits)**	S_{cost} **(in bits)**
Multimodal biometric [34]	$T0+1 +Tc+cs +T4+9=512$	$6Mn\phi$=23.3 kb
Remote registration and group authentication [35]	$T_k + 2T_h + 2T_e$=576	$T_k + 2T_h + 2T_d$=576
Lightweight Three-Factor Authentication [36]	$4T_{mp} +2T_{add} +T_h$=320	$2klogk(4k^2log^2k + 4klogk + 7)$=861
Secure user authentication and key agreement [37]	$7T_h + 2T_{e/d}$=1152	-
Secure authentication key exchange [38]	$26T_h +11T_{pm}$=4736	-
Light authentication key agreement [39]	$19T_h$=2432	$3T_h + 3T_{fe} + 3T_d + K_{fe}$=1280
BDAEC-QA Scheme	$3T_e +2T_{qk} +2T_d +T_b$=468	$2T_e + 2T_{qk} + 2T_d + T_b$=532

Table 4. Comparison of security properties with existing schemes.

Security Properties	**[34]**	**[35]**	**[36]**	**[37]**	**BDAEC-QA scheme**
Confidentiality	✓	✓	✓	✓	✓
Integrity	✓	✓	✓	**NA**	✓
Man-in-middle attack	×	✓	✓	✓	✓
Impersonation attack	✓	✓	✓	✓	✓
Anonymity	×	✓	✓	✓	✓
Eavesdropping attack	×	✓	✓	✓	✓
Blockchain data transfer	×	×	✓	**NA**	✓
Quantum attack	**NA**	**NA**	**NA**	**NA**	✓

The quantum-key approach uses symmetric key between the IoT device and the edge server. The symmetric-key exchange is faster as compared with asymmetric-key exchange. The main objective is to implement the quantum mechanics for generating Qkey and the blockchain approach to enhance the security of the proposed scheme. The Qkey is used to exchange messages between edge device and edge server. The CI and QKey of the device are stored in the blockchain with smart contract to provide the extra layer security at the edge server. The blockchain is an immutable ledger which provides the integrity of data stored in it. The proposed protocol is feasible by not only resisting the major attacks, but also by performing better compared with other schemes. The BDAEC-QA scheme is simulated and compared with existing schemes with respect to various delays, computation cost and storage cost and it performs better. The scheme is also validated with the CK adversary model with different attacks and analyzed using AVISPA tool to meet the security requirements.

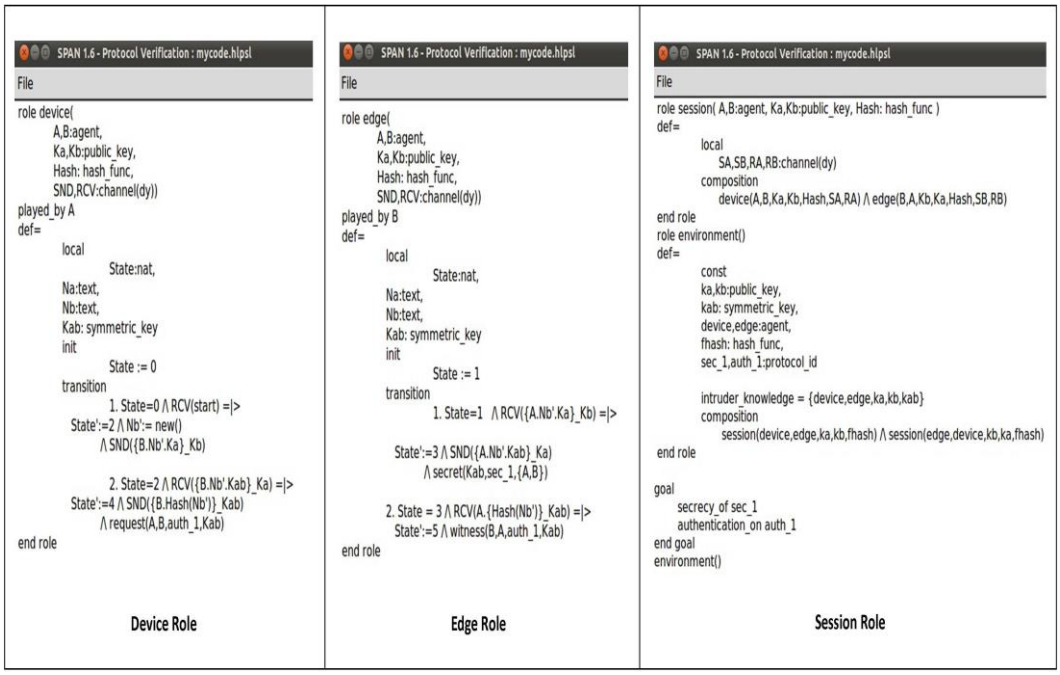


Figure 8. Device, edge and session roles.

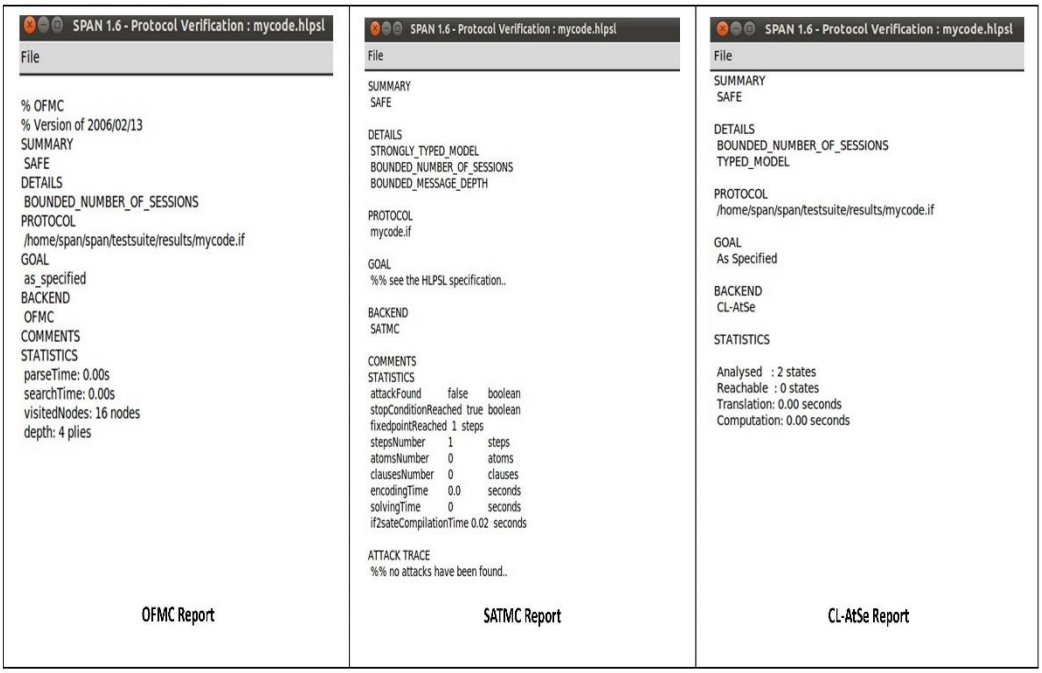


Figure 9. AVISPA simulation results.

5. DISCUSSION

The existing schemes use the complex mathematical computations for key generation and hence, they take more time to encrypt and decrypt the data. In time-sensitive application, IoT devices are deployed and need less time to communicate and authenticate themselves. The proposed scheme is compared with LTBA scheme which is based on two-factor authentication which stores biometric details using random oracle model and BBMA scheme, based on blockchain-based mutual authentication between IoT devices and edge server by using different cryptographic algorithms. The BDAEC-QA scheme takes less delay as compared with the LTBA and BBMA schemes with respect to different performance parameters. The lattice-based solutions for device-to-device authentication [29] provide the post-quantum solutions at the cost of communication overhead. The proposed work BDAEC-QA uses less overhead by considering Qkey and only the blockchain is used

to store the device information as compared with [29] where a consortium-blockchain network among edge servers is used to maintain a copy of the ledger in each edge server. The membership-service provider is added to manage access level to the ledger, hence the authentication is provided with more overhead cost. The anonymous authentication for vehicular communication [30] takes complex operations and may impact the communication overhead. The paper in [30] discusses the blockchain with smart contract and stores the public keys, which reduces the cost at the edge node. The proposed BDAEC-QA scheme also stores the IoT-device information in the blockchain using smart contract with less communication cost. The storage and computations cost of the proposed scheme are also compared with different existing schemes, as shown in Table 3 and the BDAEC-QA scheme performs better by storing less bits to store and compute the data at the edge server. As compared with LTBA scheme, our scheme uses the blockchain to store the IoT-device information at the server side; hence, it provides more security. The security analysis of the proposed scheme is done using AVISPA as compared with LTBA, BBMA and all the security requirements of the proposed scheme are met as shown in Table 4.

6. CONCLUSION

The blockchain-based device authentication using quantum approach focuses on authenticating IoT devices within and outside the edge network using quantum-key mechanism. The main objective is to implement the quantum mechanics for generating Qkey and the blockchain approach to enhance the security in case of authenticating the IoT devices. The proposed scheme works in 3 phases: IoT-device registering with the edge server and storing the context information using the quantum key. The Qkey is used to exchange messages between edge device and edge server. The CI and QKey of device are stored in the blockchain with smart contract to provide the extra-layer security at the edge server. The blockchain is an immutable ledger which provides the integrity of data stored in it. The IoT device is authenticated when the device sends the authentication message to the edge server. The proposed work is feasible, not only by resisting the major attacks, but also by performing better compared with other schemes. The BDAEC-QA scheme is simulated and compared with existing schemes with respect to various delays, computation cost and storage cost and it performs better. The scheme is also validated with CK adversary model with different attacks and analyzed using AVISPA tool to check the safety of the proposed scheme to meet the security requirements.

REFERENCES

[1] A. Kumar et al., "A Comprehensive Survey of Authentication Methods in Internet-of-Things and Its Conjunctions," Journal of Network and Computer Applications, vol. 204, Page 103414, 2022.

[2] W. Z. Khan et al., "Edge Computing: A Survey," Future Generation Computer Systems Journal, vol. 97, pp. 219–235, 2019.

[3] K. Cao, Y. Liu, G. Meng and Q. Sun, "An Overview on Edge Computing Research," IEEE Access Journal, vol. 8, pp. 85714–85728, 2020.

[4] Q. Fan et al., "A Secure and Efficient Authentication and Data Sharing Scheme for Internet of Things Based on Blockchain," Journal of Systems Architecture, vol. 117, Page 102112, 2021.

[5] P. M. Chanal and M. S. Kakkasageri, "Security and Privacy in IoT: A Survey," Wireless Personal Communications Journal, vol. 115, pp. 1667—1693, 2020.

[6] P. Memarmoshrefi, R. Seibel and D. Hogrefe, "Autonomous Ant-based Public Key Authentication Mechanism for Mobile *Ad-hoc* Networks," Journal of Mobile Networks and Applications, vol. 21, pp. 149–160, 2016.

[7] S. S. Rani, S. Pradeep, R. M. Dinesh and S. G. Prabhu, "OTP Based Authentication Model for Autonomous Delivery Systems Using Raspberry Pi," Proc. of the Int. Conf. on Intelligent Controller and Computing for Smart Power (ICICCSP), pp. 1–5, Hyderabad, India, 2022.

[8] M. Mitev et al., "Authenticated Secret Key Generation in Delay-constrained Wireless Systems," EURASIP Journal of Wireless Communication and Networking, vol. 2020, Article no. 122, 2020.

[9] Y. Wang, T. Nakachi and H. Ishihara, "Edge and Cloud-aided Secure Sparse Representation for Face Recognition," Proc. of the 27th IEEE European Signal Processing Conf. (EUSIPCO), pp. 1–5, A Coruna, Spain, 2019.

[10] H. Goumidi et al., "Lightweight Secure Authentication and Key Distribution Scheme for Vehicular Cloud Computing," Journal of Symmetry, vol. 13, no. 3, Article no. 484, pp. 1-29, 2021.

[11] J. Mulholland, M. Mosca and J. Braun, "The Day the Cryptography Dies," IEEE Security Privacy, vol. 15, no. 4, pp. 14–21, 2017.

[12] N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, "Quantum Cryptography," Reviews of Modern Physics, vol. 74, pp. 145–195, 2002.
[13] C. H. Bennett and G. Brassard, "Quantum Cryptography: Public Key Distribution and Coin Tossing," Theoretical Computer Science, vol. 560, pp. 7–11, 2014.
[14] H. Zeyu et al., "Survey on Edge Computing Security," Proc. of the IEEE Int. Conf. on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), pp. 96–105, Fuzhou, China, 2020.
[15] X. Wang et al., "Survey on Blockchain for Internet of Things," Journal of Computer Communications, Elsevier, vol. 136, pp. 10–29, 2019.
[16] M. A. Uddin, A. Stranieri, I. Gondal and V. Balasubramanian, "A Survey on the Adoption of Blockchain in IoT: Challenges and Solutions," Blockchain: Research and Applications, vol. 2, no. 2, pp. 1–49, 2021.
[17] Y. Zhang et al., "Smart Contract-based Access Control for the Internet of Things," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 1594– 1605, 2018.
[18] K. Ekerta, "Quantum Cryptography Bases on Bell's Theorem," Physical Review Letters, vol. 67, pp. 661–664, 1991.
[19] V. Scarani et al., "Security Aspect of Practical Quantum Key Distribution," Reviews of Modern Physics, vol. 81, no. 3, pp. 1301–1350, 2009.
[20] U. Khalid et al., "A Decentralized Lightweight Blockchain-based Authentication Mechanism for IoT Systems," Cluster Computing, vol. 23, pp. 2067–2087, 2020.
[21] K. Hameed, S. Garg, M. B. Amin and B. Kang, "A Formally Verified Blockchain-based Decentralized Authentication Scheme for the Internet of Things," Journal of Supercomputing, vol. 77, pp. 14461—14501, 2021.
[22] A. A. Al-Saggaf, T. Sheltami, H. Alkhzaimi and G. Ahmed, "Lightweight Two-factor-based User Authentication Protocol for IoT-enabled Healthcare Ecosystem in Quantum Computing," Arab Journal for Science and Engineering, vol. 48, pp. 2347–2357, 2023.
[23] H. Abulkasim et al., "Authenticated Secure Quantum-based Communication Scheme in Internet-of-Drones Deployment," IEEE Access Journal, vol. 10, pp. 94963–94972, 2022.
[24] J. Wu, Z. Jin, G. Li, Z. Xu, C. Fan and Y. Zheng, "Design of Vehicle Certification Schemes in IoV Based on Blockchain," World Wide Web Journal, vol. 25, pp. 2241—2263, 2022.
[25] G. Cheng, Y. Chen, S. Deng, H. Gao and J. Yin, "A Blockchain-based Mutual Authentication Scheme for Collaborative Edge Computing," IEEE Transactions on Computational Social Systems, vol. 9, no. 1, pp. 146–158, 2022.
[26] A. K. Sahu, S. Sharma and D. Puthal, "Lightweight Multi-party Authentication and Key Agreement Protocol in IoT-based E-Healthcare Service," ACM Transactions on Multimedia Computing, Communications and Applications, vol. 17, pp. 1–20, 2021.
[27] P. Nag, P. Chandrakar and K. Chandrakar, "An Improved Two-factor Authentication Scheme for Healthcare System," Procedia Computer Science, vol. 218, pp. 1079–1090, 2023.
[28] S. Sathyadevan, K. Achuthan, R. Doss and L. Pan, "Protean Authentication Scheme – A Time-bound Dynamic KeyGen Authentication Technique for IoT Edge Nodes in Outdoor Deployments," IEEE Access Journal, vol. 7, pp. 92419–92435, 2019.
[29] A. Shahidinejad and J. Abawajy, "Decentralized Lattice-based Device-to-device Authentication for the Edge-enabled IoT," IEEE Systems Journal, vol. 17, no. 4, pp. 6623–6633, 2023.
[30] A. Shahidinejad, J. Abawajy and S. Huda, "Anonymous Lattice-based Authentication Protocol for Vehicular Communications," Vehicular Communications, vol. 48, Page 100803, 2024.
[31] D. Stucki et al., "Long-term Performance of the SwissQuantum Quantum Key Distribution Network in a Field Environment," New Journal of Physics, vol. 13, no. 12, Page 123001, 2011.
[32] Toshiba,"QKD Technology to Semiconductor Chip," [Online], Available: https://news.toshiba.com/press-releases/press-release-details/2021/Toshiba-Shrinks-Quantum-Key-Distribution-Technology-to-a-Semiconductor-Chip/default.aspx.
[33] R. Canetti and H. Krawczyk, "Analysis of Key-exchange Protocols and Their Use for Building Secure Channels," Proc. of the International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2001), pp. 453–474, Springer, 2001.
[34] N. D. Sarier, "Multimodal Biometric Authentication for Mobile Edge Computing," Information Sciences, vol. 573, pp. 82–99, 2021.
[35] H. Goswami and H. Choudhury, "Remote Registration and Group Authentication of IoT Devices in 5G Cellular Network," Computers Security Journal, vol. 120, Page 102806, 2022.
[36] A. M. Almuhaideb and K. S. Alqudaihi, "A Lightweight Three-factor Authentication Scheme for WHSN Architecture," Sensors Journal, vol. 20, no. 23, Page 6860, 2020.
[37] S. Uppuluri and G. Lakshmeeswari, "Secure User Authentication and Key Agreement Scheme for IoT Device Access Control Based Smart Home Communications," Wireless Network Journal, vol. 29, pp. 1333—1354, 2023.

[38] Wu, Tsu-Yang, Zhiyuan Lee, Lei Yang, Jia-Ning Luo and Raylin Tso, "Provably Secure Authentication Key Exchange Scheme Using Fog Nodes in Vehicular *Ad Hoc* Networks," The Journal of Supercomputing vol. 77, no. 7, pp. 6992-7020, 2021.

[39] M. Hamada, S. A. Salem and F. M. Salem, "LAMAS: Lightweight Anonymous Mutual Authentication Scheme for Securing Fog Computing Environments," Ain Shams Engineering Journal, vol. 13, no. 6, p. 101752, 2022.

[40] I. Aciobanitei, R. I. Guinea and M. L. Pura, "AVISPA *versus* AVANTSSAR in the Model Checking of Secure Communication Protocols," Proc. of the 15th Int. Joint Conf. on e-Business and Telecomm. (ICETE 2018), vol. 2: SECRYPT, pp. 520-525, DOI: 10.5220/0006887905200525, 2018.

ملخص البحث:

ظهرت إنترنت الأشياء كتكنولوجيا جديدة، حيث كلُّ شيء متّصل. وينبغي تخزين كمّياتٍ ضخمة من البيانات، لذا فإنّ حوسبة الحافة يمكنها التّقليل من تخزين البيانات في بيئةٍ موزَّعة، الأمر الّذي من شأنه أن يحسن من سرعة المعالجة ويؤدّي إلى استخدام قدْر أقلّ من عرض النّطاق. ومع ازدياد استخدام أجهزة إنترنت الأشياء، ازداد ظهور قضايا مثل المُصادقة على الأجهزة، وخصوصية البيانات المخزّنة، وتكامل البيانات. وتُعدّ المُصادقة على الأجهزة مسألةً مهمّةً بالنّسبة إلى أجهزة إنترنت الأشياء المتّصلة باستخدام حوسبة الحافَة. وقد تمّت معالجة هذه المسألة عن طريق خوارزميات ترميزٍ كلاسيكية لترميز الرّسائل باستخدام المفاتيح العامّة والمفاتيح الخاصّة الّتي يتعيّن تخزينها. ويتطلب الأمر تخزين هذه المفاتيح في جهاز خادم من أجل المُصادقة على الأجهزة، علماً بأنّ تخزين عددٍ كبيرٍ من المفاتيح يؤدّي إلى ازدياد تكلفة الحوسبة وتكلفة التّخزين وإلى ازديادٍ في التّأخير. وباستخدام الحوسبة الكمّيّة والخوارزميات الكمّيّة، يُصبح من السّهل كَسْرُ خوارزميات التّرميز، الأمر الّذي يجهل النّظام هشّاً.

إنّ الآلية المقترحة في هذا البحث من شأنها أن توفِّر المُصادقة على أجهزة إنترنت الأشياء باستخدام معلومات السّياق وتوزيع المفاتيح الكمّيّة وسلاسل الكُتل. وتستخدم الطّريقة المقترحة "العُقود الذّكية" لتخزين المعلومات الخاصّة بأجهزة إنترنت الأشياء في جهاز الخادم الّذي تستخدمه سلسلة الكُتل من أجل تأمين المُصادقة المتبادلة بين أجهزة إنترنت الأشياء عبر الشّبكة.

لقد جرت مقارنة الطّريقة المقترحة بعددٍ من الطُّرق المشابهة الواردة في أدبيات الموضوع، وامتازت الطّريقة المقترحة بقيم أقلّ لكلٍّ من زمن التّسجيل وزمن توليد المفاتيح وزمن تأخير المُصادقة. كذلك تميّزت الطّريقة المقترحة بقدْرٍ أقلّ لتكلفة الحوسبة وتكلفة التّخزين، مقارنةً بغيرها من الطُّرق الواردة في دراسات سابقة أخرى. وقد تمّت محاكاة الطّريقة المقترحة باستخدام أداة (AVISPA) للبرهنة على أمانها ومُقاومتها للهجمات.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 11, No. 01, March 2025.

ENHANCING MICRO-EXPRESSION RECOGNITION: A NOVEL APPROACH WITH HYBRID ATTENTION-3DNET

Budhi Irawan[1,2], Rinaldi Munir[2], Nugraha Priya Utama[2] and Ayu Purwarianti[2]

(Received: 5-Sep.-2024, Revised: 11-Nov.-2024, Accepted: 12-Nov.-2024)

ABSTRACT

This paper proposes a unique pipeline for micro-expression recognition using a Dual-path 3D Convolutional Neural Network enhanced with Hybrid Attention and Squeeze-and-Excitation Blocks. The three main goals of the pipeline are to (1) Optimize the extraction of spatial-temporal features using advanced neural network architectures, (2) Enhance data representation by implementing targeted image augmentation and balanced class distribution and (3) Enhance feature fusion using state-of-the-art network techniques. Comprehensive experiments were conducted on four benchmark datasets: CAS(ME)^2, SMIC, SAMM and CASME II. The Hybrid Attention-3DNet model demonstrated superior recognition accuracy of 93.95% for CAS(ME)^2, 93.42% for SMIC, 93.61% for SAMM and 93.79% for CASME II, surpassing the state-of-the-art methods across these datasets. These outcomes demonstrate the efficacy and robustness of the proposed pipeline, underscoring its potential for a range of micro-expression recognition uses.

KEYWORDS

Micro-expression recognition, 3D convolutional dual-path network, Hybrid attention, Squeeze-and-excitation blocks, Deep learning.

1. INTRODUCTION

Micro-expressions are quick, uncontrollable facial movements that show true feelings that a person may try to hide. Even skilled observers may find it challenging to identify these expressions, since they are brief, frequently lasting less than 0.5 seconds [1]. Identifying micro-expressions has several uses, especially in security, psychology and medicine, where it is essential to comprehend genuine emotions [2].

New developments in deep learning have made it possible to create complex models to identify these nuanced expressions. Nevertheless, fundamental difficulties persist, including a lack of data and a notable disparity in micro-expression classes [3]. Furthermore, conventional Convolutional Neural Networks (CNNs) frequently demand data and require assistance with overfitting in situations where data is limited [4].

This study suggests a unique architecture that combines Hybrid Attention and Squeeze-and-Excitation Blocks into a Dual-path 3D Convolutional Neural Network (3DCNN) to improve micro-expression identification. Tested on benchmark datasets, including CAS(ME)^2,[5] SMIC [6], SAMM [7] and CASME II [8], the suggested model outperformed state-of-the-art techniques in terms of accuracy, indicating its potential for practical use in emotion recognition [9].

2. RELATED WORK

Recent developments in micro-expression recognition have sparked the creation of creative techniques to increase precision. Combining CNNs with other methods is one such strategy. A technique that combines Swin Transformer and ConvNeXt is presented in [10] and is based on a Dual-branch Spatiotemporal Convolutional Network (STCN). This method uses both CNN and Transformers to address issues, including the preservation of facial spatial structure and the localization of micro-expression actions. According to tests on the CASME and SMIC datasets, the STCN network improves micro-expression identification accuracy.

The Divided-block Multi-scale Convolution Network (DBMNet) is a unique multi-scale convolutional

1. B. Irawan is with Telkom University, Indonesia. Email: budhiirawan@telkomuniversity.ac.id
2. B. Irawan, R. Munir, N. P. Utama and A. Purwarianti are with Bandung Institute of Tecnology, Indonesia. Emails: budhiirawan@telkomuniversity.ac.id, rinaldi@informatika.org, utama@informatika.org and ayu@informatika.org

network proposed in another study [11]. This network is intended to learn from four different optical flow feature images produced between the micro-expression samples' onset and apex frames. With the use of the Divided-block Multi-scale Convolution Module (DBMCM), the network can efficiently capture more intricate and useful multi-scale properties of micro-expressions.

A deep-learning technique known as the Spatiotemporal Capsule Network (STCP-Net) was recently presented in [12]. This method aims to increase recognition accuracy while decreasing recognition time. The four main parts of STCP-Net are a jitter removal module, a differential feature-extraction module, a spatiotemporal capsule module and a fully connected layer.

[13] presents the Parallel Dual-branch Attention-based Spatio-temporal Fusion Network (PASTFNet). This method is influenced by the combined architecture of Long-Short-Term Memory (LSTM) and CNN for temporal modeling. The paper suggests encoding sequential frame features using an attention-based multi-scale feature-fusion network (AMFNet). The network gathers more expressive face-detail features for micro-expression recognition through multi-scale feature fusion and integrated attention.

A two-layer feature-encoding technique is suggested in [14] to depict interactions across different regions of the feature map, along with a novel multi-frame technique intended to capture subtle motion patterns. The paper also presents an Action Unit Graph Convolutional Network (AU GCN). It uses a transformer encoder, an adjacency matrix and an AU-detection module to adjust to test data.

A Triple-branch Attention Fusion Network (Triple-ATFME) is presented in [15] for micro-expression recognition. With the help of a Triple-branch ShuffleNet module, an adaptive channel attention module and pre-processing, this approach enables the model to extract multi-view features using a multi-path architecture. The framework uses optical-flow approaches to capture various optical-flow information by extracting optical-flow features from the facial region's cropped start and peak frames. The Triple-ATFME network processes these features to find hidden features. Channel features are adjusted *via* a Channel Fusion Attention Module (CFAM) to improve multi-view feature integration and lessen the model's emphasis on local information during feature fusion.

Lastly, the suggested framework in this research introduces a unique method for improving the accuracy of micro-expression identification through spatio-temporal deep learning, data augmentation and class balancing [16]. Rotation, contrast adjustment and SMOTE are examples of sophisticated pre-processing techniques that the framework uses to effectively handle data restrictions and class imbalances, enhancing model generalization and lowering overfitting. Based on the results, this method set a new standard for micro-expression analysis and created a more dependable model for emotion-detection applications. It also made notable gains, especially in accuracy and F1-scores across many datasets. In addition, this report serves as a baseline for the research.

3. PROPOSED METHOD

This study adopts a systematic approach by developing a pipeline to classify input video datasets of spontaneous micro-expressions. The datasets are categorized into three emotional classes: angry, happy and disgusted for the CAS(ME)² dataset and positive, negative and surprise for the SMIC, SAMM and CASME II datasets. The study is structured around four experimental scenarios, each tailored to the input from these four datasets.

The developed pipeline consists of several stages, including data preparation, pre-processing, classification and performance measurement, as illustrated in Figure 1. The datasets are organized based on their respective emotional classes in the data-preparation stage. Specifically, the CAS(ME)² dataset is divided into three classes: angry, happy and disgusted, while the SMIC, SAMM and CASME II datasets are categorized into positive, negative and surprise classes.

During pre-processing, the video clips from the datasets are converted into a series of image frames, followed by face detection and the identification of 68 facial landmarks. The areas around the eyes and mouth are masked and the face is cropped. The images are then resized to 128x128 pixels and converted into grayscale. Additionally, data augmentation is performed by adjusting orientation, contrast and brightness and class balancing is achieved using the class-weight method. The pre-processed facial images are subsequently divided into upper and lower face regions.

This study proposes a model for the classification task that utilizes a dual-path 3D convolutional neural

network incorporating hybrid attention and Squeeze-and-Excitation blocks. Finally, the emotion classification is evaluated using accuracy, F1-score and error rate, ensuring a comprehensive assessment of the model's performance.

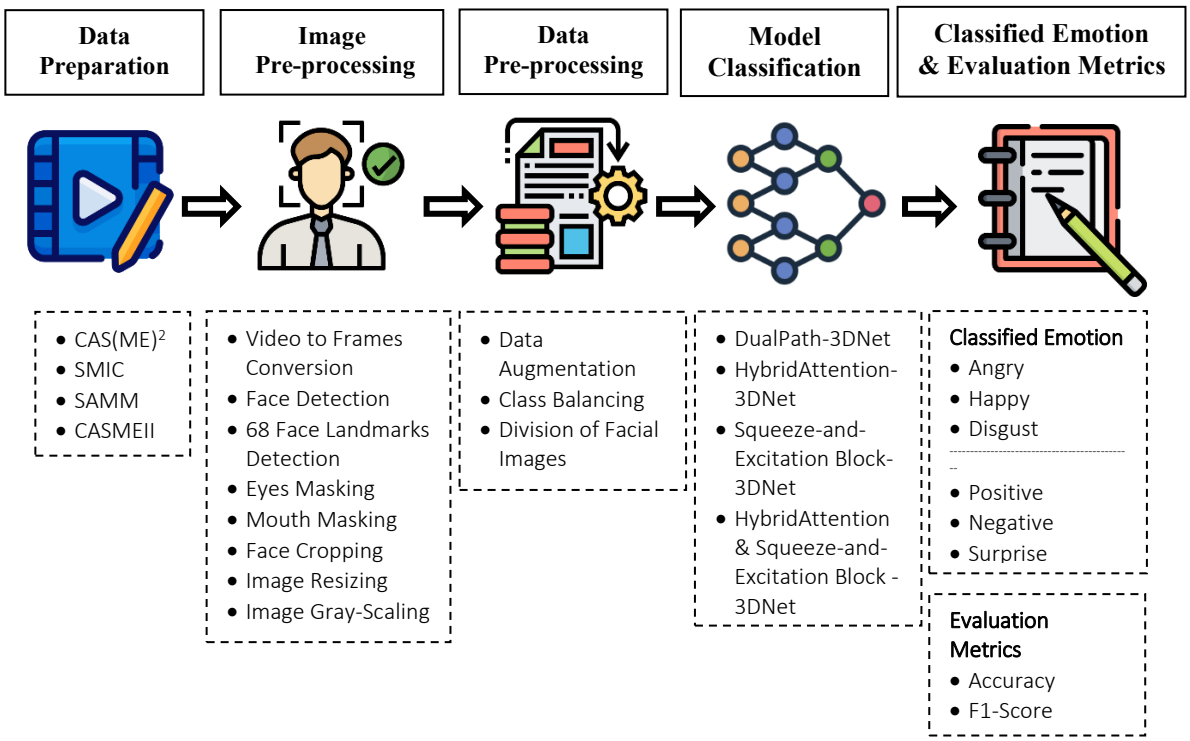


Figure 1. The proposed design of a micro-expression recognition pipeline.

3.1 Pre-processing Stages

The process begins with image pre-processing to identify spontaneous micro-expressions from extended sequences of facial videos. Initially, the emotional classes of each dataset are categorized. The CAS(ME)² dataset includes categories for emotions, such as anger, happiness and disgust, while the SMIC, SAMM and CASME II datasets cover positive, negative and surprise emotions. In the next step, video clips from these datasets are converted into sequential image frames. The original resolutions of these frames vary by dataset: 640x480 pixels for CAS(ME)² and SMIC, 960x560 pixels for SAMM and 280x340 pixels for CASME II, as depicted in Figure 2.

The provided diagram demonstrates the sequential steps of image and data pre-processing, beginning with raw video input and concluding with facial-image sequences split into upper and lower sections, each resized to 128x64 pixels. The main goal of pre-processing is to optimize raw video data for practical use in micro-expression recognition. This process transforms video data into clean, structured image sequences that highlight essential facial features, reduce noise and maintain a balanced distribution of classes. It begins by converting the raw video into individual frames, with each dataset's frames retaining specific resolutions: 640x480 pixels for CAS(ME)² and SMIC, 960x560 pixels for SAMM and 280x340 pixels for CASME II. Face detection is applied to each frame during pre-processing, followed by identifying 68 facial landmarks. These landmarks outline critical facial areas, such as the eyes, mouth and nose, guiding the masking and segmentation steps necessary for accurate micro-expression recognition.

Once the landmarks are detected, the eye and mouth areas are masked to focus on the most expressive facial regions, which aids the model in capturing subtle movements associated with micro-expressions. The face is then cropped, resized to 128x128 pixels and converted into grayscale to simplify color-data without sacrificing critical information, thereby speeding up analysis while preserving accuracy. Data augmentation techniques, including rotation, cropping and contrast adjustments, are also applied to enhance data variability. This step increases the diversity of the dataset, helping the model to generalize across various angles and lighting conditions.

Class balancing is implemented to ensure a balanced representation across classes by assigning weights to each class, reducing bias towards majority classes and improving the model's ability to recognize minority classes accurately. Finally, the pre-processed facial images are divided into upper and lower sections, each resized to 128x64 pixels and organized into frame sequences ready for classification. This segmentation enables the model to analyze distinct facial areas independently, enhancing the detection of subtle changes in the eyes and mouth regions that play a crucial role in micro-expression recognition.

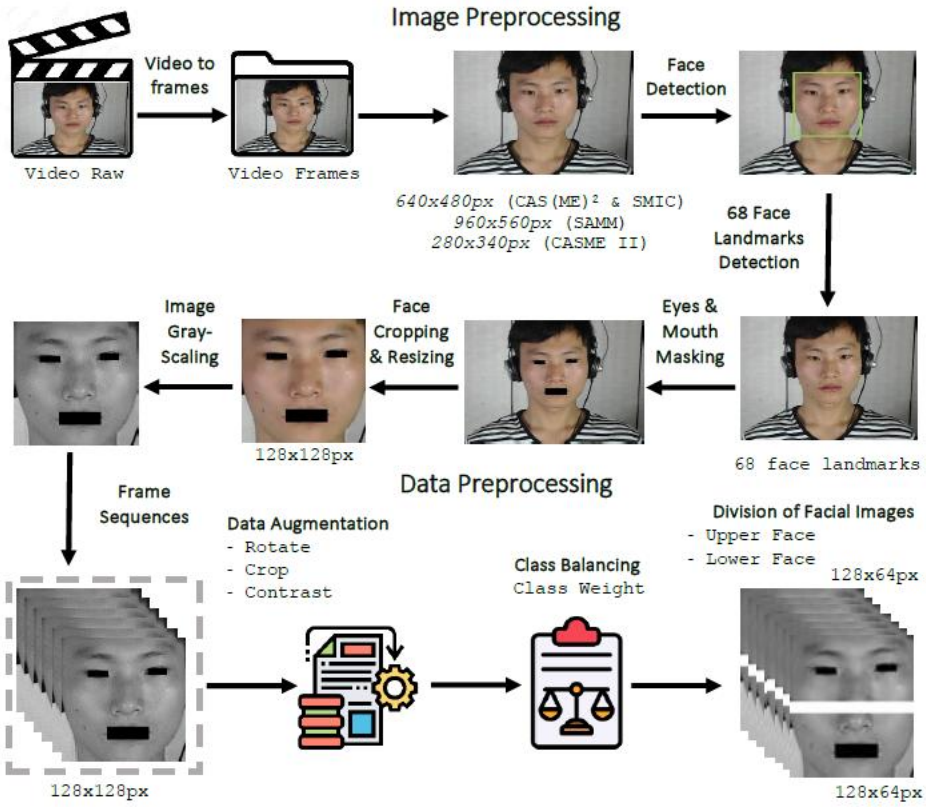


Figure 2. Pre-processing stages.

3.2 Data Optimization

The quality of a dataset is crucial in ascertaining the accuracy and efficiency of machine-learning models in data processing. Imbalanced or unrepresentative data may result in biased models and suboptimal performance. Consequently, diverse methodologies enhance data, enabling models to learn more efficiently and generate more precise predictions. Data augmentation and class balance are two essential methodologies employed in this process, which will be examined in more detail in the subsequent sections. Data-augmentation techniques are essential in machine learning and image processing, particularly for improving the quality and quantity of training datasets. These strategies entail methodically modifying existing photos to create new variants and augmenting the dataset without further data collection. Data augmentation is indispensable in micro-expression recognition, where extensive datasets are vital for enhancing model accuracy.

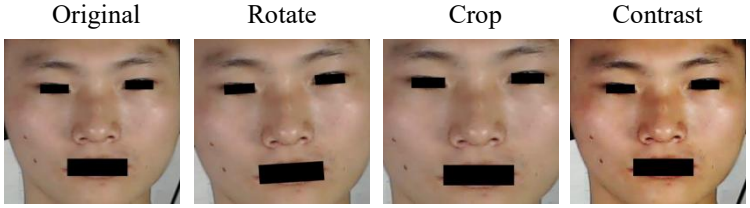


Figure 3. Data augmentation.

In this context, image frames can undergo various transformations, including rotation, cropping and modifications to brightness and contrast, as illustrated in Figure 3. The transformations generate varied representations of the original images, enhancing the model's ability to recognize and adapt to diverse

patterns and variations in the data. Data augmentation enhances model generalization by increasing the diversity of training samples, resulting in improved performance across various scenarios.

Class balancing is a technique employed to equalize the distribution of samples or observations across various classes within a dataset. In classification tasks, a class denotes the specific label or category that the model aims to predict. An imbalance occurs when there is a significant disparity in the number of samples between classes. The imbalance in datasets presents considerable challenges in machine learning, as models trained on such uneven data frequently produce biased predictions, favoring the majority classes while overlooking minority groups. This bias may hinder the model's capacity to identify and classify instances from the minority classes accurately. This study implements a method of class balancing known as the class weight method.

The class weight method is instrumental in detecting micro-expressions, particularly by addressing distribution imbalances in the dataset. In micro-expression recognition, class imbalance is a common issue; some classes have more samples than others. By assigning higher class weights to the less common classes, the model is encouraged to learn from the minority classes more effectively. This approach is beneficial in mitigating class imbalance and enhancing the model's ability to recognize minority classes. Furthermore, rare classes, such as specific facial expressions that occur infrequently, typically exhibit lower accuracy due to the limited number of samples [17]. Assigning class weights offers greater motivation for the model to accurately recognize these classes, thereby improving its performance on minority classes.

In addition, applying class weights is anticipated to minimize the risk of overfitting to the majority class. Assigning weights to each class prevents the model from overly focusing on the majority class and promotes a more balanced learning process across all classes. The equation used for calculating class weights in cases of class imbalance often involves comparing the sizes of the classes or employing simple proportional methods. The general formula for computing class weights is given by Equation (1):

$$W_k = \frac{N}{n_k} \tag{1}$$

where W_k is the class weight for class k, N is the total number of samples in the training data and n_k is the number of samples in class *k*.

This equation implies that the minority class will have a higher class weight than the majority class. This inverse relationship between the sample proportion within the class and the assigned class weight gives the minority class more importance in the learning process, aiding in overcoming class imbalance.

3.3 Division of Facial Images

The Division of Facial Images is crucial in enhancing micro-expression recognition by focusing on specific regions of facial-muscle activity. In micro-expressions, Action Units (AU) are located in distinct facial regions where subtle muscle movements occur, often concentrated around the eyes, eyebrows and mouth. These areas contain a dense, exemplary muscle network that enables intricate facial movements. For example, muscle contractions around the eyes can form wrinkles, while those around the mouth can alter lip shape. This division aims to develop a more precise and focused approach to facial micro-expression recognition by acknowledging the significance of AU locations.

In the context of a dual-stream input classification model for facial-expression recognition, dividing the facial image into upper and lower sections enables the model to concentrate on the specific distribution of AUs across the face. This division facilitates the separate processing of the upper and lower face regions, aligning with their distinct roles in expressing emotions. The original image, sized at 128x128 pixels, is split into two parts, each measuring 128x64 pixels. According to Ekman's research on facial regions and emotional expression, the upper face, which includes the eye and eyebrow areas, is predominantly associated with emotions such as positivity and surprise. Conversely, the lower face, encompassing the nose, mouth and cheeks, often conveys subtler or more complex emotional nuances, particularly negative emotions.

The process of dividing facial images in this way enables the model to capture spatio-temporal features related to micro-expression AUs more effectively. By processing these sections separately, the dual-stream classification model can focus on the distinct emotional signals conveyed by each part of the

face. This separation allows the model to detect fine-grained emotional expressions, improving accuracy in recognizing micro-expressions across different facial regions. The illustration in Figure 4 demonstrates the process of dividing the facial image into upper and lower sections, each resized to 128x64 pixels, enhancing the model's capacity to analyze and classify these regions independently. Thus, the Division of Facial Images facilitates the targeted analysis of key facial regions and contributes to a more nuanced understanding of emotion-expression dynamics, which is essential for effective micro-expression recognition.

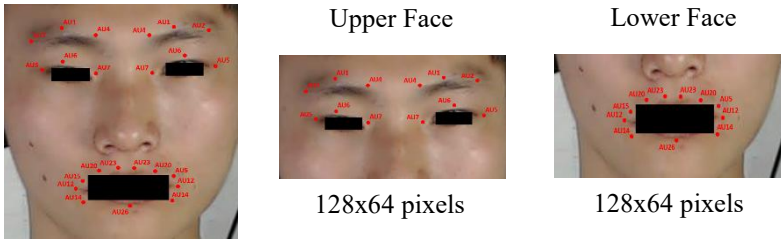


Figure 4. Division of facial images.

3.4 The 3D Convolutional Dual Path Network Model

This study utilizes a 3D convolutional dual-path network model for micro-expression recognition, which enhances attention to spatial-temporal feature weights. The model incorporates hybrid attention and squeeze-and-excitation blocks to recognize spontaneous micro-expressions. This dual-path model is an extension of the single-path model [18], which primarily focuses on general facial features without addressing the detailed features of the upper and lower facial regions. The proposed research pipeline includes four model designs: Dual Path-3DNet, Hybrid Attention-3DNet, SE Block-3DNet and Hybrid Attention-3DSENet.

3.4.1 Dual Path-3DNet

The purpose of proposing the Dual Path-3DNet model is to enhance the recognition of micro-expressions by utilizing a dual-stream approach that separately processes different facial regions, thereby capturing more detailed spatial-temporal features relevant to each region. Using two distinct input paths, this model can independently analyze the upper and lower sections of the face derived from the pre-processing steps applied to the four datasets used in this study. This dual-path strategy allows the model to focus on region-specific characteristics in each section, optimizing the detection of subtle emotional cues that may be localized to particular facial areas.

In the Dual Path-3DNet model's design, each input path is structured to process sequences of pre-processed image frames and the segmented upper-face and lower-face regions. Each path includes layers, including 3D convolution with ReLU activation, 3D max pooling, flatten, dense with ReLU and dropout layers. These layers enable effective feature extraction and reduce overfitting by discarding non-essential data points. The outputs of the two paths are then merged into a single pathway through a concatenate layer, which integrates the distinct information extracted from both facial regions.

Following this merging, the combined pathway passes through additional dense layers with ReLU activation, dropout and finally, a softmax layer for classification. This final pathway enables the model to learn complex interrelations between features from both facial regions, allowing for more accurate and nuanced emotion detection. The architecture of the Dual Path-3DNet model is illustrated in Figure 5, where each layer and process flow is visually represented to demonstrate the interaction between the dual pathways. By adopting this dual-path approach, the model is better equipped to recognize micro-expressions by simultaneously analyzing diverse facial features across regions. This ultimately contributes to higher accuracy in emotional-recognition tasks.

3.4.2 Hybrid Attention-3DNet

The proposed design of the Hybrid Attention-3DNet model shares the same basic architecture as the Dual Path-3DNet model. The critical difference in this architecture is the addition of a hybrid-attention layer (encompassing both Spatial and Temporal Attention) placed after the 3D max pooling layer. This

is followed by flatten, dense + ReLU and dropout layers, which are then merged into a single path using a concatenate layer. After this merging process, the subsequent path consists of dense layers with ReLU activation, dropout and a softmax layer. The architecture of the Hybrid Attention-3DNet model is illustrated in Figure 6.

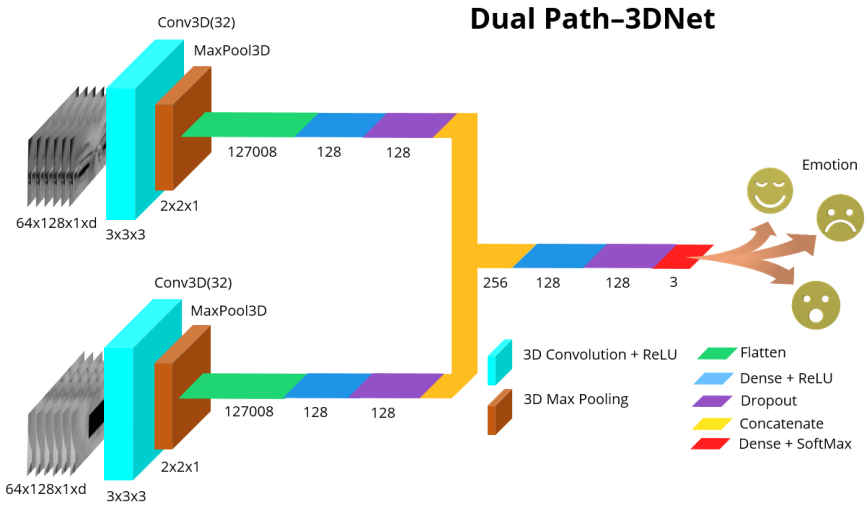


Figure 5. The architecture of the dual path-3DNet model.

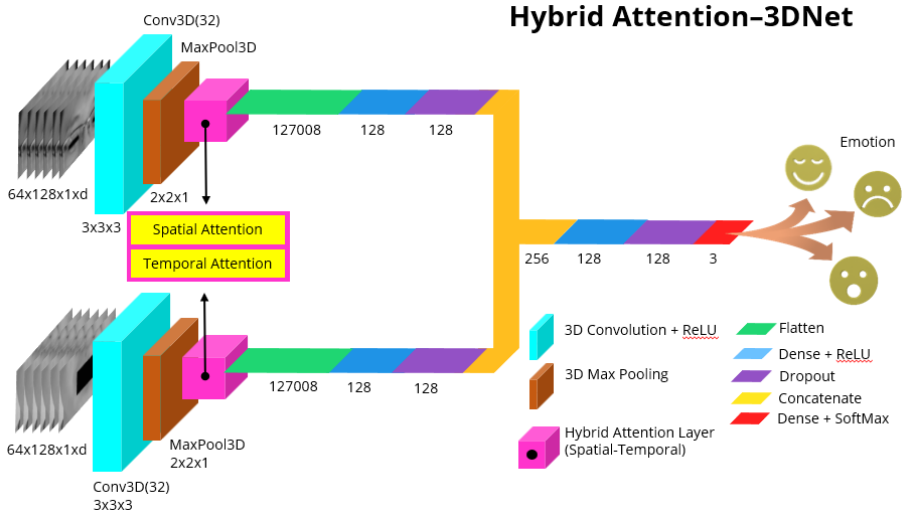


Figure 6. The architecture of the hybrid attention-3DNet model.

This model incorporates Hybrid Attention, composed of Spatial and Temporal Attention, to enhance its ability to recognize micro-expressions. Hybrid Attention is applied following the 3D max pooling layer, where each attention mechanism, spatial and temporal, contributes to the improved representation of spatial and temporal features.

Spatial attention is designed to identify critical spatial features within facial regions. This mechanism increases the weights for crucial areas, such as those around the eyes and mouth, using an attention map to determine the distribution of activations across spatial features. By selectively amplifying significant features, the model is better equipped to identify specific patterns linked to micro-expressions that could be challenging to detect without spatial attention.

Temporal attention focuses on capturing sequential changes in facial expressions, allowing the model to detect subtle variations that occur from frame to frame over short durations in micro-expression videos. This layer assigns weights based on temporal dynamics, enabling the model to recognize small changes in facial expressions that might be overlooked by traditional methods that do not incorporate temporal information.

The integration process and benefits from Hybrid Attention are incorporated after the 3D max pooling

layer and positioned before the dual-path 3DCNN concatenation stage. This layer processes inputs from the refined feature maps, ensuring that spatial and temporal attention mechanisms are employed before the final classification step. Through this combination, the model can prioritize essential features in both spatial and temporal dimensions, thus enhancing responsiveness to rapid, subtle variations in expressions, leading to more accurate micro-expression detection in sequential video data.

3.4.3 Squeeze-and-Excitation Block-3DNet

Incorporating the Squeeze-and-Excitation Block-3DNet model enhances feature selection by emphasizing critical spatial-temporal information within the data, which is essential for accurate micro-expression recognition. While the SE Block-3DNet model shares the basic structure with the Dual Path-3DNet model, it introduces a Squeeze-and-Excitation block after the 3D max pooling layer. This additional layer selectively emphasizes significant features by applying global average pooling (Global AP) to squeeze the spatial dimensions, followed by fully connected layers with ReLU and sigmoid activation to recalibrate the feature-map channels.

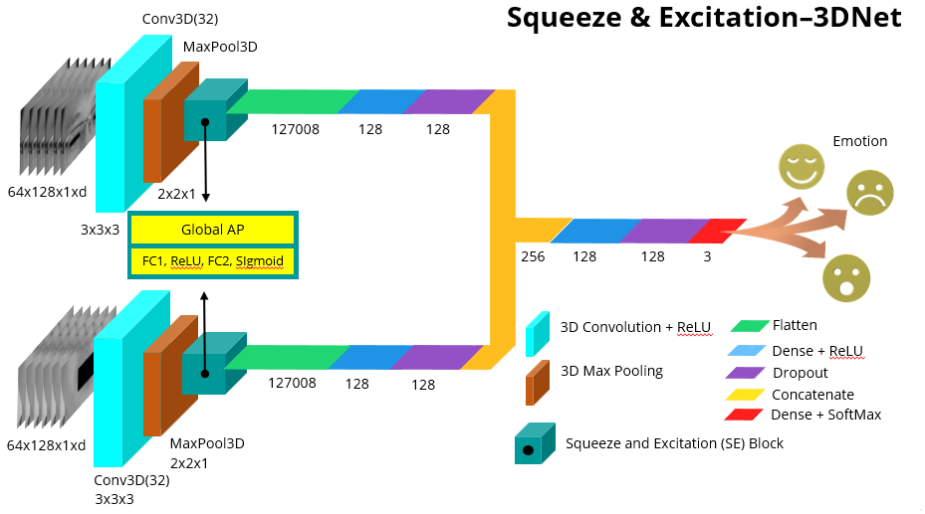


Figure 7. The architecture of the squeeze & exitation-3DNet model.

After recalibration, the Squeeze-and-Excitation Block's output undergoes the same layers as the Dual Path-3DNet model, including flatten, dense with ReLU and dropout layers. The pathways from each input are then merged using a concatenate layer. Further dense layers with ReLU activation, dropout and a softmax layer for final classification follow this merging. The Squeeze-and-Excitation Block's selective attention mechanism helps the model focus on essential micro-expression cues, optimizing the feature representation for each facial region.

As shown in Figure 7, the SE Block-3DNet model's architecture incorporates this additional Squeeze-and-Excitation Block, which enhances the model's ability to prioritize essential features, leading to improved performance in emotion-recognition tasks. This approach leverages spatial recalibration and dual-path processing to capture fine-grained details, making it a powerful method for precise micro-expression analysis.

3.4.4 Hybrid Attention-3DSENet

Introducing the Hybrid Attention Squeeze-and-Excitation Block-3DNet model aims to enhance the model's ability to capture both spatial and temporal features essential for recognizing micro-expressions. This model builds on the fundamental structure of the Dual Path-3DNet, but incorporates a hybrid attention layer that combines both spatial and temporal attention mechanisms. Placed after the 3D max pooling layer, this hybrid-attention layer selectively focuses on important spatial locations and temporal sequences, optimizing feature extraction for subtle micro-expressions.

After the attention layer, the model continues with flatten, dense + ReLU and dropout layers, which are subsequently merged using a concatenate layer to integrate features from both input paths. Following this merge, a Squeeze-and-Excitation Block layer is added, providing further refinement by recalibrating

channel importance and is then followed by dense layers with ReLU activation, dropout and finally, a softmax layer for classification.

As illustrated in Figure 8, the Hybrid Attention-3DSENet architecture effectively combines spatial and temporal attention with channel recalibration through the Squeeze-and-Excitation Block. This integration allows the model to prioritize essential micro-expression cues across both dimensions, producing more precise and robust emotion recognition. This hybrid-attention mechanism and dual-path processing make the model particularly effective for detailed micro-expression analysis.

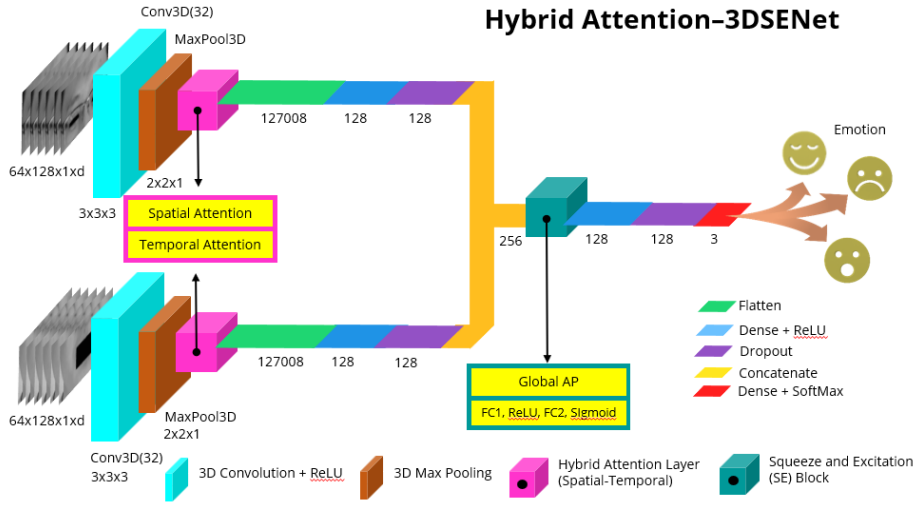


Figure 8. The architecture of the squeeze-and- exitation-3DNet model.

3.5 Hybrid Attention and Squeeze-and-Excitation Block

The attention mechanism is a technique in artificial neural networks that enables the model to focus on specific input parts when making predictions. This technique is beneficial in tasks, such as pattern recognition and object detection in vision systems. The mechanism assigns different weights to input elements, allowing the model to emphasize more relevant information while disregarding less essential details. The primary benefit of the attention mechanism is its ability to improve model performance by capturing more complex contexts and reducing computational load by concentrating resources on the most crucial information. Consequently, the model becomes more efficient and accurate in processing large and heterogeneous datasets.

One variant of the attention mechanism is hybrid attention. Hybrid attention is an approach that combines spatial and temporal attention in a 3D convolutional network model to enhance performance in tasks, such as image or video classification [19]. The primary function of the Hybrid Attention layer is to enable the model to capture important information spatially and temporally from the input data. The Hybrid Attention layer is illustrated in Figure 9.

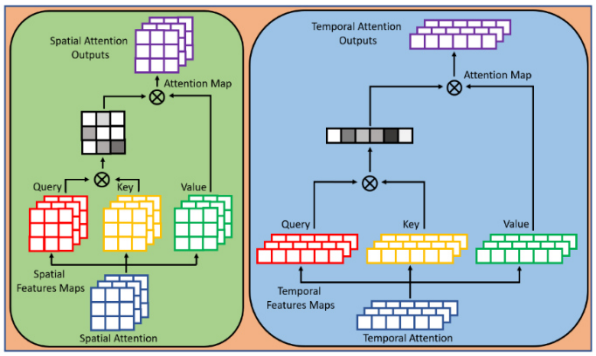


Figure 9. Hybrid attention.

The Squeeze-and-Excitation Block is widely adopted in convolutional neural networks to enhance the model's ability to extract significant image features. The Squeeze-and-Excitation Block layer functions

by assigning higher weights to important features while reducing the weights of less relevant ones. The Squeeze-and-Excitation Block operates through two main stages. The first stage, the "squeeze" stage, involves a global average pooling operation to generate descriptors or feature vectors representing the aggregate information from all feature channels. This stage reduces dimensionality and complexity, producing a more compact, yet informative, representation.

The next stage, the excitation stage, uses the feature vector generated from the squeeze stage and passes it through a series of fully connected layers, including a ReLU activation layer and a sigmoid layer. The ReLU layer enhances the representational capacity and flexibility in learning feature relationships. In contrast, the sigmoid layer produces weights or scalars that indicate the importance of each feature channel.

The Squeeze-and-Excitation Block shows the process flow from input to output within the Squeeze-and-Excitation Block layer. First, the input with dimensions HxWxC is fed into the global average pooling stage to generate feature descriptors. These descriptors then pass through two fully connected layers with ReLU and sigmoid activations, producing scalars representing each feature channel's significance. These scalars are subsequently used to scale the original features through a residual operation, resulting in an adjusted output.

By passing features through the Squeeze-and-Excitation Block, the model can adaptively select and focus attention on the most relevant and essential features within the image while disregarding less informative ones. This process helps the model obtain more robust and discriminative representations from the input data, ultimately improving performance in classification tasks. The Squeeze-and-Excitation Block is typically placed after the convolutional layer in a convolutional neural network architecture. This enables the model to effectively capture spatial-temporal features from images or sequences and apply more significant attention to the most critical features using the Squeeze-and-Excitation Block.

4. EXPERIMENT SETUP

Defining the experiments' scope within the context of developing a micro-expression recognition classification model is crucial for ensuring the model's effectiveness and generalizability. This is essential to guarantee that the experiments conducted are relevant to real-world conditions and can provide meaningful solutions to practical problems. Furthermore, the experiments' scope encompasses various scenarios and micro-expression variations, ensuring that the model can manage emotional differences' complexity. The scope also facilitates hyper-parameter optimization, enabling fine-tuning to achieve the most effective configuration.

Hyper-parameter tuning is conducted to identify the optimal combination of parameters that maximizes model performance, especially in micro-expression recognition. This involves experimenting with different parameters to find the configuration that delivers the most effective results. In this study, parameters such as batch sizes of 80 or 100 are selected to ensure comprehensive data processing, reducing the likelihood of missing critical patterns and helping prevent overfitting. Using 200 or 250 epochs allows for early monitoring of model performance, which helps avoid unnecessary overtraining. Data is divided into 80% for training, 10% for testing and 10% for validation in order to guarantee a thorough model evaluation.

Implementing the Adaptive Moment Estimation (ADAM) optimizer is essential for adjusting the learning rate in the intricate 3D convolutional neural network, resulting in faster and more stable convergence. The default learning rate of 0.001 balances stability and convergence speed. The Categorical Cross-Entropy loss function effectively manages multi-category classification tasks, ensuring precise facial micro-expression identification. This strategic combination of parameters and methods enhances the model's generalizability and improves its effectiveness in recognizing micro-expressions.

An experimental scenario is a series of plans and steps that are iteratively and alternately conducted to achieve the desired outcomes. This study's four experimental scenarios correspond to the classification models developed: Dual path-3DNet, Hybrid Attention-3DNet, SEBlock-3DNet and Hybrid Attention-3DSENet. Each scenario employs one dataset and applies image pre-processing, data pre-processing and variations in batch size and epoch. Each dataset undergoes 16 experiments, resulting in 64

experiments conducted across the four datasets in this study. A detailed explanation of each experimental scenario is provided in Table 1.

Table 1. Experimental scenarios.

Scenario	Dataset	Pre-processing	Augmentation	Class Weight	Batch Size	Epoch	Hybrid Att.	SE-Block
1	CAS(ME)^2	✓	✓	✓	80/100	200/250	×	×
	SMIC	✓	✓	✓	80/100	200/250	×	×
	SAMM	✓	✓	✓	80/100	200/250	×	×
	CASME II	✓	✓	✓	80/100	200/250	×	×
2	CAS(ME)^2	✓	✓	✓	80/100	200/250	✓	×
	SMIC	✓	✓	✓	80/100	200/250	✓	×
	SAMM	✓	✓	✓	80/100	200/250	✓	×
	CASME II	✓	✓	✓	80/100	200/250	✓	×
3	CAS(ME)^2	✓	✓	✓	80/100	200/250	×	✓
	SMIC	✓	✓	✓	80/100	200/250	×	✓
	SAMM	✓	✓	✓	80/100	200/250	×	✓
	CASME II	✓	✓	✓	80/100	200/250	×	✓
4	CAS(ME)^2	✓	✓	✓	80/100	200/250	✓	✓
	SMIC	✓	✓	✓	80/100	200/250	✓	✓
	SAMM	✓	✓	✓	80/100	200/250	✓	✓
	CASME II	✓	✓	✓	80/100	200/250	✓	✓

5. EVALUATION METRICS

In classification evaluation, specific metrics are used to evaluate a model's ability to predict the target class of a dataset. Essential terms include True Positives, False Positives, True Negatives and False Negatives. A True Positive (TP) occurs when the model correctly predicts a positive instance, aligning with the actual class. In essence, TP represents the count of positive samples accurately identified. A False Positive (FP), on the other hand, happens when the model incorrectly predicts a negative sample as positive, reflecting the number of negative instances misclassified as positive. True Negative (TN) refers to instances where the model correctly classifies a sample as negative, matching the actual class and representing the accurate identification of negative instances. Finally, a False Negative (FN) occurs when the model incorrectly predicts a positive sample as negative, signifying the number of positive instances mistakenly identified as negative.

Evaluation metrics are specific measures to gauge a deep-learning model's performance. Choosing the right metric is essential, as it influences the assessment of the model's ability to complete the task and helps compare the efficiency of various models.
Accuracy measures how well a classification model identifies the correct classes across the entire dataset [20]. It is determined by dividing the sum of true positives and true negatives by the total number of samples.

$$Accuracy = \frac{True\ Positives + True\ Negatives}{\text{Total Samples}} \quad (2)$$

The F1-score is a metric used to evaluate a model by balancing precision and recall. Precision measures the accuracy of the model's positive predictions, while recall evaluates how well the model detects all true positive instances.

$$F1 - score = 2\ \text{x}\ \frac{Precission x Recall}{Precission + Recall} \quad (3)$$

Recall assesses a model's capability to detect all actual positive instances. It is determined by dividing

the number of true positives by the total of true positives and false negatives. A high recall indicates that the model effectively captures the majority of positive cases within the dataset.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (4)$$

Precision measures how accurately a model predicts positive outcomes when they are indeed positive. It is calculated by dividing the true positives by the total number of predicted positive instances, which includes both true positives and false positives. A high precision score shows that the model makes few false positive predictions.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (5)$$

The error rate quantifies how often a model misclassifies data points. It is determined by dividing the sum of false positives and false negatives by the total number of samples. A lower error rate reflects that the model makes few classification mistakes.

$$Error\ Rate = \frac{False\ Positives + False\ Negatives}{Total\ Samples} \quad (6)$$

6. RESULTS AND DISCUSSION

The MER-DACWB3DCNNST model (Single-path 3DCNN) serves as the baseline in this study [18], offering a straightforward, yet practical, approach for micro-expression recognition by leveraging spatio-temporal features from a single input stream of facial images to enhance accuracy in detecting subtle micro-expressions. This model achieves strong results with a relatively simple structure while minimizing computational complexity. However, its primary limitation lies in its constrained ability to capture the depth of complex micro-expression features, particularly in cases where expression variations are incredibly subtle and rapid. This limitation can reduce the model's performance, detecting nuanced and fleeting emotional changes. Nonetheless, the model demonstrates robust accuracy and F1 score results across several micro-expression datasets, as shown in Table 2.

Table 2. Comparison of accuracy and F1-score of the MER-DACWB3DCNNST model (single-path 3DCNN) across different datasets.

Dataset	Accuracy (%)	F1-score
CAS(ME)²	92.75	0.9271
SMIC	91.49	0.9032
SAMM	92.20	0.9218
CASME II	93.66	0.9361

To enhance model performance, this study incorporated three additional components for testing: the Dual-path 3D CNN model, Hybrid Attention and Squeeze-and-Excitation Blocks, each utilizing dual input streams from the upper and lower facial regions. An ablation study was conducted to assess the impact of each component on model performance, using accuracy and F1-score as the primary metrics across multiple datasets. This analysis provides insights into the contribution of each component compared to the single-path baseline model.

Evaluation metrics such as accuracy and F1-score are calculated in each experimental stage. The presented graphs include information from all experimental scenarios, covering the dataset used, the application of image pre-processing and data pre-processing, batch size and epoch selection, as well as the accuracy and F1-score values. The type of graph presented is a line graph, which displays the highest accuracy and F1-score values for each experiment based on the dataset used. To provide detailed insights into the accuracy and F1-score calculations for each experimental scenario, the graphs are accompanied by a complete table of the experimental results. From these graphs, conclusions and analyses related to the obtained results can be drawn.

Figure 10 presents the accuracy and F1-score graphs for micro-expression recognition using four datasets: CAS(ME)², SMIC, SAMM and CASME II, with the proposed models: Dual Path-3DNet, Hybrid Attention-3DNet, Squeeze-and-Excitation-3DNet and Hybrid Attention-3D Squeeze-and-Excitation Net. From the graphs, Hybrid Attention-3DNet (HA-3DNet) model consistently achieves the highest accuracy across all datasets, with scores of 93.95% for CAS(ME)², 93.42% for SMIC, 93.61% for SAMM and 93.79% for CASME II. F1-scores are similarly high across datasets, with 0.9395 for

CAS(ME)², 0.9330 for SMIC, 0.9113 for SAMM and 0.9203 for CASME II.

The Dual-path 3DCNN structure performed better than the baseline Single-path model, particularly in capturing spatial-temporal features within different facial regions. The Dual-path approach enhances the model's sensitivity to subtle spatial-temporal dynamics by enabling separate pathways for upper and lower face regions. Results indicate a consistent accuracy improvement across all datasets, suggesting that the Dual-path structure provides more robust feature extraction.

Incorporating Hybrid Attention into the Dual-path 3DCNN model, which combines spatial and temporal attention mechanisms, further enhances the model's performance. This component allows the model to prioritize critical facial features like eyes and mouth while adapting to temporal variations. Experiments indicate that adding Hybrid Attention significantly boosts both accuracy and F1-score, highlighting its effectiveness in refining feature relevance. This component has proven especially effective on datasets with subtle, rapid expressions, confirming its essential role in distinguishing fleeting emotions.

Including Squeeze-and-Excitation Blocks in the Dual-path 3DCNN model enables adaptive reweighting of feature channels, allowing the model to highlight the most relevant features while downplaying less significant elements. This mechanism helps reduce noise in the micro-expression recognition process, particularly for complex expressions involving subtle changes across various facial regions. Squeeze-and-Excitation Blocks have proven effective in enhancing classification precision, as reflected in increased F1-scores across all datasets. The contribution of the Squeeze-and-Excitation Blocks is especially evident in recognizing expressions that require high sensitivity to specific features, delivering consistent and stable results in micro-expression classification.

Integrating Dual-path 3DCNN, Hybrid Attention and Squeeze-and-Excitation Blocks, the whole combination model achieves high accuracy and F1-scores across all datasets, indicating that each component contributes meaningfully to the model's overall performance. This combination offers robust feature extraction, adaptive focus and refined feature weighting. However, as shown in the result graphs, the model using only Hybrid Attention on Dual-path 3DCNN outperforms the whole combination, suggesting that the contribution of each component in this combination does not necessarily yield a more significant performance boost than Hybrid Attention alone.

Figure 10. Accuracy and F1-score graphs for micro-expression recognition using CAS(ME)², SMIC, SAMM, CASME II datasets and DP-3DNet, HA-3DNet, SE-3DNet, HA-3SENet models.

When comparing models, Hybrid Attention-3DNet consistently outperforms other models across all datasets, indicating that incorporating spatial and temporal attention mechanisms substantially enhances the model's capacity for recognizing micro-expressions. This improvement highlights the importance of capturing spatial and temporal dependencies, crucial for identifying subtle and rapid facial expressions.

Hybrid Attention, with its combined spatial and temporal attention mechanisms, effectively enhances micro-expression recognition accuracy. Our experiments demonstrate that Hybrid Attention improves accuracy and F1-score, particularly on the CAS(ME)² and SMIC datasets. This underscores the model's ability to capture expressions' complex spatial and temporal characteristics.

The spatial and temporal attention combination enables the model to detect subtle changes in expressions within very short durations, often challenging to identify in micro-expression videos. By focusing on critical facial features and temporal variations, Hybrid Attention improves the model's ability to differentiate emotions that appear briefly, yet convey meaningful information. Thus, incorporating Hybrid Attention significantly enhances micro-expression recognition, especially for fleeting expressions commonly found in micro-expressions.

Regarding the attention mechanism, the Hybrid Attention-3DNet architecture proves superior to Squeeze-and-Excitation-3DNet, suggesting that attention mechanisms offer more significant benefits in this context. This effectiveness likely stems from the specific nature of micro-expression data, where capturing temporal dynamics is essential. Although a combined model architecture with Hybrid Attention and Squeeze-and-Excitation was tested, results showed no substantial improvement over Hybrid Attention-3DNet alone, indicating that the main performance gain originates from the attention mechanism.

Dataset-performance variability across datasets indicates that, although Hybrid Attention-3DNet is highly reliable, dataset characteristics still influence its performance. The model consistently has high accuracy and F1 scores across datasets, but reflects good generalization capabilities. Additionally, pre-processing steps, such as data augmentation and class balancing, play an essential role**.** Data augmentation improves generalization by diversifying training samples, while class balancing prevents bias toward majority classes.

A visualization analysis was conducted on correct and incorrect classification results for specific micro-expressions, including "anger" and "fear," which often experience misclassification due to similar spatial and temporal patterns. This visualization highlights particular facial areas, such as the region around the eyes, that frequently cause misclassifications due to similar muscle movements in both expressions.

A case study on the "happiness" expression, which the model recognizes relatively quickly, was also performed. This recognition can be attributed to consistent spatial patterns around the mouth, aiding the model in differentiating this expression from others. This qualitative analysis provides insights into the model's strengths and weaknesses, particularly concerning subtle variations in micro-expressions.

For dataset performance, Hybrid Attention-3DNet achieved the highest accuracy with the SMIC dataset, likely due to multiple factors. SMIC has the largest sample size (about 164 video clips), enabling the model to learn and generalize patterns more effectively. Moreover, sample durations in SMIC vary significantly (from 9 to 343 seconds), allowing the model to capture spatial and temporal features. With a frame rate of 100 fps, the second highest among the datasets, SMIC provides ample spatial and temporal information, which is critical for classification. Its 640x480 resolution balances spatial detail with manageable data size.

Error analysis helps understand where and why the model makes mistakes. Based on the confusion matrix, some common errors in the CAS(ME)² dataset can be observed: for the Angry class, 7 Angry samples were classified as Disgust and 3 Angry samples were classified as Happy. For the Disgust class, 9 Disgust samples were classified as Angry and 5 Disgust samples were classified as Happy. For the Happy class, 6 Happy samples were classified as Angry and 7 Happy samples were classified as Disgust. This indicates confusion between particular classes, particularly between Angry and Disgust and between Happy and Disgust. A similar analysis for the other datasets can be found in Table 3.

In real-world applications, the developed micro-expression recognition model holds potential for various fields, including security, clinical psychology and human-computer interaction. For instance, accurately recognizing expressions of "fear" or "anger" can be crucial for detecting potential threats or

high-stress situations in security settings. However, the model's limitations in differentiating between similar micro-expressions, such as "fear" and "anger," could pose challenges in these applications, as misclassifying these expressions may impact critical decisions.

Table 3. Accuracy, F1-Score and Error Rate of the Hybrid Attention – 3DNet model with CAS(ME)², SMIC, SAMM and CASME II Datasets.

Dataset	Accuracy (%)	F1-Score	Error Rate (%)
CAS(ME)^2	93.95	0.9395	6.05
SMIC	93.42	0.9330	5.58
SAMM	93.61	0.9113	6.39
CASME II	93.79	0,9203	6.21

In clinical psychology, recognizing more apparent expressions like "happiness" or "sadness" offers applications for non-invasively assessing patients' emotional states. The model can assist in evaluating emotional responses to specific stimuli; however, subtle variations in micro-expressions could be missed, particularly when spatial-temporal patterns overlap between expressions of interest.

Another limitation in practical deployment is the model's sensitivity to the characteristics of the training dataset. Variations in lighting, facial angles or environmental conditions may affect the model's performance outside controlled laboratory settings. Additional data augmentation and adjustments for varying lighting conditions could be considered in the implementation phase, enhancing the model's reliability across diverse real-world situations. This discussion underscores the importance of further optimizing the model and conducting additional testing under real-world conditions to enhance its reliability in practical applications. It also highlights future development opportunities focused on adapting the model to a broader range of scenarios.

7. COMPARISON WITH PREVIOUS WORKS

Tables 4, 5, 6 and 7 compare the accuracy and F1-Score between the latest and proposed approaches using datasets including CAS(ME)², SMIC, SAMM and CASME II. These tables show that the proposed approach performs relatively better than the state-of-the-art methods. Enhancing the spatial-temporal feature weight attention in the 3D dual-path convolutional network model using hybrid attention and the Squeeze-and-Excitation Block improved the evaluation metrics for accuracy and F1-score.

Table 4. Comparison of accuracy and F1-score between the proposed method and state-of-the-art models on the CAS(ME)² dataset.

Year	Methods	Accuracy (%)	F1-score
2021	MSFME-IR [21]	-	0.8103
2021	RMER-3DCNN [9]	79.31	-
2021	LEARNET [22]	76.33	-
2022	MERASTC [23]	91.20	0.9070
2022	Deep3DCANN [24]	90.00	0.8800
2022	SE-DenseNet-T+EVM [25]	92.96	0.9289
2023	MER-DBNN [10]	-	0.8103
2024	Dual Path-3DNet	93.68	0.9358
2024	**Hybrid Attention-3DNet**	**93.95**	**0.9395**
2024	Squeeze-and-Excitation-3DNet	93.76	0.9368
2024	Hybrid Attention-3DSENet	93.88	0.9379

Note: '-' indicates that the data was not available in the referenced study.

Table 5. Comparison of accuracy and F1-score between the proposed method and state-of-the-art models on the SMIC dataset.

Year	Methods	Accuracy (%)	F1-score
2021	RMER-3DCNN [9]	76.92	-
2022	3DCNN-MED [26]	80.94	-
2022	MERASTC [23]	79.30	0.7900
2023	MER-DBNN [10]	-	0.6687
2023	BDCN [27]	-	0.7859
2023	RNAS MER [28]	-	0.7443
2023	FRL-DGT [29]	-	0.749
2023	DS-3DCNN [30]	78.78	0.7887
2024	Dual Path-3DNet	91.55	0.9147
2024	**Hybrid Attention-3DNet**	**93.42**	**0.9330**
2024	Squeeze-and-Excitation-3DNet	92.10	0.9198
2024	Hybrid Attention-3DSENet	92.72	0.9263

Table 6. Comparison of accuracy and F1-score between the proposed method and state-of-the-art models on the SAMM dataset.

Year	Methods	Accuracy (%)	F1-score
2021	RMER-3DCNN [9]	73.91	-
2022	MERASTC [23]	83.80	0.8440
2022	Deep3DCANN [24]	93.00	0.8900
2023	DBMNet [11]	-	0.6494
2023	BDCN [27]	-	0.8538
2023	RNAS MER [28]	-	0.7880
2023	ADMME [31]	81.43	0.8161
2023	FRL-DGT [29]	-	0.7580
2023	DS-3DCNN [30]	79.17	0.7156
2024	Dual Path-3DNet	92.76	0.9008
2024	**Hybrid Attention-3DNet**	**93.61**	**0.9113**
2024	Squeeze-and-Excitation-3DNet	93.35	0.9120
2024	Hybrid Attention-3DSENet	93.44	0.9138

Table 7. Comparison of accuracy and F1-Score between the proposed method and state-of-the-art models on the CASME II dataset

Year	Methods	Accuracy (%)	F1-score
2022	MERASTC [23]	85.40	0.8620
2022	Deep3DCANN [24]	86.00	0.8400
2022	SE-DenseNet-T+EVM [25]	82.74	0.7659
2023	DBMNet [11]	-	0.6653
2023	MER-DBNN [10]	-	0.8189
2023	BDCN [27]	-	0.9501
2023	STCPNet [12]	91.46	0.8977
2023	RNAS MER [28]	-	0.8985
2023	ADMME [31]	86.34	0.8635
2023	FRL-DGT [29]	-	0.9030
2024	Dual Path-3DNet	93.35	0.9115
2024	**Hybrid Attention-3DNet**	**93.79**	**0.9203**
2024	Squeeze-and-Excitation-3DNet	93.38	0.9112
2024	Hybrid Attention-3DSENet	93.52	0.9125

8. CONCLUSION

This study developed a pipeline with multiple processing stages to recognize spontaneous micro-expressions effectively. The results indicate that the proposed method surpasses state-of-the-art approaches, achieving accuracy and F1-score values of 93.95% and 0.9395 on the CAS(ME)² dataset, 93.42% and 0.9330 on SMIC, 93.61% and 0.9113 on SAMM and 93.79% and 0.9203 on CASME II. Among the datasets, the SMIC dataset exhibited the lowest error rate at 5.58%, followed by CAS(ME)² at 6.05%, CASME II at 6.21% and SAMM with the highest error rate of 6.39%. The differences in accuracy and F1-score values can be attributed to the distinct characteristics of each dataset, even when the same pipeline is applied. This study highlights that the implemented pipeline has successfully enhanced micro-expression recognition accuracy, primarily due to the improved attention to spatial-temporal feature weights.

ACKNOWLEDGEMENTS

The first author is a dedicated Telkom Foundation of Education employee, serving as a lecturer at the School of Electrical Engineering, Telkom University. He is advancing his academic career by pursuing a doctoral program at the School of Electrical Engineering and Informatics, Bandung Institute of Technology. Telkom University strongly supports this study, reflecting the institution's commitment to fostering academic growth and contributing to advancements in electrical engineering and informatics. The author's ongoing studies and this study project demonstrate a synergy between professional responsibilities and academic pursuits, aiming to contribute significantly to academia and industry.

REFERENCES

[1] P. Zhang, X. Ben, R. Yan, C. Wu and C. Guo, "Micro-expression Recognition System," Optik (Stuttgart), vol. 127, no. 3, pp. 1395–1400, DOI: 10.1016/j.ijleo.2015.10.217, 2016.

[2] G. Zhao and X. Li, "Automatic Micro-expression Analysis: Open Challenges," Frontiers in Psychology, vol. 10, no. AUG, pp. 1–4, DOI: 10.3389/fpsyg.2019.01833, 2019.

[3] L. Zhou, X. Shao and Q. Mao, "A Survey of Micro-expression Recognition," Image and Vision Computing, vol. 105, p. 104043, DOI: 10.1016/j.imavis.2020.104043, 2021.

[4] Y. He, S. J. Wang, J. Li and M. H. Yap, "Spotting Macro-and Micro-expression Intervals in Long Video Sequences," Proc. of the 2020 15th IEEE Int. Conf. Autom. Face Gesture Recognition (FG 2020), pp. 742–748, DOI: 10.1109/FG47880.2020.00036, 2020.

[5] F. Qu, S. J. Wang, W. J. Yan, H. Li, S. Wu and X. Fu, "CAS(ME)^2): A Database for Spontaneous Macro-expression and Micro-expression Spotting and Recognition," IEEE Transactions on Affective Computing, vol. 9, no. 4, pp. 424–436, DOI: 10.1109/TAFFC.2017.2654440, 2018.

[6] X. Li, P. Tomas, H. Xiaohua, Z. Guoying and P. Matti, "A Spontaneous Micro-expression Database: Inducement, Collection and Baseline," Proc. of the 2013 10th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG), DOI: 10.1109/FG.2013.6553717, Shanghai, China, 2013.

[7] A. K. Davison et al., "SAMM : A Spontaneous Micro-facial Movement Dataset," IEEE Transactions on Affective Computing, vol. 9, no. 1, pp. 116–129, DOI: 10.1109/TAFFC.2016.2573832, 2016.

[8] W. J. Yan et al., "CASME II: An Improved Spontaneous Micro-expression Database and the Baseline Evaluation," PLoS One, vol. 9, no. 1, pp. 1–8, DOI: 10.1371/journal.pone.0086041, 2014.

[9] Y. Jiao, M. Jing, Y. Hu and K. Sun, "Research on a Micro-expression Recognition Algorithm Based on 3D-CNN," Proc. of the 2021 3rd Int. Conf. Intell. Control. Meas. Signal Process. Intell. Oil Field (ICMSP 2021), no. Icmsp, pp. 221–225, DOI: 10.1109/ICMSP53480.2021.9513351, 2021.

[10] F. Guowen and L. Xi, "Micro-expression Recognition Based on Dual Branch Neural Network," Proc. of the 2023 Int. Conf. Artif. Intell. Comput. Inf. Technol. (AICIT 2023), no. 2020, pp. 2–5, DOI: 10.1109/AICIT59054.2023.10278020, 2023.

[11] Q. Zhou, S. Liu, Y. Wang and J. Wang, "Divided Block Multi-scale Convolutional Network for Micro-expression Recognition," Proc. of the 2022 1st Int. Conf. Cyber-Energy Syst. Intell. Energy (ICCSIE 2022), pp. 1–5, DOI: 10.1109/ICCSIE55183.2023.10175242, 2023.

[12] Z. Shang, J. Liu and X. Li, "Micro-expression Recognition Based on Spatio-temporal Capsule Network," IEEE Access, vol. 11, no. January, pp. 13704–13713, DOI: 10.1109/ACCESS.2023.3242871, 2023.

[13] H. Tian, W. Gong, W. Li and Y. Qian, "PASTFNet: A Paralleled Attention Spatio-temporal Fusion Network for Micro-expression Recognition," Medical and Biological Engineering and Computing, DOI: 10.1007/s11517-024-03041-y, 2024.

[14] A. J. Rakesh Kumar and B. Bhanu, "Relational Edge-node Graph Attention Network for Classification of Micro-expressions," Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 2023-June, pp. 5819–5828, DOI: 10.1109/CVPRW59228.2023.00618, 2023.

[15] F. Li, P. Nie, M. You, Z. Chen and G. Wang, "Triple-ATFME: Triple-branch Attention Fusion Network for Micro-expression Recognition," Arabian J. for Science and Eng., DOI: 10.1007/s13369-024-08973-z, 2024.

[16] H. Insan, S. S. Prasetiyowati and Y. Sibaroni, "SMOTE-LOF and Borderline-SMOTE Performance to Overcome Imbalanced Data and Outliers on Classification," Proc. of the 2023 3rd Int. Conf. Intell. Cybern. Technol. Appl., pp. 136–141, DOI: 10.1109/icicyta60173.2023.10428902, 2024.

[17] A. Sagoolmuang, "Power-weighted kNN Classification for Handling Class Imbalanced Problem," Proc. of the 2021 2nd Int. Conf. Big Data Anal. Pract. (IBDAP 2021), pp. 42–47, DOI: 10.1109/IBDAP52511.2021.9552164, 2021.

[18] B. Irawan, N. P. Utama, R. Munir and A. Purwarianti, "Improving the Accuracy of Facial Micro-expression Rrecognition: Spatio-temporal Deep Learning with Enhanced Data Augmentation and Class Balancing," Interdisciplinary Journal of Information, Knowledge, and Management, vol. 19, pp. 1–15, DOI: 10.28945/5386, 2024.

[19] Y. Zhou, H. Chen, J. Li, Y. Wu, J. Wu and L. Chen, "ST-Attn: Spatial-," Proc. of the IEEE Int. Conf. Data Min. Work. (ICDMW), vol. 2019-Novem, no. November, pp. 609–614, DOI: 10.1109/ICDMW.2019.00092, 2019.

[20] Y. S. Gan, S. E. Lien, Y. C. Chiang and S. T. Liong, "LAENet for Micro-expression Recognition," Visual Computer, vol. 40, no. 2, pp. 585–599, DOI: 10.1007/s00371-023-02803-3, 2024.

[21] P. Sharma, S. Coleman, P. Yogarajah, L. Taggart and P. Samarasinghe, "Magnifying Spontaneous Facial Micro Expressions for Improved Recognition," Proc. of the Int. Conf. Pattern Recognit., pp. 7930–7936, DOI: 10.1109/ICPR48806.2021.9412585, 2020.

[22] M. Verma, S. K. Vipparthi, G. Singh and S. Murala, "LEARNet: Dynamic Imaging Network for Micro Expression Recognition," IEEE Transactions on Image Processing, vol. 29, no. c, pp. 1618–1627, DOI: 10.1109/TIP.2019.2912358, 2020.

[23] P. Gupta, "MERASTC: Micro-expression Recognition Using Effective Feature Encodings and 2D Convolutional Neural Network," IEEE Transactions on Affective Computing, vol. 14, no. 2, pp. 1431–1441, DOI: 10.1109/TAFFC.2021.3061967, 2023.

[24] S. Thuseethan, S. Rajasegarar and J. Yearwood, "Deep3DCANN : A Deep 3DCNN-ANN Framework for Spontaneous Micro-expression Recognition," Information Sciences (Ny), vol. 630, no. November 2022, pp. 341–355, DOI: 10.1016/j.ins.2022.11.113, 2023.

[25] L. Cai, H. Li, W. Dong and H. Fang, "Micro-expression Recognition Using 3D DenseNet Fused Squeeze-and-excitation Networks," Applied Soft Computing, vol. 119, p. 108594, DOI: 10.1016/j.asoc.2022.108594, 2022.

[26] W. S. P. Bayu and A. Setyanto, "3D CNN for Micro Expression Detection," Proc. of the 5th Int. Conf. Inf. Commun. Technol. A New W. to Make AI Useful Everyone New Norm. Era (ICOIACT 2022), pp. 397–401, DOI: 10.1109/ICOIACT55506.2022.9972194, 2022.

[27] B. Chen, K. H. Liu, Y. Xu, Q. Q. Wu and J. F. Yao, "Block Division Convolutional Network with Implicit Deep Features Augmentation for Micro-expression Recognition," IEEE Transactions on Multimedia, vol. 25, pp. 1345–1358, DOI: 10.1109/TMM.2022.3141616, 2023.

[28] M. Verma, P. Lubal, S. K. Vipparthi and M. Abdel-Mottaleb, "RNAS-MER: A Refined Neural Architecture Search with Hybrid Spatiotemporal Operations for Micro-expression Recognition," Proc. of the 2023 IEEE Winter Conf. Appl. Comput. Vision (WACV 2023), pp. 4759–4768, DOI: 10.1109/WACV56688.2023.00475, 2023.

[29] Z. Zhai and J. Zhao, "Feature Representation Learning with Adaptive Displacement Generation and Transformer Fusion for Micro-expression Recognition," Proc. of the 2023 IEEE/CVF Conf. Comput. Vis. Pattern Recognit., pp. 22086–22095, DOI: 10.1109/CVPR52729.2023.02115, 2023.

[30] Z. Li, Y. Zhang, H. Xing and K.-L. Chan, "Facial Micro-expression Recognition Using Double-Stream 3D Convolutional Neural Network with Domain Adaptation," Sensors, vol. 23, no. 7, DOI: 10.3390/s23073577, 2023.

[31] Y. Wang, H. U. Shi and R. Wang, "Action Decouple Multi-tasking for Micro-expression Recognition," IEEE Access, vol. 11, no. June, pp. 82978–82988, DOI: 10.1109/ACCESS.2023.3301950, 2023.

ملخص البحث:

تقترح هذه الورقة البحثية نموذجاً مبتكراً للتعرّف إلى تعبيرات الوجْه الدّقيقة باستخدام شبكةٍ عصبيةٍ التفافية محسّنة عن طريق الاهتمام الهجين وكُتَل الانضغاط والإثارة. وتتلخّص الأهداف الرئيسية للنّموذج في: (1) تحسين استخلاص السِّمات باستخدام بِنية الشّبكة العصبية الالتفافية، (2) تحسين تمثيل البيانات عبر زيادة الصُّوَر على نحوٍ هادفٍ وتوزيع سِمات الصُّوَر بشكلٍ متوازنٍ، (3) تحسين انْدماج السِّمات باستخدام تقنيات الشّبكات الدّارجة في أدبيات الموضوع.

لقد تمّ إجراء تجارب للنّموذج المقترح على عددٍ من مجموعات البيانات. وقد برهن نموذج الشّبكة ثلاثية الأبعاد مع الاهتمام الهجين على قيمٍ متفوّقةٍ للدّقّة على جميع مجموعات البيانات المستخدمة مقارنةً بالنّماذج المماثلة المستخدمة في دراساتٍ سابقة أخرى على مجموعات البيانات ذاتها. وتدلُّ النّتائج الّتي تمّ الحصول عليها على أنّ النّموذج المقترح يمتاز بالفاعلية والمتانة، مع إمكانية استخدامه في مدى واسعٍ من الاستخدامات المتعلّقة بتمييز التّعبيرات الدّقيقة للوجْه، مثل الغَضَب والخَوفً والسّعادة وغيرها.

المجلة الأردنية للحاسوب وتكنولوجيا المعلومات (JJCIT) مجلة علمية عالمية متخصصة محكمة تنشر الأوراق البحثية الأصيلة عالية المستوى في جميع الجوانب والتقنيات المتعلقة بمجالات تكنولوجيا وهندسة الحاسوب والاتصالات وتكنولوجيا المعلومات.

تحتضن وتنشر جامعة الأميرة سمية للتكنولوجيا (PSUT) المجلة الأردنية للحاسوب وتكنولوجيا المعلومات، وهي تصدر بدعم من صندوق دعم البحث العلمي في الأردن. وللباحثين الحق في قراءة كامل نصوص الأوراق البحثية المنشورة في المجلة وطباعتها وتوزيعها والبحث عنها وتنزيلها وتصويرها والوصول إليها. وتسمح المجلة بالنسخ من الأوراق المنشورة، لكن مع الإشارة إلى المصدر.

الأهداف والمجال

تهدف المجلة الأردنية للحاسوب وتكنولوجيا المعلومات (JJCIT) إلى نشر آخر التطورات في شكل أوراق بحثية أصيلة وبحوث مراجعة في جميع المجالات المتعلقة بالاتصالات وهندسة الحاسوب وتكنولوجيا المعلومات وجعلها متاحة للباحثين في شتى أرجاء العالم. وتركز المجلة على موضوعات تشمل على سبيل المثال لا الحصر: هندسة الحاسوب وشبكات الاتصالات وعلوم الحاسوب ونظم المعلومات وتكنولوجيا المعلومات وتطبيقاتها.

الفهرسة

المجلة الأردنية للحاسوب وتكنولوجيا المعلومات مفهرسة في كل من:

فريق دعم هيئة التحرير

عنوان المجلة

الموقع الإلكتروني: www.jjcit.org

البريد الإلكتروني: jjcit@psut.edu.jo

العنوان: جامعة الاميرة سمية للتكنولوجيا، شارع خليل السكاكت، الجبيهة، عمان، الأردن.

صندوق بريد: 1438 عمان 11941 الأردن

هاتف: 962-6-5359949+

فاكس: 962-6-7295534+

جامعة الأميرة سميّة للتكنولوجيا
Princess Sumaya University for Technology

صندوق دعم البحث العلمي والابتكار
Scientific Research and Innovation Support Fund

المجلة الأردنية للحاسوب وتكنولوجيا المعلومات

آذار ٢٠٢٥ | المجلد ١١ | العدد ١ | ISSN 2415 - 1076 (Online) ISSN 2413 - 9351 (Print)

JJCIT

الصفحات	عنوان البحث
١- ١٥	تنظيم مجموعات البيانات من أجل تحسين تصنيف برامج التّجسُّس موسومي أحمد ميمي، هو نج، و تيموثي تزن فون ياب
١٦- ٣٢	إطار عملٍ مبتكرٌ للتّصفية التّعاونية المستندة إلى الأدلة بالاعتماد على تجاهل التّفضيلات المتضاربة خديجة بلمسوس، فوزي سباك، و محمد مطاوي
٣٣ - ٥٣	أحدث ما توصلت إليه تكنولوجيا التّعلُّم الآلي في اضطراب النّموّ العصبي: دراسة منهجية ليليان لي ين واي، أج أصري أج إبراهيم، و راينر ألفرد
٥٤ - ٧٢	طريقة جديدة تجمع بين خوارزمية RSA و خوارزمية الجمل (ElGamal): تقدم في تقنيات التّشفير والتّواقيع الرّقمية باستخدام الاعداد الصّحيحة لِ (غاوْس) يحيى عواد، دعاء جمعة، يوسف الخزّي، و رامز هندي
٧٣ - ٨٤	تحسين مُعدّل الإنذارات الكاذبة ومُعدّل الكشْف الخاطىء لجُهد العَتَبة المطلوب في الاتّصالات التّعاونية ساتيش كومار جانامانيني، و جبندو سيخار روري
٨٥ - ٩٩	الكشْف عن الملاريا مع مراعاة الخصوصية: نموذجٌ مع عدم الكشْف عن الهوية لتحليل الصُّور السّرّية غزالة حسيني، و إيمان جدي
١٠٠ - ١١٦	المصادقة على الأجهزة المستندة إلى تقنية سلاسل الكُتل في حوسبة الحافَة باستخدام النّهج الكمّي فينياك أ. تلسانج، ماهباليشوار س. ككاساجيري، و أنيل د. ديفانجافي
١١٧ - ١٣٥	تعزيز التّعرُّف إلى التّعبيرات الدّقيقة للوجْه: نهج جديد قائم على الشّبكات ثلاثية الأبعاد ذات الاهتمام الهجين بودي إيراوان، رينالدي منير، نوجراها بريفا يوتاما، و أيو بورواريانتي

www.jjcit.org

jjcit@psut.edu.jo

مجلة علمية عالمية متخصصة تصدر
بدعم من صندوق دعم البحث العلمي والابتكار