Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Deep Learning Framework for DDoS Intrusion Detection in IoMT Networks: Combining CNN, GRU, and CatBoost Classifier

Jerlin George[1], Megha Gautam[2], R. Chitra[3], G. Naveen Sundar[4*]
[1]*Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Tamil Nadu, India*
jerlingeorge@karunya.edu.in
[2]*Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Tamil Nadu, India*
meghagautam@karunya.edu.in
[3]*Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Tamil Nadu, India*
chitrar@karunya.edu
[4*]*Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Tamil Nadu, India*
naveensundar@karunya.edu
Corresponding Author: naveensundar@karunya.edu

ARTICLE INFO

Received: 24 Dec 2024

Revised: 12 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

In recent years, there are numerous malicious attack on communication and commercial services. The Intrusion Detection System (IDS) predict the anamolies and threats happens on the network with high accuracy. The conventional IDS faces more challenges in feature extraction and failed to address the spatial and temporal dependencies which inherent in medical data. In this paper a hybrid IDS framework is developed where integrated with Convolution Neural Networks (CNN), Gated Recurrent Units (GRU) for feature extraction with CatBoost for classification. The KDDCup 1999 dataset is utilized for simulates the intrusion and normal network behavior in medical network environment. The CNN capture the spatial correlation and GRU process the sequential layer which analysis the dynamic and multi-dimensional intrusion patterns. These features are fed into the CatBoost Classifier and predict the categories into normal or malicious network behavior. The performance metrics of this proposed approach is evaluated and find the accuracy, precision, recall, and F1-Score (99%) which is higher than existing approach. Therefore, this approach will be scalable and reliable solution for enhancing the network security and avoid cyber threats.

Keywords: CNN, GRU, IDS, CatBoost Classifier, KDDCup 1999 dataset

I. INTRODUCTION

The evolution of network technologies and the emergence of internet-connected digital gadgets such as Internet of Things (IoT) have facilitated access to information and communication with others. This accessibility significantly influences the economic and social growth of day-to-day life and excites cyber-criminals to disrupt the delivery of various services to customers. Most commonly the attack detection technique secure the network from this Intrusion Detections (ID) thereby it scan network traffics of malicious activity. Over the year, the ID's development is tremendous on developing a secure network for users. These rapid growth support system to handle diverse tasks carried out in network and make a complexity in securing each nodes from the attack. The interconnected network nodes transfer critical data among these network which has to secure and avoid malicious attack over the nodes. The cybercriminals or threats are raising numerously in new attack vectors which can't be handle by the traditional approaches. These networks are vulnerable to security breaches and cause a catastrophic implications for an organization. This breaches not only leads to financial losses but also leakage of sensitive information and damage the organization reputation, erode the customer trust in marketplaces. There are several attacks such as MitM, DDOS, Packet Sniffing, IP Spoofing, ARP spoofing, Port Scanning etc.. Let us discuss about basic attacks how it used to affect

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

the network privacy, MitM is called as Man in the middle attack where the attacker intercepting the login credential using the unencrypted Wi-Fi communication. The tools like Wireshark is used to sniff the unencrypted FTP or HTTP to extract the sensitive data such as session token or financial data. In IP Spoofing they send the fake packets to trick the channel to extract the data from it via attacking the bypass firewalls. Address Spoofing where the request is redirect to the attacker's MAC address with the IP address of those legitimate device. In DNS, the direct cache entries are corrupted and make that as phishing site.

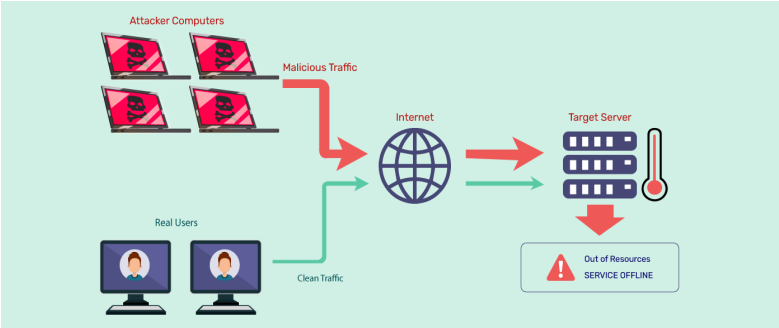

Fig. 1 Working of DDOS attack

These cyber disruptions are becoming a security concern, and the most prevalent threat that internet services encounter is Distributed Denial of Service (DDoS) attacks. DDoS assaults are launched to slow the performance of an application, deplete network resources, and overload online services by bombarding them with enormous traffic generated from multiple sources, as a result, the end-user cannot obtain essential information or response. Consequently, the CPU, network bandwidth, and memory become overwhelmed. Since the attack originates from multiple sources, it isn't easy to pinpoint from where it is originated. Similarly, if the malicious requests are generated from a single source, it is referred as a Denial of Service (DoS) attack. The objective of this paper is to secure the IOT data and health data across the cloud due to severe DDOS attacks and threats over network. The goal is to develop a efficient IDS for securing the IoMT environment in all phases of execution process.

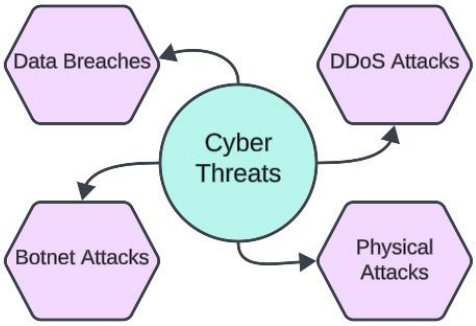

Fig. 2 Cyber Threats

The COVID-19 pandemic's challenges, depends on the digital technology to support healthcare professionals' mental health remotely [1,2], accelerated up this change and led to an malicious attacks on the Health sector [3]. Biosensors, such as blood pressure, temperature, motion, breathing, and vision sensors, are essential for linking people to healthcare systems [4,5]. The IOT sensor which is connected with patient produce large volumes of data because of it's real time analysis which support the doctor to observe the patient current state. There is exponential expansion in the IoMT landscape. By 2032, the worldwide Internet of Medical Things (IoMT) industry is projected to have grown from its 2022 valuation of $48.7 billion [6] to $370.9 billion. Normally the intruder used to provide some false request which is enough to attack the security of the network or device. In this system Deep Learning (DL) approach is utilised to overcome this threats and sequentially all the request is evaluated and provide a necessary response.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

This research focused on efficiently designing an anomaly-based detection model that identifies unknown attack patterns. Among various DDoS attacks, the volumetric flood technique has remained the most frequent in recent years. These are easy and successful, as they do not require a vulnerability (a compromised third-party system or technical experience). Since DDoS tools or publicly available attack services can launch an assault that sends more traffic to the victim than their network bandwidth can handle. The attacker's IP addresses are hidden when used with UDP reflection attacks. As a result, DDoS attacks will become the preferred method for many malicious actors. So these security threat to IoMT system make a numerous issues such as false diagnosis that affects the patient's data privacy or changing any IOT data may leads to death.

Computational statistics and mathematical optimization are closely linked fields that frequently intersect with machine learning (ML) [8]. For last decade, the Machine Learning (ML) approach provide a efficient response to the cybersecurity issues using some hybrid networks. These networks handle both anomaly and false detection where the usual activity on the network is traced and provide a potential to find the intruder. The most promising method to combat previously zero-day attacks handle via ML and provide a security to IoMT network in health care system [9]. By merely keeping on data tampering or identifying shifts in the traffic characteristics of the network, it can detect assaults.

The existing [10] approach concentrates on filter-based feature selection methods such as XGBoost and Mutual Information (MI), which could miss temporal and spatial connections that are important in medical network data (such as IoMT settings). The detection accuracy for intricate attack patterns may be decreased if feature selection relies too heavily on statistical correlation, which could lead to the loss of important context-sensitive information. Only binary categorization (malicious vs. normal) is supported by the model. Multi-class categorization, which is essential for differentiating between different attack types (such as DDoS, data manipulation, or phishing) and implementing focused defences, is not covered. The CICIDS2017 dataset, which is used to assess the system, is extensive but might not accurately capture the complexity and dynamic nature of actual IoMT environments. The CICIDS2017 dataset evaluation has a low false alarm rate (FAR), however the research doesn't address performance in more complicated scenarios or with different kinds of attacks, where tree-based models may produce more false positives. A new model called XSRU-IoMT [11] was put forth in the field of explainable AI (XAI) to identify complex attack vectors in IoMT networks. The bidirectional Simple Recurrent Unit (SRUs) is utilised which vanish the gradient issues and speed up the training process with skip connection. But this study lack of providing information about the number of epoch ,batch size utilised to affect the model's accuracy and it's majorly focus on the justifying the prediction based on the assumptions.

One popular method is to run patient biometric data and network flow data through multiple hidden deep learning layers [12,13]. For the best features from the temporal and spatial aspects of DL, this method uses a global attention layer. It also uses a cost-sensitive learning strategy to rectify unbalance data [14]. This approach didn't address the difficulties of fixed epoch and batch size utilised in deep learning algorithm.

A swarm-neural network-based approach for IDS in IoMT systems is presented in another paper [15]. This approach recognizes that because IoMT devices have limited storage and processing power, there are sensitive data which is vulnerable to attack due to unauthorised access.

To solve these issues, have implemented a novel hybrid IDS framework for IoMT networks. This framework is integrated with CNN, GRU and CatBoost classification algorithm.

1. Utilised KDDCup 1999 Dataset

   The proposed system simulates intrusion and typical network behavior in an IoMT medical network context by utilizing the KDDCup 1999 dataset. s

2. Feature Extraction with CNN and GRU

   The CNN component finds static intrusion patterns by capturing spatial correlations in network traffic data. In order to properly handle temporal dependencies, the GRU analyzes dynamic and multi-dimensional incursion patterns while processing consecutive data layers.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

3. CatBoost Classification

   The CatBoost classifier receives the extracted features and uses them to predict with high accuracy and few false alarms whether the network behavior is malicious or normal. A thorough evaluation is conducted of the system's performance parameters, which include accuracy, precision, recall, and F1-score. The findings show that the suggested framework performs noticeably better than current methods, obtaining greater detection rates and lowering computational overhead. In the IoMT ecosystem, this hybrid strategy is scalable and provides a dependable way to improve network security and prevent cyberattacks.

In order to highlight current issues and research gaps, the study starts by examining the body of research on intrusion attacks in networks. The proposed hybrid CNN-GRU system's originality is then highlighted, along with its design and special method of resolving these problems. The performance and benefits of the system are thoroughly examined in the presentation of the implementation procedure and outcomes that follows. The report ends with a discussion of the findings' significance and recommendations for future research aimed at improving security in IoMT networks.

II. LITERATURE SURVEY

An enhanced evolutionary algorithm called TA-MaEA is presented in the study presented by Cao et al. [16] for hybrid microgrid system optimization, with an emphasis on cost, dependability, carbon emissions, and power source. The system expenses are much lower with this method than with current algorithms. A new heterogeneous temporal anomaly reconstruction GAN (HTA-GAN) was presented by Chen et al [17] which improves data quality and provides effective anomaly detection in IoT. Lie et al [18] implemented a dynamic event triggering protocol for online prediction which secure the network address and provide a stability in performance. It also dropout packets which are consider as cyber-attack. The Saheed et.al. [19] analysis various ML algorithm over IoT and tried to tackle the privacy and security concern over internet. The UNSW-NB15 dataset is used to stimulate the attacks and enhance the IoT security via different Machine learning algorithm such as NB, SVM, XGBoost, CatBoost KNN. Compare to these algorithms CatBoost provide a higher accuracy in training model and prediction.

Omuya et.al.[20] presented text classification is negatively impacted by the redundant and irrelevant characteristics of the text corpus. A hybrid filter-based feature selection technique that blends principal component analysis and IG was introduced. During their analysis, they found that their proposed feature selection technique significantly reduces the data dimension by choosing a suitable feature subset and reducing training time.

Thamilarasu et al. [21] concern on IoMT network such as Wireless Body Area Networks (WBANs), sensors, and other devices which are affected by IDS. This study simulated a network topology of hospital and launch the device level and network intrusion detection. This approach successfully detect the device at 98% and network 99%.

To specifically identify anomalies in heart rate data, Šabić et al. [22] developed an anomaly identification approach. An evaluation of data fit was conducted using five algorithms, both supervised and unsupervised. In their work, Random Forests and ensemble algorithms perform particularly well in modeling such systems, achieving over 99% accuracy, suggesting that they can be used for reliably identifying anomalous heart rate data. But it has a issues in handling larger hospital with diverse sensors and users where this simulation become complex. If dynamic condition such as heterogenous thread or unpredictable network traffic cause this simulation failure. The computational and communication overhead in real time cause big issues it may cause a delay also which leads to serious consequences in hospital.

Hady et al.,[23] developed an advance Health security system which is designed to analyse the unauthorised data and threats happening on the health care application. Nearly 14000 non-attack and 2000 attack samples are utilised for training the model. There are four different algorithms are utilised for prediction such as SVM, KNN,ANN, RF. ANN perform well and achieve 25% improvement than other algorithms. The major drawback of these implementation is imbalance dataset 7:1 ratio potentially leads to false negative rate and identify as attack. Similarly the dataset may limited with diverse threat type which can't detect the emergency security treats.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

He. et al. [24], proposed an IDS on Stacked AutoEncoder SAE from anomaly detection where the normalizing and discretizing the data before extracting features using a support analysis elicitation (SAE). The retrieved attributes are subsequently fed into machine learning algorithms for classification. Using the patient's data some attacks are simulated and trained with SVM, NB, KNN, and XGBoost. The XGBoost performance better than other algorithms which provide 97.83% accuracy, 2.35% FPR, and 1.65% FNR.

Ahmed et al. [25] created and released a dataset known as the ECU-IoHT. Cyberattacks targeting a variety of vulnerabilities were launched against this dataset. These attacks were intended to collect data on attack patterns and aid in the development of efficient Défense strategies. Their study found that the KNN algorithm performed better in anomaly identification than both kernel-based algorithms and statistical clustering.

Adel et al [26],ML algorithms like ANN, KNN, NB, SVM are all included in the research examination. Training and testing of all ML models under consideration were conducted on the publicly available benchmark dataset, Bot-IoT, which contains a wide range of threats. The developed system also compared the performance of machine learning methods in terms of their ability to detect intrusions in IoMT networks using a wide range of evaluation indicators. Spatial and temporal connections in network traffic data are not automatically modeled by traditional machine learning algorithms. They are less useful for identifying dynamic intrusion patterns in real-time IoMT environments because of this restriction.

Rbah et al[27],introduce different ML algorithm and study about the IDS taxonomy with its data source to get an common idea on ID's. This provide an idea for understanding the IDS application for different data sources including packets, logs, sessions and flows. This study mainly focuse on the selection of right source and data for characteristic the attack type. This logs holds the semantic information about attacks such as R2L, U2R, and SQL. The R2L attacks concentrate on the unauthorised access on target machine remotely. The U2R attacks exploits vulnerability to the root access. The SQL injection affected with malicious queries with targeted database system.

Additionally, packets offer conversation information that can be used to identify R2L and U2L attacks. The entire network environment is represented by Flow, which is able to identify DOS and probe attacks. Sessions, which show client-server communication, can be used to identify Trojan, R2L, U2L, and tunnel attacks. The research focuses on machine learning methods, particularly deep learning algorithms, and application scenarios for IDSs that use these various data types. The biggest challenge might be the lack of available datasets. Thus, there are many opportunities for development with unsupervised learning and progressive learning methodologies.

The articles by I. Idrissi et al. [28], which claim that IoMT infiltration is increasing. These research concentrate on improving the IDS's performance through development. These articles, however, underutilize the enhanced detection rate. However, the detection rate and delay have not been examined in these works.

R. Chitra [29], Based on XGBoost, an IoMT malware detection system developed achieves a 97% accuracy rate. However, after creating the classifier, this study did not provide a reaction mechanism to take the necessary action. In order to optimize this classifier, the Genetic Algorithm (GA) is used, which optimizes the learning process. They didn't optimize the detection rate to lower the false alarm rate. There was no optimization of detection rate to reduce false alarms. Using Gradient-Boosted-Trees, W. Lu created a 95.4% accurate intrusion detection system based on IoMT [30]. However, the system's dependability is called into question because it was trained using just 11 features. A group of different biosensors make up the IoMT system. Healthcare professionals utilize these sensors to keep an eye on their patients' health. For IoMT networks, the authors [32] created the weighted majority algorithm (MOA-WMA), which is based on multi-objective optimization. Unfortunately, the classification performance of this optimization model was poor.

Karan et al. [33] address the issues of IoMT devices with cyberthreats which widely affects the patient privacy and harm their health by altering the sensor or medical equipment values. Many Lives are dependent on these IoMT devices which has to secure with robust model for network integrity. The author concentrate on the tree classifier for anomalies detection which 94.23% accuracy. Optimising the dimension of the input data improves the acceleration of the model. Despites it strength, the model have sever drawbacks such as the practical implementation of these model is complex , the tree classifier not perform well in handling different IoMT attacks and it may become efficient only

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

with integration of any new neural networks or ensemble classifier. The dimensionality reduction may loss the critical feature which may play a potential role in detection the nuanced attacks.

According to [34], the authors developed a Lyapunov Exponent Analysis and Echo State Network (LEAESN) using the Darpa98 dataset and recurrent neural echo state networks (SCESN).Furthermore, the explainable recurrent units (XSRU) were created by the authors of [35] to forecast attacks in the context of IoMT. When it came to identifying MitM attacks, the XSRU-IoMT framework outperformed DoS attacks. According to the report, different deep learning and optimization techniques were frequently used in classical DDoS attack detection approaches. Nevertheless, these approaches lacked security procedures or preventative steps.

However, the inadequate design and lack of authentication in IoMT system significant the vulnerability in networks. During the medical data transaction from sensor to fog layer to cloud there may be an intruder who can exploit these network and make it weakness for various cyber attack such as spoofing, Ransome attack, DoS, and Jamming. In Figure 1.2 depicted the cyber-attack scenarios contain biosensors. Weak security in the biosensor authentication process can be targeted by an attacker. An adversary could therefore take advantage of the security holes in the IoMT environment. This may have an effect on two important healthcare system assets.

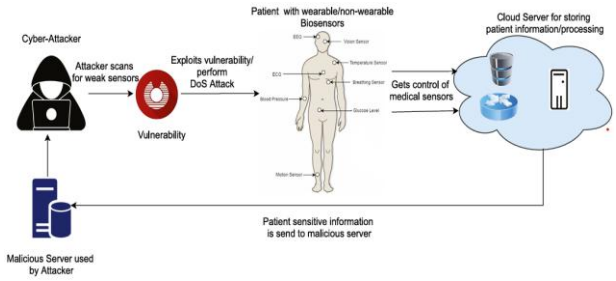

Fig. 3 Attack Scenario from attacker

If a malicious user manages to turn off vital medical devices that are currently in use, or if drug inventory systems are compromised or the operating system has a power outage, it could have a negative impact on a patient's health. The most valuable resources in hospitals are medical records. These records contain a variety of important patient data, including name, address,, DOB, bank details such as debit or credit card, user identifiable information, and information about health care providers. Due to these cyberattack, the intruder or unauthorised person can easily access this patient data. The proposed system abridge the gaps and provide better Intrusion detection system for IoMT network.

III. ARCHITECTURE SYSTEM

A Lightweight Hybrid approach is proposed to focus on the intrusion detection accuracy for securing the IoMT data. Feature extraction is performed using CatBoost for classification, with Convolutional Neural Networks (CNN) and Gated Recurrent Units (GRU) used for classification. The KDDCup 1999 [31] dataset was used to train and testing where binary classification is carried whether it's malicious or not. Among this dataset, 80 percent is used for training, and 20 percent is used for testing. Models are built and optimised using this training set, while models are validated using unseen data. To obtain better result the raw data undergoes preprocessing steps where Pandas Library is used for Dataframe of CSV files. The missing values and normalisations are done properly for practical applications.

A. Dataset

KDDCup provides header information for IP packets, ICMP/TCP/UDP segments and overall packet size, which is more relevant to network traffic analysis and intrusion detection. This IP packet header information provide details about the source and destination of the packets, protocols and flag details. The size of IP Packet help to identify the traffic pattern that deviate from the normal behaviours. For ICMP packets headers contain the echo request and replies which is used for network testing and exploited the ping floods or other attacks.

Compared to content feature extraction, header information extraction is far simpler. Reassembling data streams is a computationally and memory-intensive step in conducting in-depth packet data analysis. Furthermore, domain

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

knowledge which must be supplied by a human expert is typically needed for data analysis. The extraction and analysis of content features on a large scale can only be performed on a relatively inexpensive, high-speed network infrastructure that supports real-time, high-speed communication. Using specially designed training sets from the KDD Cup '99 datasets and a unique architecture for network feature preprocessing, it uses KDD Cup '99 datasets as input. Class distributions are thoroughly shown and examined to facilitate this process. Decision tree pruning and backward removal are the foundations of the feature reduction technique used. Using one-classifier training with extremely few features, this work ends by giving several basic feature sets for identifying both individual and all attacks. Totally 42 features are exist in the given dataset.

Numerous attributes that represent network connection properties are included in the dataset. These consist of the protocol type (e.g., TCP, UDP, ICMP), the service type (e.g., HTTP, FTP), and the connection length in seconds. It also indicates how many data bytes were transferred and received between the source and the destination, as well as whether the connection produced regular traffic or an error condition. Specific criteria, like whether the connection is to the same host or port and the quantity of urgent packets or hot indicators, are captured by additional features.

In addition to actions on access control files and outgoing commands in an FTP connection, the dataset documents the quantity of unsuccessful login attempts, root shell access, and compromised situations. Additionally recorded are login statuses, including successful and visitor logins. The percentage of connections with SYN, REJ, and RST errors is one example of a traffic behavior metric that summarizes connection counts and error rates over predetermined time intervals. Additionally, the features monitor host-specific activity, including the number of services used or connections to the same host and port, as well as fluctuations in service utilization. Together, these characteristics define network traffic patterns and aid intrusion detection systems in determining if a behavior is malicious or typical.

B. Data preparation and Preprocessing

A popular benchmark dataset to evaluate intrusion detection systems (IDS) is the KDD Cup 1999 dataset, which includes the dataset that is being provided. Records of both normal and malicious network behaviour are included in this dataset. The network connection's properties are represented by a variety of numerical and category attributes in the dataset. The type of service, protocol, amount of packets, error rates, and connection duration are all described by these features. The record's label is shown in the final column where malicious category contain attacks like smurf, neptune etc.. In Preprocessing, the missing values are handled using mean imputation with equation (1),

$$X = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i} \qquad (1)$$

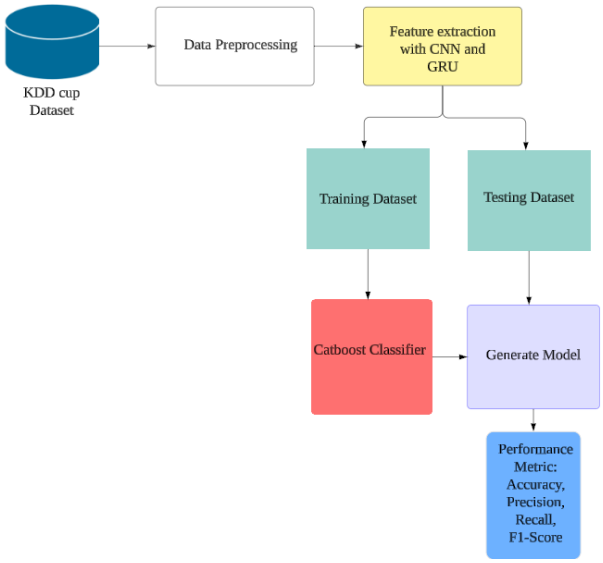

Fig. 4 : Architecture Diagram

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Then One Hot encoding is implemented with categorical features where labels are encoded. Z-Score Normalisation is applied to normalise the data using equation (2),

$$z = \frac{x-\mu}{\sigma} \qquad (2)$$

C. Model Training Workflow

In order to detect attacks, traditional intrusion detection models ignore spatial features and focus more on time series features.

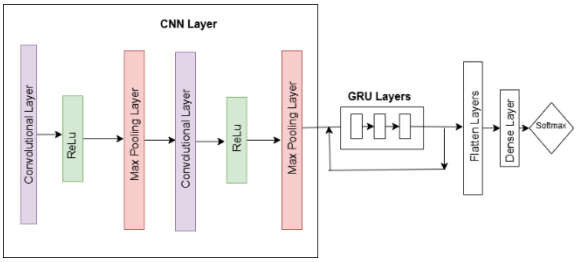

Fig. 5 CNN+GRU Feature Extraction Algorithm

GRU can easily capture the long term dependencies in sequential data using it's gating mechanism. This reset and update gate control the information flow where the long sequence data is utilised to generate the model and vanish the gradient problems. This GRU is simpler than LSTM due to its fewer gates and can handle large number of parameters based on the input size and hidden units.

Pseudocode of Intrusion Detection Workflow using Hybrid CNN-GRU Framework

1. Loading KDDCup 1999 dataset

2. Data ← $ImputerMissingValue(data)$

3. Data ← $Removeduplicate(data)$

4. enc ← $oneHotEncode(data)$

5. data_normalise ← $MinMaxScale(\text{enc})$

6. train,test ← $split(\text{data}_{\text{normalise}}, \text{ratio} = 0.8)$

7. $feature\_extraction\_cnn\_GRU$

$$X \in R^{n*t*f}$$

Where n: batch size,t: time steps, f-features

Convolution Layer (Conv1D)

Apply W_{conv}: $convolution\ kernel, B_{conv}$: $bias$

$$Z_{conv} = ReLU(X * W_{conv} + B_{conv})$$

$$Z_{conv} \in R^{n*t*k}$$

MaxPooling Layer

$$Z_{pool} = Maxpool(Z_{conv} + Pool\ size)$$

$$Z_{pool} \in R^{n*t*k}$$

GRU Layer

$$h_t = GRU(Z_{pool}[t] + h_{t-1})$$

$$H \in R^{n*u}$$

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Flatten Layer

$$Z_{flat} = Flatten(H)$$

$$Z_{flat} \in R^{n*(t^n*k)}$$

Fully Connected Dense Layers

$$Z_{dense1} = \sigma(Z_{flat}.W_{dense1} + B_{dense1})$$

$$Z_{dense1} \in R^{n*d_1}$$

$$Z_{dense2} = Softmax(Z_{dense1}.W_{dense1} + B_{dense1})$$

$$Z_{dense2} \in R^{n*d_2}$$

8. model ← $traincatboost(t_f, t_l)$

9. pred ← model . predict($test_f$)

10. scores ← $crosssVali(model, k = 5)$

11. metrics ← $evaluate(model, tetset)$

In contrast, the CNN architecture failed to extract the long distance information for spatial features. As a result, the CNN is stacked with GRU to overcome this limitation. As a result, the CNN extracts the first spatial features from the input data, whereas the GRU extracts the second and third. Then these features are passed to GRU to capture the temporal dependencies where it can learn both spatial and temporal features.

a) CNN Architecture:

Convolutional neural networks (CNNs) are mainly used to extract features from data. As a result of the CNN module of the intrusion detection model, spatial characteristics of the input dataset are analyzed in order to identify patterns in various dimensions of network traffic data. To extract spatial characteristics, these layers use convolution techniques. In order to create feature maps, the kernel performs matrix multiplication as it moves across the input data. Three consecutive timestamp of patterns are captured by a convolution with a 3 kernel size. By setting negative values to zero, the Rectified Linear Unit creates non-linearity and guarantees that the network can x

b) GRU Architecture:

GRU is the RNN which handles the sequential data of intrusion in effective manner. The CNN-generated feature maps' temporal relationships are captured by GRU in the intrusion detection system. The CNN module's output feature maps, together with batch_size, timesteps, and features, are passed into the GRU layers as input. Every GRU cell maintains a hidden state that records temporal dependencies while processing one timestep at a time. Two gates are used by GRU cells: Update the gate regulates the amount of data that is carried over from the previous hidden state. This reset gate determine the removal of past information from the memory for further proceed. The gates effectively eliminate unnecessary data while ensuring the model retains long-term dependencies.

The sequential nature of the data is captured by updating the concealed state at every timestep. Timesteps with batch size and hidden units are included in the output shape. When time-series data patterns reveal threats in network intrusion scenarios, GRU's ability to identify temporal dependencies is crucial.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Layer Name	Input Shape	Output Shape
Input Layer	(None, timesteps, features)	(None, timesteps, features)
Conv1D_1	(None, timesteps, features)	(None, timesteps - kernel_size + 1, 64)
ReLU_1	(None, timesteps - kernel_size + 1, 64)	(None, timesteps - kernel_size + 1, 64)
MaxPooling_1	(None, timesteps - kernel_size + 1, 64)	(None, reduced_timesteps, 64)
Conv1D_2	(None, reduced_timesteps, 64)	(None, reduced_timesteps - kernel_size + 1, 128)
ReLU_2	(None, reduced_timesteps - kernel_size + 1, 128)	(None, reduced_timesteps - kernel_size + 1, 128)
MaxPooling_2	(None, reduced_timesteps - kernel_size + 1, 128)	(None, reduced_timesteps // pool_size, 128)
GRU_1	(None, reduced_timesteps // pool_size, 128)	(None, hidden_units)
Flatten Layer	(None, hidden_units)	(None, flattened_size)
Dense_1	(None, flattened_size)	(None, 256)
Dense_2 (Output Layer)	(None, 256)	(None, num_classes)

Fig. 6 Layers of CNN+GRU

Because it has fewer parameters and simpler gate mechanisms than conventional RNNs or LSTMs, it is computationally more efficient. GRU avoids the vanishing gradient issue by managing long-term dependencies efficiently. The GRU processes the spatial characteristics that the CNN extracts for temporal analysis. When combined, they offer a thorough depiction of the information. The hybrid approach is appropriate for intrusion detection problems where both spatial and temporal linkages are crucial, as GRU captures sequential dependencies while CNN concentrates on feature extraction across dimensions. The layers of CNN+GRU is depicted in Fig. 6.

- Two layers of Conv1D and MaxPooling are applied to the input. Spatial patterns are extracted by each convolution, and dimensionality is decreased via pooling.
- The GRU receives the reduced feature maps from the last MaxPooling layer. Here, temporal relationships are captured by taking advantage of the data's sequential character.
- For feature vector creation and classification, the GRU's output is compressed into a dense vector and run through two dense layers.

D. CatBoost Classifier

CatBoost is a ML technique that uses DT with gradient boosting. A series of decision trees are constructed one after the other during training. Compared to the trees that came before it, each new tree is constructed with less loss. The initial parameters determine how many trees are used. This algorithm has an ability to handle the heterogenous data effectively and reduce the loss function using learning process. During each iteration the algorithm calculate the negative gradient of loss function for current prediction. These negative gradient measure the direction and magnitude and reduce the loss thereby added to the existing prediction which improve the prediction process. It flows the line search technique that finds the values which reduce the loss function and update the model which is also called optimal scaling factor. The gradient based optimisation technique is utilised for constructing the decision tree in CatBoost, where the negative gradient get fitted to predict the loss functions. These region is iteratively refine and achieve more accurate prediction. Additionally it can handle the categorical features and advance technique such as boosting to prevent overfitting and easily handle the larger dataset.

After extracting feature vector from CNN+GRU architecture is fetched to Catboost classifier, the processing step explained below,

1. Input Features from CNN+GRU

The feature such as shape and dimensions are taken from the input data where feature vector contain 128features.

2. CatBoost Training Process

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Let X be the feature matrix corresponding to target labels Y. The label of classification is normal or malicious. The algorithm build the ensemble decision tree for iteration. In each iteration the prediction for all samples set to mean of target labels. The residual for each prediction with sample is calculated with the equation 1.

$$R_i = y_i - P_i \quad (3)$$

P_i is the current prediction for sample i. The decision tree is develop to predict the residuals R_i. The feature values of X determine the split points and leaf values represent the average residuals of samples. The learning rate μ:

$$P_i^{new} = P_i + \mu . T(X_i) \quad (4)$$

Where $T(X_i)$ of current output from current tree of sample i. The iteration get repetition to refine the output.

The output from the CNN+GRU are numerical vector while the catboost handle both categorical and numerical data for prediction. This catboost uses ordering boosting to avoid overfitting for small datasets and creates a subset for training data which utilised to calculate gradients and construct tree. The dataset splitting is carried on information gain or Gini Impurity. The L2 regularization prevents overfitting and learning rate control the contribution for each trees. Based on the probability score of each class the binary classification is take place.

The catboost algorithm contain inbuilt cross validation and evaluation. Cross-validation is a statistical model that machine learning professionals can use to compare and choose machine learning models for a given application. The method can assist you in resolving issues like overfitting a model, which can lead to problems like less-than-ideal performance in practical situations. By assessing the model's performance across several validation data sets during training, cross-validation helps to prevent this problem as well as others like underfitting. The following steps describe how to apply cross-validation over 10 k-folds using k-fold cross-validation. Ten folds, or subsets, should be used in the data collection. There is roughly the same amount of data in each fold. Five or ten are popular values for k, although you can change it depending on the size and general needs of the data set.

The dataset is divided into k folds where each fold is the subset of data which containing the equal proportion of the samples that balance the distribution of data across folds. In first iteration, the test set is selected and model trained on the K-1 folds. After that remaining folds are used for training, till the k=10 the process is in repetition. These model's performance is evaluated using performance metrics and display once the iteration completed. After the k iteration, the single summary metrics are listed with average performance of single train-test split. In the next section, The developed system will explain how the performance metrics reveal the strengths and weaknesses of the model, as well as how they represent the potential for improvement.

IV. RESULT ANALYSIS AND DISCUSSION

Presented Developed systems detailed evaluation of the CNN+GRU and Catboost Classifiers in this section. The whole application is implemented in the python language with deep learning algorithm such as tensor flow framework. The experiment is carried out on the high-performance Tyrone PC which is robust hardware specification support the intensive computations. With a CPU speed of 2.20 GHz, 2 CPUs, 128 GB of RAM, and a hard drive capacity of 2 TB, this application has an Intel(R) Xeon(R) Silver 4114 CPU. This combination supports the application to handle large heterogenous dataset and generate the complex model in efficient manner. The results are presented in both graphical and tabular format for providing the in-depth views of the proposed model across the evaluation metrics. It is simpler to evaluate the model's behavior and efficacy in tackling the difficulties of IoT network intrusion detection thanks to these quantitative and visual representations, which also improve the results' interpretability.

A. Performance Metrics

Several metrics have been used in evaluating the performance of the proposed model. Most often, evaluation metrics include accuracy, detection rate, false positive rate, precision, and F1 score. Based on these measures, have considered four factors.

- TP: In TP, the prediction of true malicious attack in network which is classified accurately as true label.
- TN: In TN, the prediction of true malicious attack in a network which is classified as normal action.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

- FP: In FP ,system incorrectly classify normal behaviour as malicious
- FN: This indicates that the system misclassified the actual malicious network as normal

Using equation, all the metrics are calculated using the TP, TN,FP,FN which is represented,

Accuracy: It indicates the proportion of observations in the testing set that the model properly divided into instances.

$$Accuracy = \frac{TP+TN}{FN+TP+FP+TN} \quad (5)$$

Detection Rate: It is the ratio of actual attacks to the correctly identified by the model also called as Recall

$$DR = \frac{TP}{FN+TP} \quad (6)$$

Precision (PR) evaluate the proportion of correctly identified malicious attacks among all the observed data which are classified.

$$PR = \frac{TP}{TP+FP} \quad (7)$$

False Alaram Rate: It categories the normal traffic as malicious to all normal traffic.

$$FAR = \frac{FP}{FP+TN} \quad (8)$$

F1-Score: The F1 Score determines the average of PR and DR. As FP and FN are accounted for in calculations, it is more useful than accuracy, especially when class distribution is imbalanced.

$$F1 = 2 * \frac{RC*PR}{RC+PR} \quad (9)$$

ROC Curve: Receiver operating characteristics (ROC) curves are graphs that exhibit the true positive and false positive rates to demonstrate how well classifiers operate. In machine learning, the effectiveness of the algorithm can be gauged by the area under the ROC curve (AUC). When the false positive rate (FPR) is plotted on the x-axis of a graph and the true positive rate (TPR) on the y-axis, the ROC curve is formed. The outcomes of the predicted data against the labeled data are shown in a confusion matrix after a classification model has examined training data. It utilize this information the TPR and the FPR to create the ROC curve, which can assist in assessing the effectiveness of machine learning model.

B. Evaluation of Feature Extraction

a) Correlation Matrix for All Features

Finding connections between each feature in the dataset is made easier by the correlation matrix. While low correlation denotes independence, high correlation features may offer redundant information. Insights on the most important characteristics for prediction or aspects that can be combined may be revealed by strong correlations. Before feeding the data into intricate models like CNN or GRU, it is essential to comprehend these correlations to make sure the model isn't overloaded with duplicated information. To determine the preprocessing procedures, determine which features are substantially associated with the target or with each other. During feature extraction, features that have a high correlation with the target are probably going to contribute more. To increase the diversity of the input data, features that have few connection with one another are maintained which is depicted in Fig.6. The potential for dimensionality reduction using methods like Principal Component Analysis (PCA) may be indicated by highly connected features. This guarantees that the model is not overloaded with redundant data and can concentrate on the most instructive features of the dataset. The diversity of the input data is increased by weakly correlated features, which do not significantly overlap with other features. They may also help identify complex patterns that could otherwise be overlooked.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

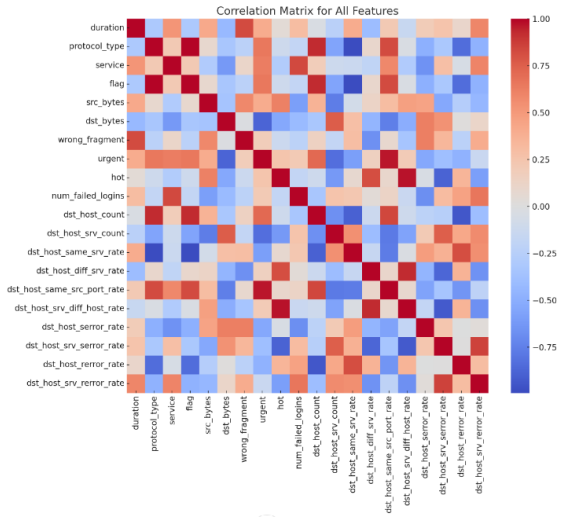

Fig. 7 Correlation matrix for All Features

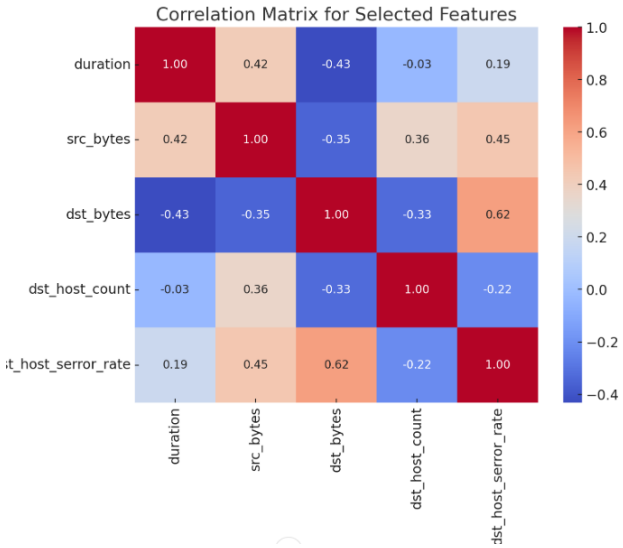

Fig. 8 Correlation matrix for Selected Features

b) Correlation Matrix for Selected Features

Focuses on a subset of features that have been chosen for their relevance to the target variable based on domain expertise, feature significance measures, or past observations. The most informative features while reducing dimensionality. Makes the dataset easier to use for downstream modeling, particularly for computationally demanding processes like feature extraction from CNN and GRU. These particular properties will serve as the main input for GRUs (Gated Recurrent Units) and CNNs (Convolutional Neural Networks). The chosen characteristics' spatial correlations and patterns will be captured by CNNs. If there are temporal relationships in the data, GRUs will take advantage of sequential dependencies. Training speed and interpretability are enhanced by the pre-selected characteristics, which guarantee a condensed and targeted input space which is depicted in Fig. 8.

Focusing on this smaller collection of characteristics yields a number of benefits, such as improved training speed, which is especially helpful when training on high-performance hardware because smaller input sizes guarantee faster computations. Additionally, this optimization makes it possible to experiment with various model designs or hyperparameters without incurring undue computing costs. Improved Interpretability which

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

is easier to comprehend how each variable influences the model's predictions when the feature set is smaller and easier to read and evaluate. In applications like intrusion detection, where explainability can be essential for determining the underlying reasons of abnormalities or attacks, this is very beneficial. Furthermore, this focused feature selection guarantees that the deep learning models only receive the most important attributes, freeing them up to concentrate on identifying important patterns and dependencies. Since they are no longer required to process duplicate or unnecessary data, CNNs and GRUs are in a better position to operate effectively. The predictive system's accuracy and efficiency are increased when feature selection and model design are in harmony. The chosen features shown in Fig. 8 strike a careful balance between keeping important details and simplifying things. This method improves the deep learning pipeline's overall efficacy while also increasing the system's scalability and interpretability. These adjustments guarantee that, even in the face of difficult computational and temporal limitations, the suggested approach can handle real-world datasets with a high degree of correctness and dependability.

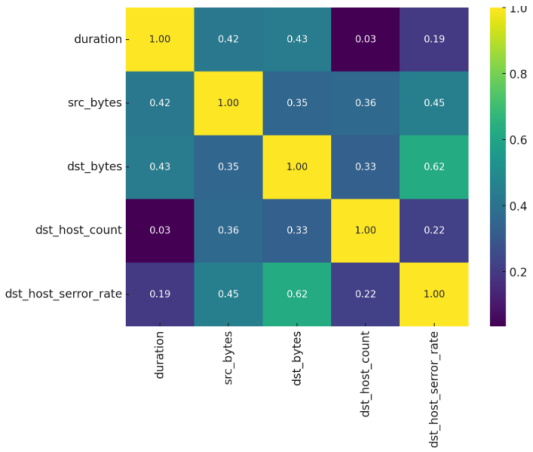

Fig. 9 Cross entropy for selected Feature

c) *Cross Entropy for selected Feature*

A non-linear similarity metric called cross entropy is used to evaluate intricate relationships between variables. Here, feature interdependence is approximated using a simplified variant that uses absolute correlation values. Aids in locating non-linear connections that conventional linear correlation analysis could overlook. Ensures that the features that have the strongest correlations are kept for modeling that comes later. Cross entropy-based feature selection will improve model resilience by supplying information that accounts for non-linear dependencies. Higher-order representations are extracted from these characteristics using CNNs and GRUs, and the final prediction is then provided by CatBoost classifier.

TABLE 1 PERFORMANCE METRICS COMPARED WITH ML ALGORITHMS

Algorithms	**Performance Metric**			
	Accuracy	***Precision***	***Recall***	***F1-score***
NB	99.1	99	99	99
MLP	99.8	99.9	99	99
SVM	99	100	100	100
Proposed	100	100	100	100

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

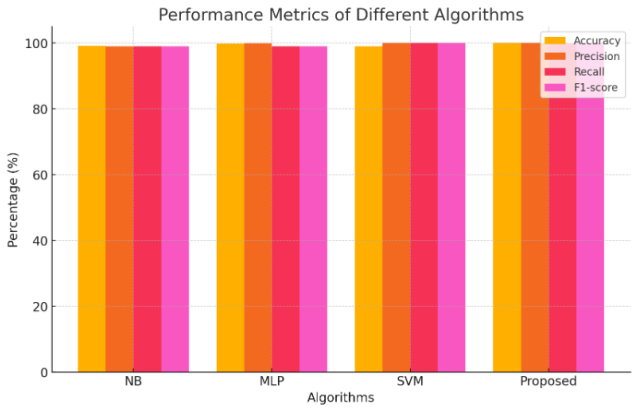

Fig. 10. Performance Metrics of ML algorithm

The existing algorithm [26] performance is compared with the proposed system where it achieves 100% accuracy than remaining ML algorithms using same dataset which is tabulated in table 1.

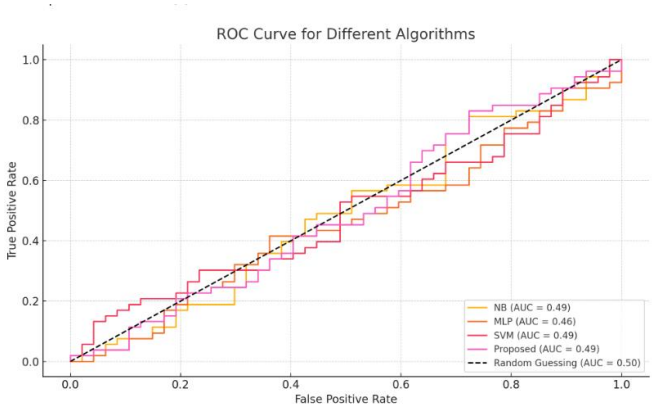

Fig. 11 ROC Curve of ML algorithms

This table compares four machine learning algorithms on four important metrics -- Accuracy, Precision, Recall, and F1-score -- Naive Bayes (NB), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), and a proposed algorithm. Using Naive Bayes, 99.1% of the instances were correctly classified, meaning 99.1% of the instances were correctly classified. With an accuracy of 99.8%, the MLP neural network model fared somewhat better. Although the SVM's performance was slightly lower than the MLP's, it was still rather good at 99%. Notably, the proposed technique performed better than any previous algorithm, properly identifying every case in the dataset with a perfect 100% accuracy rate.

Naive Bayes earned 99% when measuring Precision, which quantifies the percentage of genuine positives among all cases projected as positive. SVM performed exceptionally well with 100% precision, which means it produced no false positive errors, while the MLP performed even better with 99.9% precision. Additionally, the proposed approach equaled the SVM in this statistic with 100% precision.

With a recall score of 99%, which measures the model's capacity to recognize every positive instance, Naive Bayes was able to accurately identify 99% of all real positive cases. At 99%, MLP obtained the same recall level. The SVM classifier identified all positive instances in the dataset without missing any, demonstrating flawless performance with 100% recall. The proposed approach showed that it could detect all positive occurrences by matching SVM with 100% recall.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Lastly, both SVM and the proposed approach achieved 100% on the F1-score, a parameter that strikes a compromise between precision and recall. In terms of the trade-off between precision and recall, the MLP and Naive Bayes both achieved 99%, which is a good but little less balanced result which is tabulated in table 4.1 and drawn bar graph and ROC curve Fig. 10 and Fig. 11.

TABLE 2 PERFORMANCE METRICS COMPARED WITH DL ALGORITHMS

Metric	MOA-WMA [32]	DT-IDS [33]	XSRU [35]	LEAESN [34]	Proposed
Sensitivity (%)	90.1	92.76	94.76	95	99.12
Accuracy (%)	90.42	92.76	95.76	95.1	99.01
FRR (%)	90.44	92.99	94.99	95.4	99.02
Specificity (%)	91.09	93.33	95.33	95.8	99.18
Precision (%)	91.21	92.14	92.14	94.5	99.191
F1-score (%)	91.39	92.4	95.98	94.8	99.355
Recall (%)	91.41	92.53	94.53	95.1	99.479
FMR (%)	91.47	93.68	95.68	96	99.24
FAR (%)	91.67	93.69	96.69	96.4	99.34
FNMR (%)	91.92	93.97	97.97	96.4	99.56

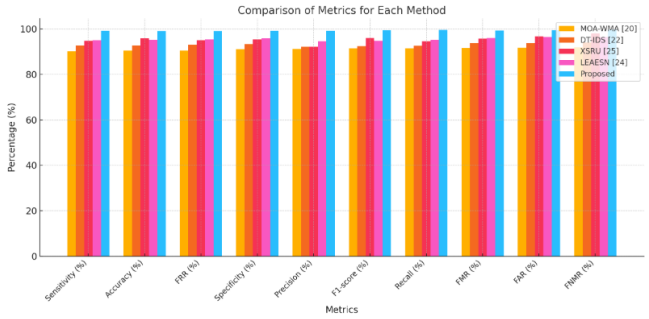

Fig. 12 Performance Metrics of DL Algorithm

Performance metrics for several intrusion detection systems, such as MOA-WMA, DT-IDS, XSRU, LEAESN, and the proposed system, are compared in the table 2. Sensitivity, accuracy, F1-score, recall, specificity, precision, false rejection rate (FRR), false match rate (FMR), false acceptance rate (FAR), and false non-match rate (FNMR) are the metrics that are assessed.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

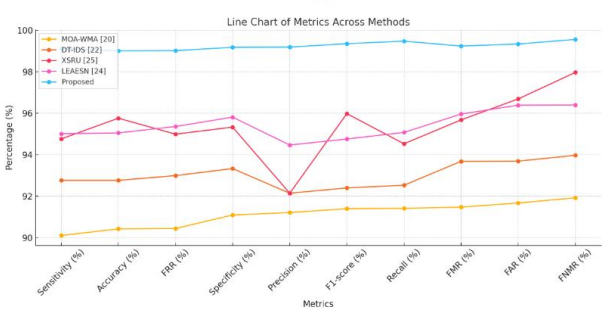

Fig. 13 Line chart of DL Algorithm

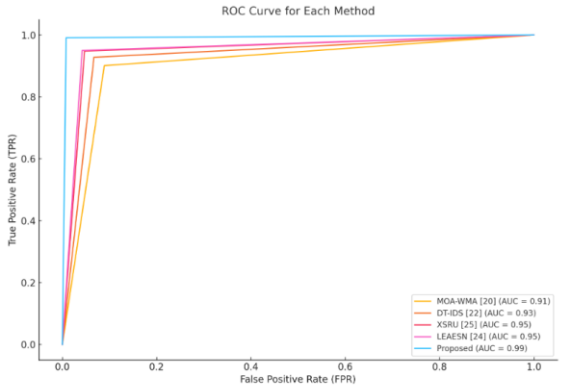

Fig. 14 ROC Curve of DL algorithm

Sensitivity, which measures how well the system detects positive instances (i.e., intrusion attempts), is where the proposed system excels. It achieves an astounding 99.12%, much above XSRU's 94.76% and LEAESN's 95.01%. At 99.01%, the proposed system outperforms competing systems such as XSRU (95.76%) and DT-IDS (92.76%) in terms of accuracy, which measures the system's total correctness in identifying both malicious and normal occurrences.

The proposed system has a lower rate of false rejections than others, as evidenced by its 99.02% False Rejection Rate (FRR), which quantifies the rate at which legal instances are mistakenly categorized as malicious. The Proposed system once again has the highest specificity, which measures its ability to accurately detect non-intrusions (normal behavior), at 99.18%, exceeding XSRU's 95.33%.

In comparison to other approaches, such LEAESN (94.466%), the Proposed system has the fewest false positives, as seen by its maximum precision of 99.191%, which represents the percentage of real positive findings among all expected positives. The proposed system's F1-score, which combines precision and recall, is 99.355%, overcoming the other approaches, such as XSRU (95.98%) and LEAESN (94.756%).

With a recall of 99.479%, the proposed system outperforms XSRU (94.526%) and DT-IDS (92.526%) in terms of the system's potential to accurately identify all positive instances.

The Proposed system has the lowest False Match Rate (FMR), which measures the rate at which malicious instances are mistakenly classified as non-malicious, at 99.24%, and the highest False Acceptance Rate (FAR), which measures the frequency with which valid instances are mistakenly accepted as malicious, at 99.34%. Last but not least, the Proposed method outperforms all other systems in minimizing the False Non-Match Rate (FNMR), which indicates how frequently adverse instances are erroneously rejected, at 99.56% which is tabulated in table 4.2 and drawn bar graph, line and ROC curve Fig.12, Fig.13, Fig.14.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

V. CONCLUSION AND FUTURE SCOPE

In conclusion, our proposed method secure the IoMT network from intrusion attack. Through the integration of a CNN-GRU model for feature extraction and the CatBoost classifier for classification, so have created a reliable system that can precisely detect and react to intrusion attempts. Data cleaning, encoding, and normalization were among the preprocessing procedures that made sure the dataset was ready for model training and allowed the system to function effectively with the available data. While the GRU layer handled temporal dependencies and enabled the model to assess sequences of network action across time, the CNN layer successfully captured spatial patterns in the feature space. In order to keep the IoMT network safe from possible security risks, the CatBoost classifier then correctly identified whether a particular intrusion attempt was malicious or not. Our evaluation and cross-validation processes demonstrated the approach's resilience and dependability in real-world situations by demonstrating that it performed well.

As a result of our work, IoMT networks can be made more secure by detecting and reducing intrusion attacks. The proposed system's scalability enables it to be modified for use in greater healthcare settings, guaranteeing the ongoing security of private medical information. Future research can concentrate on enhancing the model's real-time detection capabilities, further refining it to handle increasingly complicated assault situations, and investigating methods for smoothly integrating the solution into the current healthcare IT infrastructure. In the end, our strategy supports continued efforts to protect IoMT networks, guaranteeing patient data security and privacy in a world growing more interconnected by the day.

REFERENCES

[1] Singh, L.; Kanstrup, M.; Depa, K.; Falk, A.C.; Lindström, V.; Dahl, O.; Göransson, K.E.; Rudman, A.; Holmes, E.A. Digitalizing a brief intervention to reduce intrusive memories of psychological trauma for health care staff working during COVID-19: Exploratory pilot study with nurses. JMIR Form. Res. 2021, 5, e27473.

[2] Rbah, Y.; Mahfoudi, M.; Balboul, Y.; Fattah, M.; Mazer, S.; Elbekkali, M.; Bernoussi, B. Machine learning and deep learning methods for intrusion detection systems in iomt: A survey. In Proceedings of the 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 3–4 March 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–9.

[3] Kilincer, I.F.; Ertam, F.; Sengur, A.; Tan, R.S.; Acharya, U.R. Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization. Biocybern. Biomed. Eng. 2023, 43, 30–41.

[4] Kim, J.; Campbell, A.S.; de Ávila, B.E.F.; Wang, J. Wearable biosensors for healthcare monitoring. Nat. Biotechnol. 2019, 37, 389–406.

[5] Caldwell, Z.B. The case for a security metric framework to rate cyber security effectiveness for Internet of Medical Things (IoMT). In Women Securing the Future with TIPPSS for Connected Healthcare: Trust, Identity, Privacy, Protection, Safety, Security; Springer: Berlin/Heidelberg, Germany, 2022; pp. 63–81

[6] Market.us. Internet of Medical Things Statistics. 2024. Available online: https://media.market.us/internet-of-medical-things-statistics/ (accessed on 1 April 2024).

[7] H. Fotouhi, A. Causevic, K. Lundqvist, and M. Björkman, "Communication and Security in Health Monitoring Systems--A Review," in Proc. COMPSAC, Atlanta, GA, USA, pp. 545-554, 2016.

[8] Y. Xin, K. Lingshuang, L. Zhi, C. Yuling, L. Yanmiao, Z. Hongliang, G. Mingcheng, H. Haixia, and W. Chunhua, "Machine learning and deep learning methods for cybersecurity," IEEE ACCESS, vol. 6, pp. 35365-35381, May 2018.

[9] A. L. Buczak, E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," IEEE COMMUN SURV TUT, vol. 18, no. 2, pp. 1153-1176, October 2015.

[10] Balhareth, G.; Ilyas, M. Optimized Intrusion Detection for IoMT Networks with Tree-Based Machine Learning and Filter-Based Feature Selection. Sensors 2024, 24, 5712. https://doi.org/10.3390/s24175712

[11] RM, S.P.; Maddikunta, P.K.R.; Parimala, M.; Koppu, S.; Gadekallu, T.R.; Chowdhary, C.L.; Alazab, M. An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Comput. Commun. 2020, 160, 139–149.

[12] Kumar, A.K.; Vadivukkarasi, K.; Dayana, R. A Novel Hybrid Deep Learning Model for Botnet Attacks Detection in a Secure IoMT Environment. In Proceedings of the 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), Coimbatore, India, 9–11 February 2023; pp. 44–49. [Google Scholar]
[13] Si-Ahmed, A.; Al-Garadi, M.A.; Boustia, N. Survey of Machine Learning based intrusion detection methods for Internet of Medical Things. Appl. Soft Comput. 2023, 140, 110227. [Google Scholar]
[14] Ravi, V.; Pham, T.D.; Alazab, M. Deep Learning-Based Network Intrusion Detection System for Internet of Medical Things. IEEE Internet Things Mag. 2023, 6, 50–54. [Google Scholar]
[15] Chaganti, R.; Azrour, M.; Vinayakumar, R.; Naga, V.; Dua, A.; Bhushan, B. A Particle Swarm Optimization and Deep Learning Approach for Intrusion Detection System in Internet of Medical Things. Sustainability 2022, 14, 12828.
[16] Cao, B.; Dong, W.; Lv, Z.; Gu, Y.; Singh, S.; Kumar, P. Hybrid Microgrid Many-Objective Sizing Optimization With Fuzzy Decision. IEEE Trans. Fuzzy Syst. 2020, 28, 2702–2710. [Google Scholar] [CrossRef]
[17] Chen, P.; Liu, H.; Xin, R.; Carval, T.; Zhao, J.; Xia, Y.; Zhao, Z. Effectively Detecting Operational Anomalies In Large-Scale IoT Data Infrastructures By Using A GAN-Based Predictive Model. Comput. J. 2022, 65, 2909–2925. [Google Scholar] [CrossRef]
[18] Li, B.; Zhou, X.; Ning, Z.; Guan, X.; Yiu, K.-F.C. Dynamic event-triggered security control for networked control systems with cyber-attacks: A model predictive control approach. Inf. Sci. 2022, 612, 384–398.
[19] Saheed, Y.K.; Abiodun, A.I.; Misra, S.; Holone, M.K.; Colomo-Palacios, R. A machine learning-based intrusion detection for detecting internet of things network attacks. Alex. Eng. J. 2022, 61, 9395–9409. [Google Scholar] [CrossRef]
[20] Omuya, E.O.; Okeyo, G.O.; Kimwele, M.W. Feature selection for classification using principal component analysis and information gain. Expert Syst. Appl. 2021, 174, 114765. [Google Scholar] [CrossRef]
[21] Thamilarasu, G.; Odesile, A.; Hoang, A. An intrusion detection system for internet of medical things. IEEE Access 2020, 8, 181560–181576. [Google Scholar] [CrossRef]
[22] Šabić, E.; Keeley, D.; Henderson, B.; Nannemann, S. Healthcare and anomaly detection: Using machine learning to predict anomalies in heart rate data. AI Soc. 2021, 36, 149–158. [Google Scholar] [CrossRef]
[23] Hady, A.A.; Ghubaish, A.; Salman, T.; Unal, D.; Jain, R. Intrusion detection system for healthcare systems using medical and network data: A comparison study. IEEE Access 2020, 8, 106576–106584. [Google Scholar] [CrossRef]
[24] He, D.; Qiao, Q.; Gao, Y.; Zheng, J.; Chan, S.; Li, J.; Guizani, N. Intrusion detection based on stacked autoencoder for connected healthcare systems. IEEE Netw. 2019, 33, 64–69. [Google Scholar] [CrossRef]
[25] Ahmed, M.; Byreddy, S.; Nutakki, A.; Sikos, L.F.; Haskell-Dowland, P. ECU-IoHT: A dataset for analyzing cyberattacks in Internet of Health Things. Ad Hoc Netw. 2021, 122, 102621.
[26] Binbusayyis, A., Alaskar, H., Vaiyapuri, T. et al. An investigation and comparison of machine learning approaches for intrusion detection in IoMT network. J Supercomput 78, 17403–17422 (2022). https://doi.org/10.1007/s11227-022-04568-3
[27] Y. Rbah et al., "Machine Learning and Deep Learning Methods for Intrusion Detection Systems in IoMT: A survey," 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 2022, pp. 1-9,
[28] Idrissi, I.; Boukabous, M.; Grari, M.; Azizi, M.; Moussaoui, O. An Intrusion Detection System Using Machine Learning for Internet of Medical Things. In Proceedings of the 3rd International Conference on Electronic Engineering and Renewable Energy Systems: ICEERE 2022, Saidia, Morocco, 20–22 May 2022; Springer: Berlin/Heidelberg, Germany, 2023; pp. 641–649.
[29] Chitra, R. A Novel Autoencoder Based Feature Independent Ga Optimised Xgboost Classifier for Iomt Malware Detection. SSRN 2023, 1–29.
[30] Lu, W. Applied Machine Learning for Securing the Internet of Medical Things in Healthcare. In Proceedings of the Advanced Information Networking and Applications: Proceedings of the 37th International Conference on Advanced Information Networking and Applications (AINA-2023), Juiz de Fora, Brazil, 29–31 March 2023; Springer: Berlin/Heidelberg, Germany, 2023; Volume 2, pp. 404–416.

Research Article

[31] KDD Cup 1999 Data
[32] Hameed, Shilan S., Wan Haslina Hassan, and Liza Abdul Latiff. "Anefficient fog-based attack detection using an ensemble of MOA-WMAfor Internet of Medical Things." International Conference of ReliableInformation and Communication Technology. Springer, Cham, 2021.
[33] Gupta, Karan, et al. "A tree classifier based network intrusion detectionmodel for Internet of Medical Things." Computers and Electrical Engi-neering 102 (2022): 108158.
[34] Salemi, Hossein, et al. "LEAESN: Predicting DDoS attack in healthcaresystems based on Lyapunov exponent analysis and echo state neural net-works." Multimedia Tools and Applications 81.29 (2022): 41455-41476
[35] Khan, Izhar Ahmed, et al. "XSRU-IoMT: Explainable simple recurrentunits for threat detection in Internet of Medical Things networks." FutureGeneration Computer Systems 127 (2022): 181-193.