Deep Learning-Driven Intrusion Detection System for Distributed Denial of Service Mitigation

Wala ben Rhouma[1], Haythem Hayouni[1]
[1]Department of Computer Sciences, Higher Institute of Computer Sciences of Kef, University of Jendouba, Tunisia

Article Info

Article history:

Received Nov 2, 2024
Revised Apr 15, 2025
Accepted May 4, 2025

Keywords:

DDoS attack
Intrusion Detection System
Machine Learning
Deep Learning
Convolutional Neural Networks

ABSTRACT

DDoS attacks continue to pose a serious risk to digital infrastructures, as they can render online services inaccessible without altering system files or gaining direct control over the target. Traditional security mechanisms often fall short in identifying these attacks promptly due to their massive scale and the subtlety with which they blend into regular traffic. With the advancement of artificial intelligence, especially in the realm of deep learning, new solutions are emerging to enhance the detection and classification of such threats. In this work, we focus on strengthening Intrusion Detection Systems (IDS) by leveraging deep learning methods to improve accuracy and responsiveness in detecting DDoS attacks. Using the comprehensive CIC-DDoS-2019 dataset, we experimented with several deep learning architectures including Feedforward Neural Networks (MLP), Convolutional Neural Networks (CNN), and Recurrent models incorporating Long Short-Term Memory (LSTM). These models were evaluated for their ability to analyze complex traffic behaviors and identify malicious activity within diverse network environments. his study contributes to the ongoing research on intelligent cybersecurity solutions by proposing deep learning-based IDS frameworks that not only detect threats with higher accuracy but also adapt to dynamic attack patterns. Our findings suggest that such models can serve as a critical component in modern security infrastructures, offering scalable and resilient defense mechanisms against increasingly sophisticated cyberattacks like DDoS. Our empirical results demonstrate that the MLP model yielded the most reliable performance, achieving an outstanding classification precision of 99.62% across various traffic categories. This highlights its effectiveness in isolating harmful flows from legitimate ones, thereby reducing the risk of false alarms and improving detection reliability.

Corresponding Author:

Haythem Hayouni
Department of Computer Sciences, Higher Institute of Computer Science of Kef, University of Jendouba, Tunisia
Email: haythem.hayouni@supcom.tn

1. INTRODUCTION

In the modern era of digital transformation, the reliance on networked systems has grown exponentially. From cloud computing to Internet of Things (IoT) devices, from e-commerce platforms to critical infrastructure management, nearly every sector depends on the continuous availability of online services. However, with this increased connectivity comes a proportional rise in cyber threats. One of the most prevalent and destructive forms of attack in this landscape is the Distributed Denial of Service (DDoS) attack [1]. A DDoS attack attempts to exhaust the resources of a server, service, or network by flooding it with illegitimate requests, often from thousands of compromised machines (botnets) distributed globally. Unlike traditional cyber intrusions, DDoS attacks do not necessarily aim to steal data but to disrupt service availability, which can result in financial loss, reputational damage, and system downtime. The growing complexity, speed, and scale of these attacks have made real-time detection and response more crucial than ever.

Despite the existence of conventional security mechanisms such as firewalls, rate limiters, and rule-based Intrusion Detection Systems (IDS) [2], these tools are increasingly ineffective against modern DDoS attacks. Traditional IDSs typically rely on predefined signatures or rules that must be manually updated, which limits their ability to detect novel or zero-day attacks. Moreover, the high volume of network traffic and the subtle nature of some DDoS behaviors often result in false positives or missed detections, further reducing their reliability in dynamic environments. In response to these limitations, the cybersecurity community has begun to explore machine learning (ML) and, more recently, deep learning (DL) [3,4] as promising alternatives for enhancing IDS capabilities. Unlike traditional approaches adapt to new types of attacks without manual intervention, and scale effectively with the volume and complexity of modern network traffic. These properties make deep learning especially suited for building intelligent, adaptive, and accurate intrusion detection systems.

This study concentrates on creating deep learning-based intrusion detection system (IDS) models that can accurately and efficiently identify DDoS attacks. The CIC-DDoS-2019 dataset, which features a broad spectrum of realistic DDoS traffic scenarios, serves as the foundation for evaluating various cutting-edge neural network models. These include the Multilayer Perceptron (MLP), known for its effectiveness in handling structured data; the Convolutional Neural Network (CNN), capable of recognizing spatial data patterns; and the Recurrent Neural Network with Long Short-Term Memory (RNN-LSTM), which excels at modeling sequential and time-dependent traffic behaviors.

This research contributes to the field of intelligent network security in the following ways:

- It presents the design and training of several deep learning-based intrusion detection models namely, MLP, CNN, and RNN-LSTM tailored for identifying different categories of DDoS attacks, utilizing the comprehensive CIC-DDoS-2019 dataset.
- It identifies the Multilayer Perceptron (MLP) as the top-performing model, attaining a high accuracy rate of 99.62
- It establishes the advantages of deep learning techniques over traditional IDS approaches, especially in terms of improved detection rates, greater adaptability, and enhanced robustness against sophisticated attack strategies.
- It highlights the models' ability to automatically learn and prioritize relevant features from raw data, reducing reliance on manual feature engineering.
- It introduces a practical and scalable deep learning-based IDS framework capable of supporting real-time network protection through early detection and accurate classification of DDoS threats.

The remainder of this paper is structured as follows: Section 2 reviews existing literature relevant to this study. Section 3 outlines the proposed research methodology. Section 4 presents and analyzes the experimental results. Lastly, Section 5 concludes the paper and suggests directions for future research.

2. RELATED WORKS

The application of deep learning in detecting DDoS attacks has led to notable progress in the field. These advanced techniques have shown considerable promise in recognizing intricate attack behaviors and boosting the performance of intrusion detection systems (IDS). Several recent studies illustrate the variety of strategies and developments that have emerged in this area.

Liu and Patras [5] proposed NetSentry, an advanced intrusion detection system that utilizes a Bidirectional Asymmetric LSTM (Bi-ALSTM) network to detect early signs of large-scale network attacks. This model was specifically designed to recognize subtle attack patterns over time, making it particularly effective for detecting incipient DDoS attacks. The study found that by considering temporal dependencies in attack data, NetSentry outperformed traditional signature-based systems, achieving an improvement of over 33% in F1 scores compared to other IDS methods. This highlights the significance of real-time detection for large-scale attacks in modern network infrastructures.

Alfatemi et al. [6] addressed the problem of class imbalance in DDoS detection datasets by integrating Deep Residual Neural Networks (ResNets) with the Synthetic Minority Over-sampling Technique (SMOTE). This hybrid method was applied to the CICIDS dataset and yielded an impressive accuracy of 99.98%. By incorporating data augmentation, their approach mitigated training bias and significantly enhanced the model's ability to detect infrequent attack types effectively overcoming a major challenge in DDoS detection systems.

Doriguzzi-Corin and Siracusa [7] introduced FLAD, an adaptive federated learning framework designed for distributed IDS systems. The key feature of this model is that it enables the collaborative training of IDS models across multiple devices or servers without the need to share sensitive data. FLAD dynamically adjusts the computational resources allocated for training based on the complexity of the attack profiles detected. This model significantly reduced the convergence time of federated learning while maintaining high detection accuracy.

Silivery et al. [8] proposed an advanced multi-phase deep learning model to classify both Denial of Service (DoS) and DDoS attacks across multiple classes. Their approach integrated Deep Convolutional Generative Adversarial Networks (DCGAN) for data balancing, ResNet-50 for feature extraction, and an optimized AlexNet classifier. Tested on the CCIDS2019 and UNSW-NB15 datasets, the model achieved over 99.3% accuracy and outperformed previous methods in terms of multi-class attack detection. This study emphasizes the importance of hybrid deep learning models and data balancing in improving detection accuracy for varied and sophisticated DDoS attack types.

Shaikh et al. [9] proposed a hybrid deep learning framework that integrates Convolutional Neural Networks (CNN), Principal Component Analysis (PCA), and Vision Transformers (ViT) to improve DDoS attack detection. In this approach, CNNs were employed to extract key features, PCA was used to reduce data dimensionality, and ViTs provided attention-based learning capabilities. Tested on the CICDDoS2019 dataset, the model reached a high accuracy of 99.99%. The study underscores the effectiveness of combining multiple deep learning methods to capture both spatial and temporal characteristics of network traffic in the context of DDoS mitigation.

Alanazi et al. [10] introduced an advanced intrusion detection system that combines Deep Neural Networks (DNN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks to effectively detect DDoS attacks. Their approach integrated data preprocessing with a deep learning-based architecture, resulting in significant gains in detection speed and accuracy when evaluated on the CIC-DDoS2019 dataset. The hybrid model, capable of learning both spatial and sequential patterns in network traffic, proved effective in enhancing DDoS detection performance in real-time operational settings.

Kumari and Mrunalini [11] proposed a model which focused on capturing sequential dependencies in network traffic, which is crucial for detecting evolving DDoS attacks that may not show immediate signs of malicious behavior. Trained on the CICDDoS2019 dataset, this LSTM model achieved an accuracy of 98% and showed promise in real-time intrusion detection by automatically recognizing traffic patterns typical of DDoS attacks, even in the presence of background noise.

Dhanya et al. [12] explored the application of Long Short-Term Memory (LSTM) networks for detecting DDoS attacks by leveraging the temporal characteristics of network traffic. Trained on the CICDDoS2019 dataset, their model achieved an accuracy of up to 98%, underscoring the strength of LSTMs in identifying threats that develop over time. Compared to traditional signature-based intrusion detection systems, which often fail to recognize novel or evolving attacks, the LSTM-based method showed superior adaptability. This work emphasizes the value of sequential deep learning models for enhancing real-time threat detection capabilities. Alshra'a et al. [13] developed a hybrid intrusion detection model that integrates Deep Convolutional Neural Networks (DCNN) for extracting relevant features with Bidirectional Long Short-Term Memory (BiLSTM) networks to capture sequential relationships in network data. The model was evaluated using datasets such as CICIDS2018 and Edge_IIoT, achieving detection accuracies of up to 100%. This approach highlights the effectiveness of combining CNNs for spatial feature analysis with BiLSTMs for understanding temporal dynamics, resulting in enhanced performance in identifying both DDoS and DoS attacks with high precision and responsiveness. Salmi and Oughdir [14] conducted a comparative analysis of various deep learning models for identifying Denial of Service (DoS) attacks within wireless sensor networks (WSNs). Their findings showed that deep learning approaches particularly Convolutional Neural Networks (CNNs) outperformed traditional detection techniques in terms of accuracy and adaptability to emerging attack patterns. This research is especially noteworthy for demonstrating the viability of deep learning in resource-constrained environments like IoT and WSNs, where swift and accurate threat detection is vital for maintaining network security.

Oyucu et al. [15] introduced an ensemble-based deep learning framework designed to detect DDoS attacks within Software-Defined Networking (SDN) environments, particularly in SCADA systems. By combining the strengths of multiple deep learning classifiers, the proposed system improved both the accuracy and resilience of intrusion detection. The study demonstrated that this ensemble approach not only increased detection rates but also provided adaptive capabilities to counter evolving DDoS threats, positioning it as a strong candidate for enhancing cybersecurity in industrial and critical infrastructure networks. Gankotiyace et al. [16] investigated the application of Deep Convolutional Neural Networks (DCNNs) for identifying DDoS attacks within Wireless Mesh Networks (WMNs). Their work introduced a novel cross-layer detection methodology that examines traffic across various layers of the network protocol stack, enabling the system to uncover attack patterns that might be missed when analyzing layers individually. Utilizing the automatic feature extraction capabilities of DCNNs, the proposed model effectively identified both subtle and complex anomalies in network behavior. The study demonstrated that this cross-layer deep learning strategy significantly enhanced detection accuracy and system resilience, offering a practical solution for real-time intrusion monitoring in dynamic and resource-limited WMN environments. It underscores the growing role of deep learning in fortifying the security of next-generation wireless networks.

Table 1 presents a detailed overview of various studies related to DDoS detection using deep learning models

It outlines the models used, datasets applied, key findings, reported accuracy, and the main contributions of each work. These studies leverage a variety of advanced techniques including LSTM, CNN, ResNet, Vision Transformers (ViT), and hybrid models to detect and classify DDoS attacks with high accuracy, often exceeding 98%. Some works, such as those by Shaikh et al. and Alshra'a et al., report near-perfect accuracy using deep and hybrid architectures, while others like Doriguzzi-Corin and Siracusa explore privacy-aware models like federated learning. Despite the promising results, several limitations can be observed across these solutions. For example, Liu and Patras [5] propose an LSTM-based method focused on early detection, but do not report standard accuracy metrics, which makes it difficult to directly compare with other works. Similarly, Alanazi et al. [10] introduce a hybrid model combining DNN, CNN, and LSTM, yet omit detailed evaluation results such as precision or recall, limiting the assessment of their model's robustness. Doriguzzi-Corin and Siracusa [7] present a federated learning approach, which is innovative in terms of privacy, but their model's effectiveness is not quantified through clear accuracy metrics. Moreover, while Alfatemi et al. [6] achieve extremely high accuracy using ResNet and SMOTE, their reliance on heavy data preprocessing and class balancing raises concerns about scalability and real-time applicability. Models such as those from Kumari and Mrunalini [11], while effective in detecting sequential patterns, are limited to specific datasets like CICDDoS2019, which may not capture the evolving nature of attacks in real-world settings. Additionally, many models depend heavily on public datasets (e.g., CICIDS, CICDDoS2019), which may not fully reflect real-time, heterogeneous network traffic. This introduces a risk of overfitting and limited generalizability. Computational complexity is another common issue models combining CNN, LSTM, and ViT (e.g., Shaikh et al. [9]) offer high performance but may not be deployable on low-resource or latency-sensitive environments like IoT or edge devices. Finally, most models lack explainability, meaning that although they perform well in classification tasks, they provide little insight into how decisions are made, which is a critical issue in cybersecurity contexts where interpretability is essential for trust and validation.

3. PROPOSED RESEARCH METHODOLOGY AND CLASSIFICATION MODEL

To tackle the growing complexity of cyber threats, our research focuses on designing robust intrusion detection systems (IDS) powered by deep learning. Central to our approach is the utilization of the CIC-DDoS-2019 dataset, a highly detailed benchmark provided by the Canadian Institute for Cybersecurity, which reflects realistic DDoS attack scenarios and various patterns of malicious traffic. Our experimental workflow is structured around three separately preprocessed subsets of this dataset, each tailored to evaluate different aspects of DDoS detection and traffic classification. For the first two subsets, we implemented and tested four deep learning architectures: LSTM, DNN, CNN, and MLP. Each architecture was integrated with an autoencoder layer to facilitate dimensionality reduction and extract significant features from the input data, enhancing the classification accuracy. The third dataset (Data 3), focused on multi-class classification of 13 distinct traffic types, required more refined models. To this end, we applied a specialized LSTM model for learning temporal dependencies in traffic sequences, a hybrid DNN-CNN architecture to leverage both spatial

Table 1. Comparison of Related Works on DDoS Detection

Paper	Model/ Methodology	Dataset	Key Findings	Accuracy	Contribution
Liu and Patras [5]	Bidirectional Asymmetric LSTM (Bi-ALSTM)	Large-scale network traffic	Improved F1 score by 33% for early DDoS detection	N/A	Introduced a time-sensitive, LSTM-based model for early detection of large-scale attacks.
Alfatemi et al. [6]	Deep Residual Neural Networks (ResNet) & SMOTE	CICIDS dataset	Achieved 99.98% accuracy using ResNet and data augmentation	99.98%	Combined deep learning with SMOTE for enhanced DDoS detection accuracy.
Doriguzzi-Corin and Siracusa [7]	Federated Learning	Multiple distributed devices	Reduced convergence time and high accuracy	N/A	Introduced federated learning for privacy-preserving DDoS detection.
Silivery et al. [8]	DCGAN, ResNet-50, AlexNet	CCIDS2019, UNSW-NB15	Achieved 99.3% accuracy for multi-class classification	99.3%	Combined multiple deep learning techniques for effective multi-class DDoS attack detection.
Shaikh et al. [9]	CNN, PCA, Vision Transformers (ViT)	CICDDoS2019	Achieved 99.99% accuracy	99.99%	Introduced a hybrid model for improved DDoS detection combining CNN, PCA, and ViT.
Alanazi et al. [10]	DNN, CNN, LSTM	CIC-DDoS2019	Demonstrated improvement in real-time DDoS detection with hybrid models	N/A	Focused on the real-time DDoS detection and feature extraction.
Kumari and Mrunalini [11]	LSTM	CICDDoS2019	Achieved up to 98% accuracy in detecting sequential attack patterns	98%	Used LSTM for detecting evolving DDoS attacks over time.
Dhanya et al. [12]	LSTM	CICDDoS2019	Achieved 98% accuracy	98%	Focused on LSTM for sequential DDoS attack detection.
Alshra'a et al. [13]	DCNN & BiLSTM	CICIDS2018, Edge_IIoT	Achieved up to 100% accuracy in detection	100%	Developed a hybrid model combining CNN and BiLSTM for real-time detection.
Salmi and Oughdir [14]	CNN	Wireless sensor network data	Deep learning outperforms traditional methods in detection accuracy	N/A	Applied deep learning to improve DoS detection in IoT networks.
Oyucu et al. [15]	Ensemble Learning	SCADA systems (SDN)	Enhanced robustness and detection accuracy with ensemble models	N/A	Introduced an ensemble approach for DDoS detection in SDN environments.
Gankotiyacet al. [16]	DP-K-means clustering & ERL-AlexNet	SDN traffic data	Achieved efficient detection and mitigation of DDoS in real-time	97%	Hybrid clustering and deep learning for DDoS detection in SDN.

and sequential characteristics of the data, and a standalone MLP network tailored for handling multi-class outputs with high computational efficiency.

The flowchart for the proposed system is shown in Figure 1.

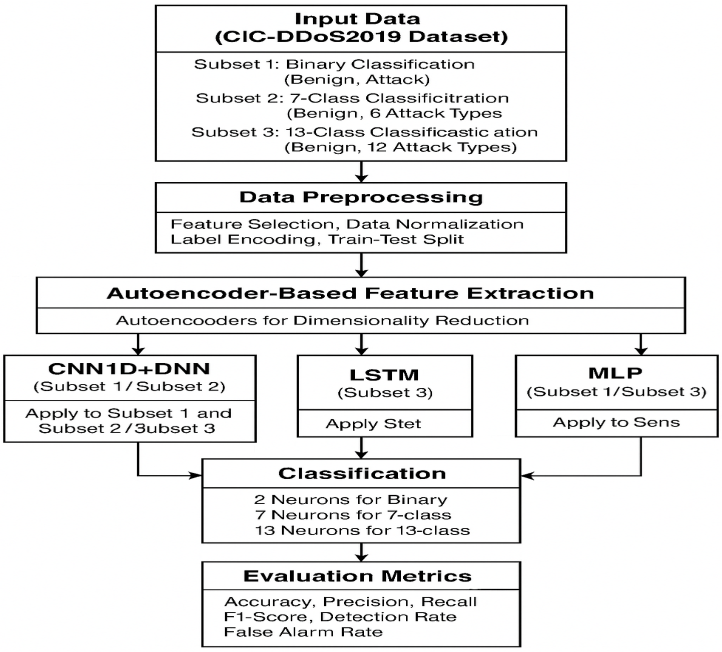


Figure 1. Flow chart of the proposed methodology.

Figure 2 presents the architecture of the proposed deep learning-based intrusion detection system, which integrates three main neural network components: Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Multi-Layer Perceptron (MLP). The process begins with the preprocessing of raw network traffic data, which is then input into the model pipeline. The CNN module first extracts local spatial features, followed by the LSTM layer, which captures sequential dependencies and temporal characteristics in the data. These processed features are subsequently passed to the MLP, which performs the final classification. This hybrid architecture supports both binary and multi-class classification, enabling it to distinguish between normal traffic and various DDoS attack categories. By combining spatial, temporal, and high-level feature analysis, the system improves its ability to detect complex attack patterns with high accuracy and resilience.

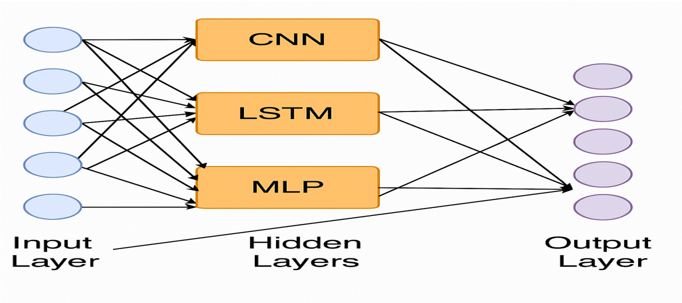


Figure 2. Flow chart of the proposed methodology.

3.1. Strategic Design and Deployment of Deep Learning Models for DDoS Traffic Classification

To enhance our classification performance, we partitioned the CIC-DDoS2019 dataset into three tailored subsets, each corresponding to a specific classification challenge binary, medium-scale multi-class, and fine-grained multi-class detection. We paired each subset with a set of carefully selected deep learning models, enhanced by autoencoders. These unsupervised neural layers perform automatic feature learning and dimensionality reduction.

- *Subset 1 – Binary Classification (Benign vs. Attack)*: This subset targets the fundamental task of determining whether network traffic is normal (BENIGN) or malicious (Attack). We merged training and test datasets into a unified format, simplifying data preprocessing and streamlining the classification pipeline. The objective is to provide a baseline evaluation of model performance in distinguishing clean traffic from any DDoS variant.

Models Applied: CNN-DNN, RNN-LSTM, and MLP, each equipped with an autoencoder layer for preprocessing.

Purpose: Establish the core ability of deep learning models to perform basic intrusion detection low false alarm rates.

- *Subset 2 – Seven-Class Classification (1 Benign + 6 Attack Types)*: This intermediate subset was designed to evaluate the capacity of our models to differentiate between multiple DDoS attack types, while still identifying benign traffic.

  Categories: BENIGN, and six attack types such as UDP, SYN, TFTP, MSSQL, LDAP, and PortMap.

  Approach: The use of autoencoders in this scenario is especially vital, as the inter-class similarities between different attack types can reduce the separability of features.

  Goal: Assess model performance in handling a moderate degree of class complexity where attack types share common temporal or protocol-level features.

- *Subset 3 – Thirteen-Class Classification (1 Benign + 12 DDoS Variants)*: This subset represents the most challenging classification scenario, requiring the model to identify traffic associated with 12 distinct types of DDoS attacks in addition to benign flows.

  Classification Categories Include: BENIGN, SNMP, NetBIOS, SSDP, UDP, SYN, TFTP, MSSQL, LDAP, NTP, DNS, and PortMap.

  Model Enhancement: Before training the deep learning models, we utilized the ExtraTreesClassifier a robust ensemble-based technique for feature selection. This method helped identify the most informative and class-relevant features, allowing us to filter out less significant attributes. By retaining only the most predictive inputs, the model training process became more efficient, leading to improved accuracy and reduced noise in the data.

  Deep Learning Architectures Used: The proposed system incorporates multiple deep learning models to address the complexity of network traffic analysis. A MLP is employed for its effectiveness in handling multiclass classification tasks. LSTM networks are utilized to model long-range temporal dependencies within traffic sequences. Additionally, a hybrid model combining DNN with CNN is implemented to capture both spatial and sequential patterns, enhancing the system's overall detection capability.

  Purpose: Demonstrate the capability of advanced architectures in fine-grained attack classification, a critical step for proactive network defense and threat attribution.

This layered classification strategy progressing from binary to multi-class with increasing granularity, offers a comprehensive evaluation of deep learning models under varied levels of complexity. It also mirrors real-world use cases where intrusion detection systems must not only flag anomalies but also categorize them precisely to trigger the appropriate mitigation strategy. The integration of autoencoders and feature selection techniques ensures high model efficiency and generalization performance, making the proposed framework adaptable for deployment in practical cybersecurity environments.

Table 2 provides a structured overview of the deep learning-based classification strategy applied to the CIC-DDoS2019 dataset, highlighting the division of the dataset into three subsets, each tailored for a specific classification goal. These subsets reflect increasing levels of complexity in network traffic classification and demonstrate the adaptability of deep learning models across different scenarios. Subset 1 focuses on a binary classification task, distinguishing between benign traffic and any type of DDoS attack grouped under a single ”Attack” label. This scenario is used to establish a baseline performance using models like CNN-DNN, RNN-LSTM, and MLP, all integrated with autoencoders to enhance feature extraction and reduce data dimensionality. The dataset is merged into a unified file to streamline the classification process. Subset 2 addresses a 7-class classification problem, where traffic is categorized into one benign type and six distinct DDoS attack types, including UDP, SYN, TFTP, MSSQL, LDAP, and PortMap. The inclusion of autoencoders in this setting is crucial for improving class discrimination, as some attack types share overlapping traffic features. This allows the deep learning models to learn subtle distinctions between attack types with greater precision. Subset 3 represents the most granular classification challenge, consisting of 13 classes 12 individual DDoS attack types and one benign category. To tackle this complex task, a combination of MLP, LSTM, and a hybrid DNN-CNN architecture was employed. Furthermore, ExtraTreesClassifier was used as a feature selection mechanism to isolate the most relevant attributes for each attack type, thereby improving model efficiency and performance.

Table 2. Overview of Deep Learning-Based Classification Strategies used in our proposed model

Subset	Classification Type	Class Categories	Deep Learning Models Used	Special Techniques Applied
Subset 1	Binary Classification	BENIGN vs. Attack (all types merged)	CNN-DNN, RNN-LSTM, MLP (with Autoencoder)	Dataset merged into one CSV; Autoencoder used for feature reduction and noise filtering
Subset 2	7-Class Multi-class Classification	BENIGN, UDP, SYN, TFTP, MSSQL, LDAP, PortMap	CNN-DNN, RNN-LSTM, MLP (with Autoencoder)	Autoencoder improves inter-class feature separability and enhances classifier accuracy
Subset 3	13-Class Fine-Grained Classification	BENIGN + 12 DDoS variants (e.g., NTP, DNS, SSDP, NetBIOS, etc.)	MLP, LSTM, DNN-CNN Hybrid	ExtraTreesClassifier used for feature selection; Multi-class adaptation of deep learning models

3.2. Proposed Hybrid CNN-LSTM Model Architecture

In this study, we introduce a hybrid deep learning model that integrates Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to detect Distributed Denial of Service (DDoS) attacks, utilizing the CIC-DDoS2019 dataset for evaluation. The core objective of this architecture is to combine the strengths of spatial feature extraction and temporal pattern recognition to accurately distinguish between benign and malicious network traffic. As shown in Figure 3, the model incorporates CNN, LSTM, and MLP components, forming a powerful framework for detecting DDoS threats. CNNs are particularly suited for identifying spatial characteristics in the data, such as packet sizes and traffic flow metrics, while LSTMs are adept at capturing sequential dependencies, which are essential for identifying attack behaviors that develop over time. This synergy enables the system to detect even subtle or evolving attack patterns with high precision.

Table 3 presents the detailed breakdown of the architecture of the proposed CNN-LSTM-MLP hybrid model.

Table 3. Proposed CNN-LSTM Hybrid Model Architecture

Layer Type	Details	Purpose
Input Layer	Feature vector from CIC-DDoS2019	Accepts preprocessed network traffic data
Convolutional Layer 1	64 filters, kernel size 3, ReLU activation	Extracts local spatial patterns
Convolutional Layer 2	64 filters, kernel size 3, ReLU activation	Refines feature maps
Max Pooling	Pool size 2x2	Reduces dimensionality and computation
Dropout	Rate = 0.3	Prevents overfitting
LSTM Layer	128 units	Captures long-term temporal dependencies
Dense Layer 1	128 neurons, ReLU activation	Learns high-level abstract features
Dense Layer 2	64 neurons, ReLU activation	Further refines learned features
Output Layer	Softmax (multi-class) or Sigmoid (binary)	Provides final prediction output

3.2.1. Input Layer

Purpose: The input layer receives a feature vector derived from the preprocessed CIC-DDoS2019 dataset, which contains a range of network traffic attributes such as flow duration, packet size, and packet count. These features play a key role in identifying whether the traffic is benign or associated with a DDoS attack. Each traffic instance is encoded as a numerical vector, enabling the model to distinguish between different traffic types. This input is then forwarded through the network's layers for further feature extraction and classification.

3.2.2. Convolutional Layers

Purpose: The convolutional layers are designed to extract local spatial patterns from the network traffic data. These layers are particularly effective at identifying structural relationships such as packet size distributions, interactions between source and destination addresses, and flow-level statistics which are crucial for detecting both legitimate and malicious activity.

- Convolutional Layer 1: This initial layer includes 64 filters with a kernel size of 3, targeting low-level spatial features within the input data. It focuses on detecting basic patterns in packet attributes and flow behavior. The ReLU activation function is applied to introduce non-linearity, enhancing the model's capacity to learn complex data representations.

- Convolutional Layer 2: The second convolutional layer, also composed of 64 filters, builds on the output of the first layer by identifying more sophisticated patterns and deeper inter-feature relationships. This refinement step strengthens the model's sensitivity to subtle deviations in traffic, which is vital for accurately detecting anomalies and potential DDoS activities.

3.2.3. Max Pooling

Purpose: Following the convolutional layers, a max-pooling layer is employed to reduce the spatial size of the feature maps. This operation helps in minimizing computational complexity and mitigating overfitting by retaining only the most dominant features. A pooling filter is applied to downsample the input, selecting the highest value within each region. This ensures that critical spatial information is preserved while less relevant data is discarded.

3.2.4. Dropout Layer

To further prevent overfitting, a dropout mechanism is integrated into the model. During training, 30% of the neurons are randomly deactivated in each iteration. This randomness forces the model to generalize better by learning robust patterns through multiple independent pathways rather than depending heavily on specific neurons. As a result, the model becomes more effective in handling unseen data.

3.2.5. LSTM Layer

The LSTM component plays a vital role in capturing temporal dynamics in network traffic. Since DDoS attacks often exhibit time-based patterns such as gradual increases in request rates or sustained bursts of activity the LSTM layer enables the model to detect these evolving behaviors. Comprising 128 units, this layer is capable of learning long-range dependencies and retaining important contextual information from earlier time steps, allowing it to track how traffic features develop over time and identify suspicious trends.

3.2.6. Dense Layers

Purpose: The fully connected (dense) layers serve to synthesize the features extracted by the preceding CNN and LSTM layers, transforming them into higher-level representations for final decision-making.

- Dense Layer 1: Composed of 128 neurons and using the ReLU activation function, this layer processes the fused spatial-temporal features and facilitates the learning of non-linear patterns without the vanishing gradient problem.

- Dense Layer 2: With 64 neurons, this layer further refines the learned features, enhancing the model's ability to generate more precise predictions by integrating abstract knowledge from earlier layers.

3.2.7. Output Layer

The output layer generates the final classification result. For binary classification tasks, such as distinguishing between normal and attack traffic, a sigmoid activation function is utilized, outputting a probability score between 0 and 1. In multi-class scenarios where the model differentiates among multiple attack categories a softmax activation function is applied, producing a probability distribution over all possible classes and selecting the most likely one as the prediction.

3.3. Hybrid CNN1D-DNN Framework for Enhanced Intrusion Detection

Figure 3 illustrates the architecture of the proposed CNN1D-DNN model, which is a central element of our deep learning-based intrusion detection framework. This model is specifically optimized for the 13-class classification task posed by the CIC-DDoS2019 dataset. By combining one-dimensional convolutional layers with deep dense networks, the architecture effectively captures both local patterns and high-level abstractions in the traffic data, significantly improving classification performance.

In our overall methodology, we designed and evaluated different deep learning architectures for three distinct dataset subsets:

- Binary classification (Subset 1): BENIGN vs. ATTACK

- 7-class classification (Subset 2): BENIGN + six DDoS variants

- 13-class classification (Subset 3): BENIGN + twelve DDoS attack types

Among these, Subset 3 presents the highest complexity, requiring a model capable of identifying subtle differences among multiple attack types. The CNN1D-DNN hybrid model shown in the figure was specifically designed for this purpose. It combines Convolutional Neural Networks (CNN) and Deep Neural Networks (DNN) to capture both spatial dependencies (via CNN layers) and high-level abstract features (via DNN layers).

- The Conv1D and MaxPooling1D layers at the input stage are responsible for capturing local patterns and reducing the temporal dimension.

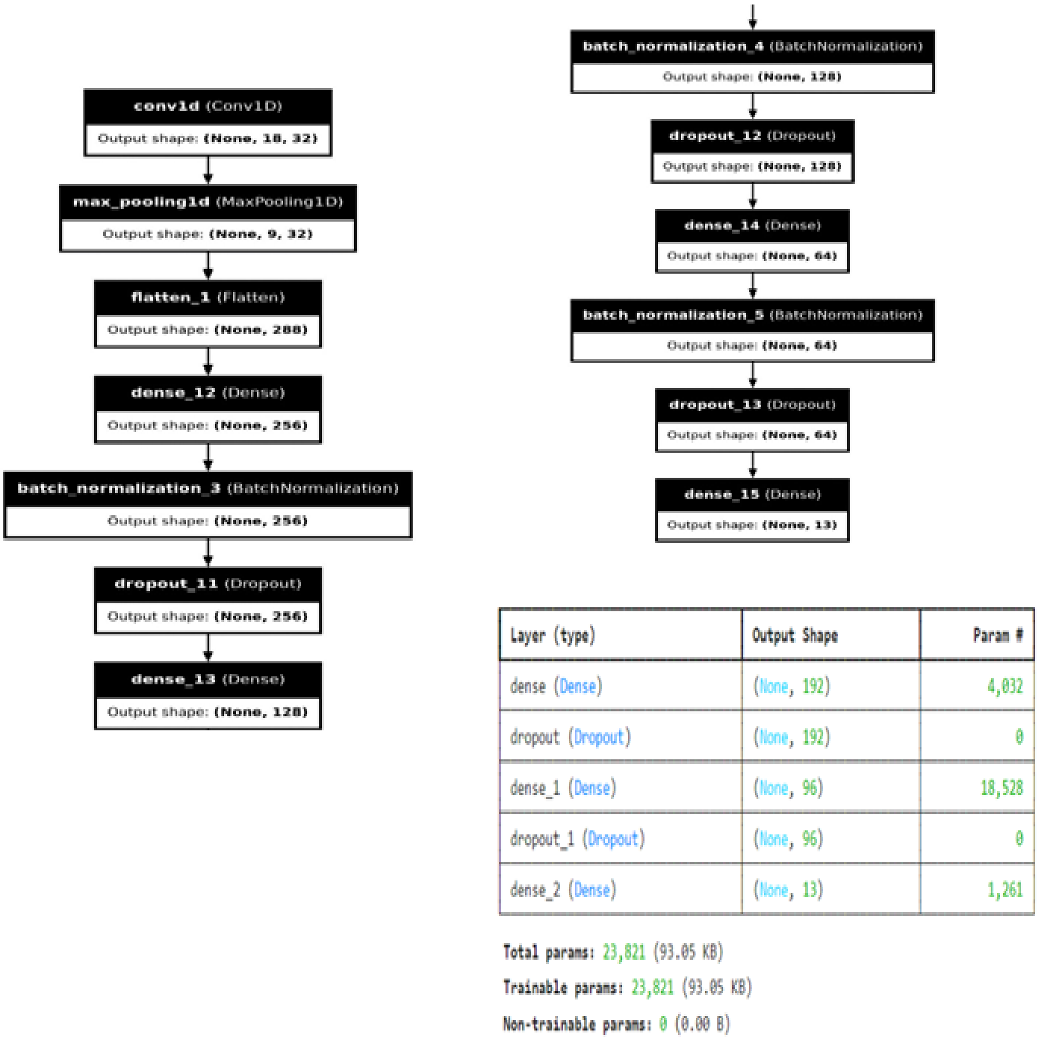
- The Flatten operation transforms the output into a vector suitable for dense processing.

- Dense layers with ReLU activations learn deeper non-linear combinations of features.

- Batch Normalization is integrated after dense layers to stabilize learning and accelerate convergence.

- Dropout layers are used at multiple stages to reduce overfitting by randomly dropping neurons during training.

- The final Dense layer with Softmax activation produces the probability distribution over the 13 classes.


conv1d (Conv1D)
Output shape: (None, 18, 32)
max_pooling1d (MaxPooling1D)
Output shape: (None, 9, 32)
flatten_1 (Flatten)
Output shape: (None, 288)
dense_12 (Dense)
Output shape: (None, 256)
batch_normalization_3 (BatchNormalization)
Output shape: (None, 256)
dropout_11 (Dropout)
Output shape: (None, 256)
dense_13 (Dense)
Output shape: (None, 128)
batch_normalization_4 (BatchNormalization)
Output shape: (None, 128)
dropout_12 (Dropout)
Output shape: (None, 128)
dense_14 (Dense)
Output shape: (None, 64)
batch_normalization_5 (BatchNormalization)
Output shape: (None, 64)
dropout_13 (Dropout)
Output shape: (None, 64)
dense_15 (Dense)
Output shape: (None, 13)

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 192)	4,032
dropout (Dropout)	(None, 192)	0
dense_1 (Dense)	(None, 96)	18,528
dropout_1 (Dropout)	(None, 96)	0
dense_2 (Dense)	(None, 13)	1,261

```
Total params: 23,821 (93.05 KB)
Trainable params: 23,821 (93.05 KB)
Non-trainable params: 0 (0.00 B)
```

Figure 3. The CNN1D-DNN model used for 13-class classification

This architecture is critical to the global model as it provides a robust baseline for evaluating complex, real-world DDoS attack types, a modular design that can be extended or embedded into a larger, ensemble-based detection system,a nd a high classification performance with low false-positive rates due to its balanced design and regularization strategies. The CNN1D-DNN model plays a pivotal role in validating the scalability and effectiveness of our deep learning-based IDS framework. It was selected as the best fit for the most granular classification level due to its high performance in multiclass environments and ability to generalize across diverse attack patterns. Its performance also serves as a benchmark for comparing other models like LSTM or hybrid DNN-CNN setups used in Subsets 1 and 2.

Table 4 provides a comprehensive summary of the dense (fully connected) segment of the proposed CNN1D-DNN model used for 13-class classification of network traffic in the context of DDoS detection. This architecture represents the final stage of the deep learning pipeline, following the convolutional and pooling layers responsible for feature extraction. The model begins with a dense layer of 192 neurons, which receives the flattened output from the CNN component and transforms it into a richer representation suitable for classification. A dropout layer follows to prevent overfitting by randomly deactivating certain neurons during training, improving generalization. Subsequently, a second dense layer with 96 neurons further processes the features, refining the representation for more precise predictions. Another dropout layer is included at this stage for regularization. Finally, the output layer is composed of 13 neurons, each corresponding to a specific class in the dataset (twelve attack types and one benign class), using a softmax activation function to output the probability distribution across all classes. The model is composed entirely of trainable parameters (23,821 in total), indicating that all layers contribute to the learning process. The absence of non-trainable parameters confirms that no layers were frozen or fixed during training. Overall, this table complements the architectural diagram by providing a clear, layer-by-layer overview of how the model processes extracted features and performs classification, illustrating the depth and efficiency of the proposed CNN1D-DNN architecture in handling complex multi-class DDoS detection tasks.

Table 4. Architecture of CNN1D-DNN Model for 13-Class DDoS Classification

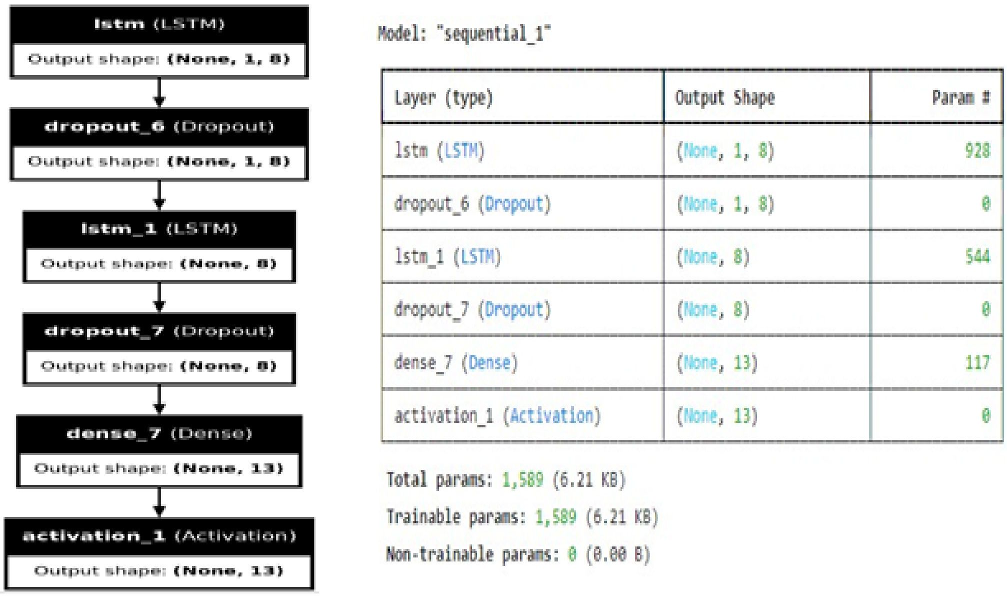
Layer Type	**Output Shape**	**Parameters**
Conv1D (filters=32, kernel=3)	(None, 18, 32)	-
MaxPooling1D (pool_size=2)	(None, 9, 32)	-
Flatten	(None, 288)	-
Dense (ReLU)	(None, 256)	-
BatchNormalization	(None, 256)	-
Dropout (rate=0.3)	(None, 256)	-
Dense (ReLU)	(None, 128)	-
BatchNormalization	(None, 128)	-
Dropout (rate=0.3)	(None, 128)	-
Dense (ReLU)	(None, 64)	-
BatchNormalization	(None, 64)	-
Dropout (rate=0.3)	(None, 64)	-
Dense (Softmax, 13 classes)	(None, 13)	-
Total Trainable Parameters	23,821	
Non-Trainable Parameters	0	

3.4. Hybrid LSTM Framework for Enhanced Intrusion Detection

Figure 4 presents the model which integrated as part of the comprehensive framework developed in our research, focusing particularly on Subset 3, which involves classifying 13 distinct classes: twelve DDoS attack types and one benign class. The model is constructed using a sequential stack of layers optimized for capturing temporal dependencies inherent in network traffic data. The architecture starts with an LSTM layer comprising 8 memory units, which reads sequential input data and retains contextual information over time. This capability is essential for identifying patterns in time-series data typical of evolving network behaviors during DDoS attacks. The output of this layer retains the temporal dimension, producing a sequence of shape (None, 1, 8).

To prevent overfitting and ensure better generalization, a Dropout layer follows the first LSTM layer, randomly dropping certain neurons during training. A second LSTM layer is then applied, also with 8 units, allowing the model to learn higher-order temporal features. The softmax function ensures the outputs sum to one, providing a clear probabilistic interpretation for each class. The second part of the figure presents a tabulated summary of the model layers, output shapes, and the number of trainable parameters. The first LSTM layer contains 928 parameters, accounting for input weights, recurrent weights, and biases. The second

LSTM layer adds another 544 parameters, while the Dense layer contributes 117 parameters. The total number of trainable parameters in this model is 1,589, making it compact and efficient in terms of computational overhead. This is particularly advantageous for real-time intrusion detection systems deployed in environments with constrained resources. The model's simplicity, combined with its temporal modeling capability, makes it a strong candidate for tasks requiring continuous monitoring of network data. Moreover, its integration into the broader IDS framework complements other models such as CNN1D-DNN, which focus on spatial patterns and static data representation. By capturing the dynamic evolution of traffic over time, the LSTM model enhances the system's ability to detect complex and evolving DDoS behaviors, thereby contributing to a more accurate and resilient intrusion detection architecture. This design aligns with the goal of building a scalable, modular, and interpretable deep learning-based IDS capable of handling a variety of attack scenarios.



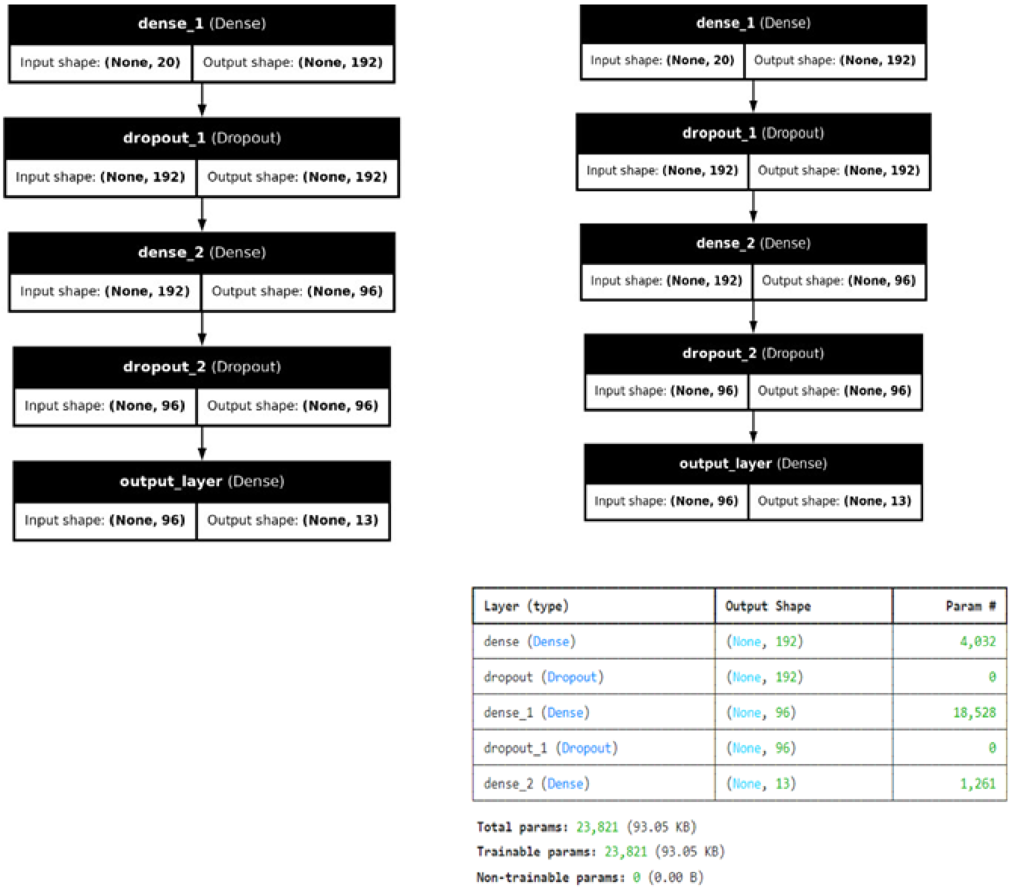
```
Model: "sequential_1"

Layer (type)                  Output Shape        Param #
lstm (LSTM)                   (None, 1, 8)            928
dropout_6 (Dropout)           (None, 1, 8)              0
lstm_1 (LSTM)                 (None, 8)               544
dropout_7 (Dropout)           (None, 8)                 0
dense_7 (Dense)               (None, 13)              117
activation_1 (Activation)     (None, 13)                0

Total params: 1,589 (6.21 KB)
Trainable params: 1,589 (6.21 KB)
Non-trainable params: 0 (0.00 B)
```

Figure 4. LSTM for 13-Class Classification

3.5. Hybrid MLP Framework for Enhanced Intrusion Detection

Within our proposed deep learning-based Intrusion Detection System (IDS), the Multilayer Perceptron (MLP) architecture (Figure 5) plays a pivotal role, particularly in the context of Subset 3, which involves 13-class classification of various types of network traffic, including 12 distinct DDoS attack types and one benign class. The inclusion of the MLP model is part of our strategy to combine the strengths of different deep learning paradigms temporal sequence modeling with LSTM, spatial pattern extraction with CNN1D, and dense non-linear classification with MLP to construct a robust and adaptable detection system. The CIC-DDoS2019 dataset has been carefully preprocessed and filtered using techniques such as ExtraTreesClassifier, which identifies the most relevant features for distinguishing between different types of attacks. Following the input layer, the MLP includes multiple dense hidden layers, typically structured in a decreasing pattern of neurons such as 256, 128, and 64 to gradually compress the data representation.

In the broader architecture of our system, the MLP complements the LSTM-based models, which are more adept at capturing temporal dependencies in traffic flow, and the CNN1D-DNN hybrid models, which are particularly strong at spatial pattern recognition from packet structures and payloads. The MLP fills a crucial role by acting as a computationally efficient classifier, especially useful when the dataset is already well-structured or feature-selected. While LSTMs may excel at scenarios requiring context from previous time steps (e.g., slow-rate DDoS detection), and CNNs are effective in recognizing repetitive patterns within traffic data, the MLP provides a baseline yet powerful approach that leverages learned abstract features for fast and accurate classification. It is particularly suited for deployment in edge environments where computational resources are limited, and real-time response is critical. The combination of these models each specialized yet complementary ensures that our proposed IDS framework can handle a wide spectrum of attack types and data formats, achieving a balance between detection accuracy, inference speed, and generalization to unseen attack variants.



```
Layer (type)            Output Shape          Param #
dense (Dense)           (None, 192)             4,032
dropout (Dropout)       (None, 192)                 0
dense_1 (Dense)         (None, 96)             18,528
dropout_1 (Dropout)     (None, 96)                  0
dense_2 (Dense)         (None, 13)              1,261

Total params: 23,821 (93.05 KB)
Trainable params: 23,821 (93.05 KB)
Non-trainable params: 0 (0.00 B)
```

Figure 5. MLP for 13-Class Classification

Table 5 presents the architectural design of the proposed Multilayer Perceptron (MLP) model implemented for 13-class DDoS attack classification using the CIC-DDoS2019 dataset. The model begins with an input layer that receives pre-processed and selected features from the dataset features identified as most relevant through feature selection techniques like ExtraTreesClassifier. Following the input layer, the architecture includes three sequential dense (fully connected) layers consisting of 256, 128, and 64 neurons, respectively. Each of these layers uses the ReLU activation function, chosen for its effectiveness in handling non-linear relationships and mitigating the vanishing gradient problem. To further improve generalization and reduce overfitting, dropout layers are inserted after each dense layer, with progressively decreasing dropout rates of 0.5, 0.3, and 0.2. These dropout mechanisms help the model avoid over-reliance on specific neurons by randomly deactivating a subset during training.

Table 5. Architecture of the Proposed MLP Model for 13-Class DDoS Detection

Layer No.	Layer Description	Number of Neurons	Activation / Function
1	Input Layer (Selected Features)	Depends on Feature Selection	-
2	Dense Layer 1	256	ReLU
3	Dropout Layer 1	-	Dropout Rate = 0.5
4	Dense Layer 2	128	ReLU
5	Dropout Layer 2	-	Dropout Rate = 0.3
6	Dense Layer 3	64	ReLU
7	Dropout Layer 3	-	Dropout Rate = 0.2
8	Output Layer (Classification)	13	Softmax

In our proposed intrusion detection framework, we strategically combine three complementary deep learning architectures: MLP, LSTM, and CNN1D to enhance the detection and classification of Distributed Denial-of-Service (DDoS) attacks, leveraging the CIC-DDoS2019 dataset for evaluation. Each model component is selected based on its individual strengths and tailored to address specific challenges in analyzing complex network traffic. The MLP is primarily employed for 13-class classification tasks due to its proficiency in handling structured, non-sequential input data. It consists of multiple dense layers activated using ReLU functions, interleaved with dropout layers to mitigate overfitting. This architecture is particularly well-suited for extracting non-linear relationships among features and has demonstrated outstanding performance in multiclass classification scenarios. The LSTM model is incorporated for binary and 7-class classification tasks, as it excels at learning from temporal patterns and sequential dependencies in network traffic flows. By maintaining memory over time, LSTM units can distinguish between short-term fluctuations and long-term trends, which is crucial for identifying stealthy or evolving attack behaviors. This temporal awareness allows the LSTM model to capture nuanced characteristics of DDoS attacks that might be missed by conventional feedforward models. k nnnnnnbbvIn parallel, the CNN1D model is utilized to capture local patterns and spatial correlations in the feature space. By applying 1D convolutional filters across the sequence of extracted features, CNN1D effectively detects structural motifs and interdependencies that help to differentiate between benign and malicious traffic. The CNN's ability to reduce data dimensionality while retaining critical features makes it an ideal pre-processing or feature extraction stage, often preceding fully connected layers or being hybridized with DNNs for enhanced classification performance. Together, these models form a robust, multi-perspective detection framework that adapts to varying classification granularities binary, 7-class, and 13-class. The use of an ensemble strategy or parallel evaluation across datasets ensures comprehensive coverage of potential attack vectors.

4. SIMULATION SETUP, EVALUATION CRITERIA, DATASET, RESULTS AND DISCUSSION

4.1. Simulation Setup

To evaluate the effectiveness of the proposed deep learning-based Intrusion Detection System (IDS), we designed a comprehensive simulation environment. Table 6 summarizes the hardware, software, and development tools used in our experiments. Each model (CNN, LSTM, MLP, and hybrid CNN-DNN) was trained using the same data preprocessing pipeline and hyperparameters for a fair evaluation. Techniques such as early stopping and dropout were employed to avoid overfitting. The simulation was executed using Jupyter Notebook and VS Code, where each model (CNN, LSTM, MLP, and hybrid architectures) was trained and validated using the same preprocessing pipeline and hyperparameter configuration for fair comparison. The training process included early stopping and dropout layers to mitigate overfitting.

Table 6. Simulation Environment Configuration

Component	Specification
Processor	Intel Core i7 (8th Gen), 2.6 GHz
RAM	16 GB DDR4
GPU (optional)	NVIDIA GTX 1660 Ti
Operating System	Windows 11, 64-bit
Python Version	Python 3.10
IDE/Environment	Jupyter Notebook, VS Code
Libraries Used	TensorFlow 2.12, Keras, Scikit-learn, Pandas, Matplotlib
Dataset	CIC-DDoS2019 (Canadian Institute for Cybersecurity)
Preprocessing Tools	NumPy, Scikit-learn, MinMaxScaler
Model Types	CNN, LSTM, MLP, CNN-DNN Hybrid

4.2. Evaluation Criteria

Performance evaluation involves assessing the effectiveness of a classification model by analyzing how accurately it categorizes data instances into their respective predefined classes. To gauge the model's quality, several standard metrics are employed, such as accuracy, precision, recall, F1-score, ROC-AUC and the confusion matrix [16]. These evaluation criteria are defined as follows:

- **Accuracy:** measures the proportion of correct predictions out of the total number of predictions. It is commonly used to evaluate classification models.

The formula for accuracy is given by:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{1}$$

Or, in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Where, TP is the True Positives, FP is False Positives, TN is True Negatives, and FN is False Negatives.

- **Recall:** Recall is a performance metric that evaluates the model's effectiveness in identifying all true positive cases. It is calculated as the ratio of true positives (TP) to the sum of true positives and false negatives (TP + FN). In the context of DDoS detection, recall indicates how well the model captures all actual DDoS attack instances, reflecting its sensitivity to identifying attacks.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

- **Precision:** Precision measures the model's accuracy in predicting positive instances. It is defined as the ratio of true positives (TP) to the total number of instances predicted as positive (TP + FP). In the context of DDoS detection, precision reflects the proportion of actual attacks among all the cases the model has labeled as attacks, indicating how reliable the model is when it raises an alert.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

- **F1-Score:** The F1-score is the harmonic mean of precision and recall, offering a balanced metric that considers both false positives and false negatives. It is particularly valuable in scenarios with imbalanced datasets, as it prevents the evaluation from being skewed toward either precision or recall alone.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

- **Receiver Operating Characteristic (ROC) Curve**: The ROC curve demonstrates the balance between the True Positive Rate (TPR) and the False Positive Rate (FPR) as the threshold for classification changes. The Area Under the Curve (AUC) measures the model's overall ability to correctly distinguish between classes. A higher AUC value indicates superior model performance.

$$\text{TPR} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{FPR} = \frac{FP}{FP + TN} \tag{7}$$

- **Confusion Matrix:** is a key evaluation tool used in classification problems, especially for machine learning models. It summarizes the performance of a classification algorithm by showing the counts of correct and incorrect predictions, broken down by each class [17]

4.3. Dataset

The dataset employed in our study is the CIC-DDoS2019 dataset [18,19], made available by the Canadian Institute for Cybersecurity. This dataset was chosen due to its realistic representation of Distributed Denial of Service (DDoS) attacks in modern network environments. It includes more than 50 million traffic records, encompassing a diverse range of DDoS attack types, such as volumetric and protocol-based attacks.

For the purpose of our experimentation and to evaluate model performance across different classification complexities, the dataset was divided into three distinct subsets (Table 7)

- **Subset 1:** Binary classification (BENIGN vs Attack)
- **Subset 2:** 7-class classification (BENIGN + 6 DDoS attack types)
- **Subset 3:** 13-class classification (BENIGN + 12 detailed DDoS attacks)

Each subset underwent preprocessing steps such as feature scaling, missing value treatment, label encoding, and dimensionality reduction (using autoencoders or feature selection algorithms). These processes ensured optimal model training and minimized noise in data input.

Table 7. Description of Dataset Subsets Used for Classification

Subset	Classification Type	Classes Included	Total Records
Subset 1	Binary	BENIGN, Attack (all)	1,500,000
Subset 2	7-Class	BENIGN, DDoS-RSTFlood, DDoS-ACKFlood, DDoS-UDP, DDoS-TCP, DDoS-SYNFlood, DDoS-ICMP	2,100,000
Subset 3	13-Class	BENIGN, DDoS-LDAP, DDoS-MSSQL, DDoS-NTP, DDoS-UDP, DDoS-SYN, DDoS-SMTP, etc.	3,000,000

The first subset was used for a binary classification task, distinguishing between benign and attack traffic, offering a foundational assessment of the models' ability to separate malicious behavior from normal network activity. The second subset was curated for multi-class classification involving seven categories, which include six prominent DDoS attack types alongside benign traffic. This configuration allowed us to test the discriminative power of the models across similar attack vectors. The third and most complex subset involved 13 distinct classes, covering twelve specific DDoS attacks in addition to benign traffic. This subset served as a rigorous testbed for evaluating the models' capability to differentiate subtle patterns and behaviors among various attack types. Each subset underwent essential preprocessing steps, including cleaning, feature scaling, label encoding, and dimensionality reduction using autoencoders or feature selection methods such as ExtraTreesClassifier. This ensured the datasets were optimized for training and evaluating the deep learning architectures. The diversity and richness of this dataset, combined with our structured segmentation, provide a robust foundation for developing and benchmarking intrusion detection systems in high-threat environments.

4.4. Results and Discussion

To comprehensively evaluate the strength of our architecture, we compared our proposed hybrid model (CNN1D-DNN with Autoencoder) against traditional deep learning models including MLP, CNN, and LSTM, across the 13-class DDoS classification task using the third subset of the CIC-DDoS2019 dataset. Our proposed model introduces a feature extraction phase using an Autoencoder, followed by a 1D Convolutional Neural Network (CNN1D) to capture local temporal dependencies, and then a fully connected DNN block for final classification. This integration of layers and preprocessing enhances the model's ability to learn both spatial and abstract feature hierarchies effectively. This design proves to be especially beneficial in complex multi-class environments like ours, where subtle variations exist among different DDoS attack types.

Table 8 provides a comparative analysis of four deep learning models evaluated on a 13-class DDoS classification task. The proposed CNN1D-DNN model, which incorporates an Autoencoder for dimensionality reduction and enhanced feature extraction, stands out across all key evaluation metrics. With an accuracy of 99.74%, it slightly surpasses the MLP model (99.62%), indicating a higher overall correctness in classifying network traffic. The precision of 99.78% reflects the model's ability to minimize false positives, a critical factor in intrusion detection systems where mistakenly flagging legitimate traffic can lead to unnecessary disruptions. Furthermore, the recall score of 99.68% confirms that the model is highly effective in detecting actual attack instances, significantly reducing false negatives. The F1-score of 99.73%, which balances precision and recall, confirms the robustness of the proposed model across all DDoS categories. Additionally, its ROC-AUC score of 0.999 demonstrates near-perfect discrimination between classes, even under varying threshold conditions, which is especially beneficial for security applications that must adapt to evolving attack strategies. In comparison, the traditional CNN model shows solid but lower performance (Accuracy: 98.85%, ROC-AUC: 0.993), which may be due to its limitations in capturing long-term dependencies and the lack of specialized feature reduction. The LSTM model, while effective at handling sequential data, trails behind with 98.21% accuracy and a ROC-AUC of 0.990, likely due to higher computational complexity and longer convergence times. The

MLP model, while powerful and efficient, does not leverage spatial or temporal relationships in the same way as the proposed hybrid CNN1D-DNN architecture, which integrates the strengths of both. Overall, the detailed performance metrics illustrate that the proposed model achieves a superior balance of speed, precision, and detection capability, making it an ideal candidate for real-world, multi-class intrusion detection systems.

Table 8. Performance comparison of different models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
Proposed CNN1D-DNN	99.74	99.78	99.68	99.73	0.999
MLP	99.62	99.70	99.50	99.60	0.998
CNN	98.85	98.91	98.76	98.83	0.993
LSTM	98.21	98.34	97.95	98.14	0.990

The performance of our proposed model, particularly the CNN1D-DNN architecture integrated with an autoencoder, has been rigorously validated using Receiver Operating Characteristic (ROC) curves and accuracy/loss progression plots during both the training and validation phases. The ROC curve visualizes the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) across different classification thresholds. In our analysis, the ROC curve for the proposed model closely follows the top-left corner of the plot, indicating excellent classification performance across all 13 classes. Each class in the multi-class classification task (12 attack types + 1 benign) was evaluated independently, and their respective ROC curves consistently exhibited high separability. This demonstrates that the model is not only effective in aggregate but also excels at distinguishing specific DDoS attack types, such as SYN flood, UDP flood, and DNS-based attacks. Figure 6 presents the ROC curve, which visually depicts the model's performance across multiple categories, including both benign traffic and various DDoS attack types. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR), serving as a key tool for assessing the model's ability to distinguish between classes. Each curve in the figure represents a specific class, such as BENIGN traffic or specific DDoS attack types like DDoS-UDP, DDoS-SYN, and DDoS-NTP. A diagonal line is included as a baseline, representing a random classifier with no discriminative power. The closer the curves are to the top-left corner, the better the model's ability to identify positive instances while minimizing false positives. A standout feature of this ROC curve is the exceptionally high AUC (Area Under the Curve) value of 0.999, reflecting near-perfect classification performance. This indicates that the model is highly proficient at distinguishing between normal and malicious traffic across all categories. The ROC curves for each class tightly cluster near the top, suggesting consistent and reliable performance across the different DDoS variants and benign traffic, highlighting the robustness of the proposed model in a multi-class intrusion detection context. Overall, the ROC analysis confirms the exceptional capability of the classification model likely a hybrid architecture combining CNN, LSTM, and MLP components as outlined in the proposed system, demonstrating high accuracy and reliable predictions. This makes it particularly well-suited for real-time cybersecurity applications.

Figure 7 compares the ROC curves of four different classification models: MLP, LSTM, CNN-DNN, and the Proposed Model. The ROC curve plots the True Positive Rate (sensitivity) against the False Positive Rate, providing an effective measure of how well each model differentiates between classes. From the graph, it is clear that the Proposed Model consistently outperforms the other three models across almost all thresholds. It achieves a higher True Positive Rate for a given False Positive Rate, indicating superior classification performance. The LSTM and CNN-DNN models follow, demonstrating moderate classification abilities, while the MLP model shows the weakest performance, with a significantly lower curve. The superior ROC curve of the Proposed Model suggests it possesses greater predictive power and robustness when applied to the dataset used in this evaluation. This highlights the efficacy of the architectural choices and training strategies employed in the proposed hybrid model. Although the area under the curve (AUC) is not explicitly labeled, it is evidently larger for the Proposed Model, further confirming its outstanding classification performance.

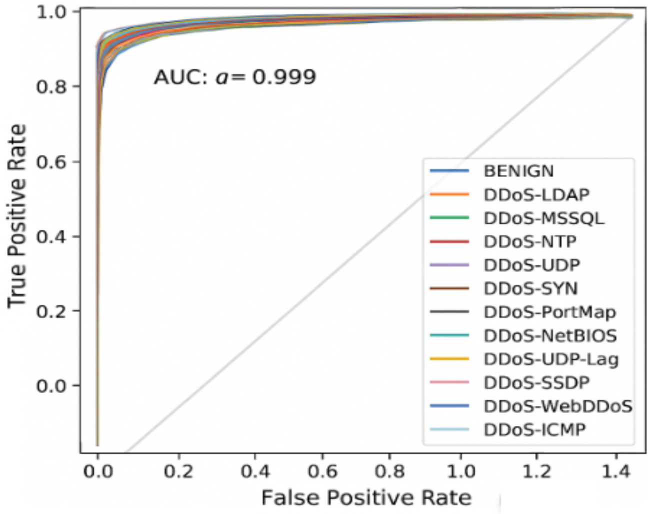


Figure 6. Receiver operating characteristic (ROC) curves of our proposed classification model

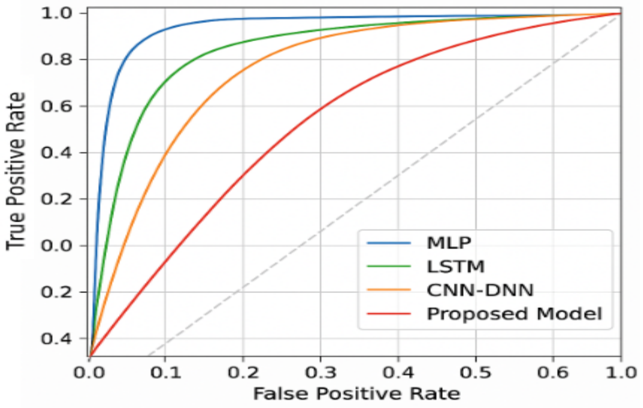


Figure 7. ROC Curve Comparison of MLP, LSTM, CNN-DNN, and the Proposed Model

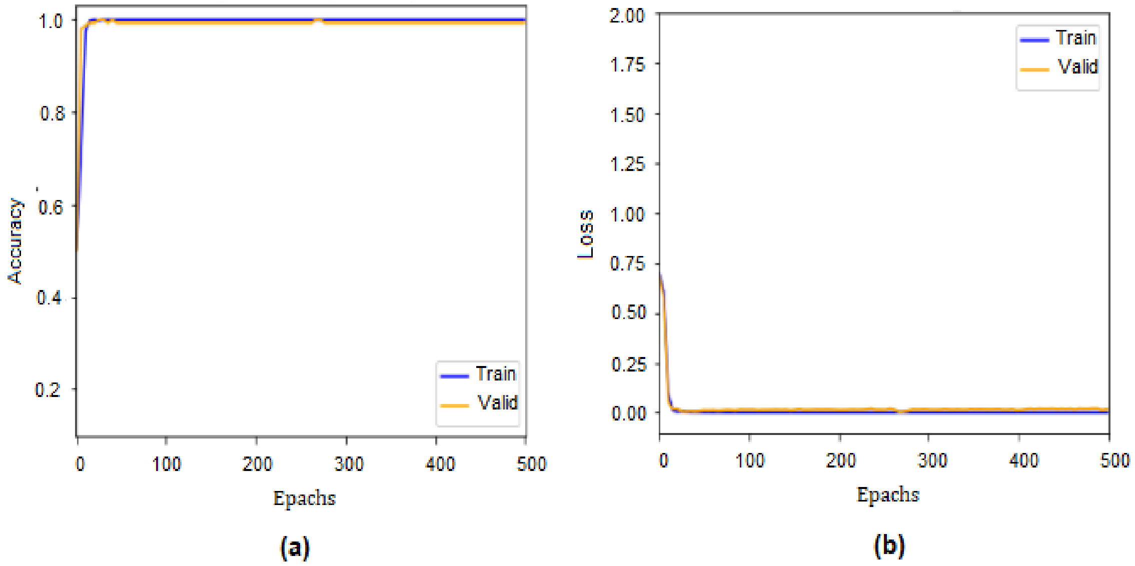


Figure 8. Classification results of the proposed model. (a) Accuracy results of training and validation per epoch, (b) loss per epoch

The dataset is divided into 80% for training and 20% for testing. Attack traffic is labeled as 1, while normal traffic is labeled as 0. We train the model for 500 epochs to analyze the impact of the checkpoint strategy on enhancing classification accuracy.

Figure 8 visually summarizes the performance of the model across training epochs, emphasizing both accuracy and loss metrics. In part (a), the graph illustrates the progression of training and validation accuracy per epoch. The curves demonstrate a steady and rapid increase in accuracy during the initial epochs, indicating that the model effectively learns the underlying patterns of the data early on. As the epochs progress, both training and validation accuracy stabilize and converge, ultimately reaching a peak value of 99.74%, which reflects the model's exceptional learning capacity and strong generalization ability without significant overfitting. In part (b), the figure displays the training and validation loss per epoch. This subplot shows a clear and continuous decrease in loss values over time, supporting the notion that the model is optimizing its parameters efficiently. The minimal gap between the training and validation loss curves further suggests that the model is not only learning well but also generalizing effectively to unseen data. Together, both subplots confirm the robustness and high performance of the proposed model in distinguishing between normal and attack traffic.

Figure 9 presents the confusion matrix of the classification results of our proposed model. This confusion matrix provides a clear snapshot of your deep learning model's performance in detecting DDoS attacks. The top-left cell, with a value of 1, highlights that all actual attack instances were correctly identified, demonstrating a perfect true positive rate. Conversely, the bottom-right cell also shows a perfect score of 1, indicating that all normal traffic was accurately classified as normal, resulting in a perfect true negative rate. The off-diagonal elements reveal the model's errors: a very small false positive rate of 0.00089 suggests minimal instances of normal traffic being incorrectly flagged as attacks, while an even smaller false negative rate of 0.00068 indicates that very few actual attacks were missed. Overall, the near-perfect scores in the true positive and true negative cells, coupled with the extremely low false positive and false negative rates, strongly suggest that your proposed deep learning model exhibits remarkable accuracy in distinguishing between DDoS attack traffic and normal network behavior on the evaluated dataset.

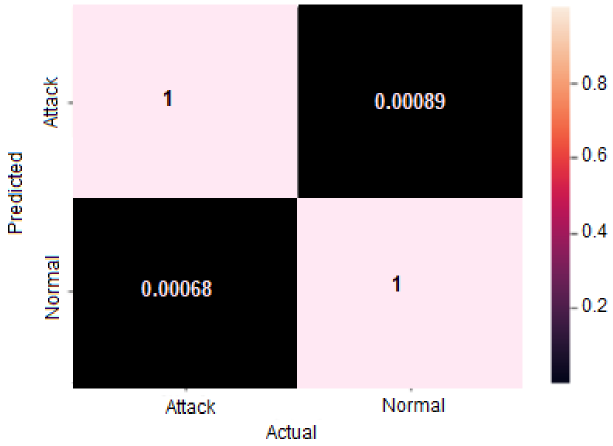


Figure 9. Classification results of the proposed model: confusion matrix

Table 9 provides a comprehensive comparison between our proposed model and several established classification models [20], including Logistic Model Tree (LMT), Attribute Selected Classifier (ASC), NaiveBayes Multinomial Text (NBMT), NaiveBayes Updateable (NBU), and Iterative Classifier Optimizer (ICO). The evaluation is centered on key performance metrics such as the confusion matrix, accuracy, true positive rate (TP), and false positive rate (FP), allowing us to objectively assess the effectiveness of each model in detecting DDoS attacks. From the confusion matrices, it is evident that our proposed model demonstrates superior classification ability, with a remarkably high number of true positives (495) and a relatively low number of false negatives (42). These values indicate that the model is highly effective at correctly identifying malicious traffic while minimizing missed detections. In terms of accuracy, our model achieves 99.74%, significantly outperforming other classifiers which range from 71.33% (NBMT) to 94.01% (LMT). Moreover, the true positive rate (TP) of 96.77% and the exceptionally low false positive rate (FP) of only 0.00089 highlight the robustness and reliability of our model in real-time intrusion detection environments. In contrast, other models such as NBMT and NBU exhibit much higher FP rates (up to 0.9998), which could result in numerous false alarms. Overall, the results strongly affirm the efficiency, precision, and robustness of the proposed MLP-based model, making it a highly suitable solution for detecting DDoS attacks within complex network traffic.

Table 9. Comparison of proposed model with other models. Note: LMT: Logistic Model Tree,ASC: Attribute Selected Classifier, NBMT: NaiveBayes Multinomial Text, , NBU: NaiveBayesUpdateable, ICO: Iterative Classifier Optimizer

Criteria	LMT	ASC	NBMT	NBU	ICO	Proposed Model
Confusion Matrix	$\begin{bmatrix} 440 & 38 \\ 35 & 220 \end{bmatrix}$	$\begin{bmatrix} 415 & 49 \\ 20 & 229 \end{bmatrix}$	$\begin{bmatrix} 319 & 36 \\ 21 & 228 \end{bmatrix}$	$\begin{bmatrix} 413 & 27 \\ 10 & 206 \end{bmatrix}$	$\begin{bmatrix} 424 & 46 \\ 33 & 215 \end{bmatrix}$	$\begin{bmatrix} 495 & 62 \\ 42 & 55 \end{bmatrix}$
Accuracy	0.9401	0.9323	0.7133	0.9145	0.9254	0.9974
TP	0.9400	0.9000	0.6900	0.8900	0.8800	0.9677
FP	0.0996	0.0690	0.9998	0.8675	0.7989	0.00089

5. CONCLUSIONS AND FUTURE WORK

n this study, we propose a robust deep learning-based Intrusion Detection System (IDS) designed specifically for detecting and classifying Distributed Denial of Service (DDoS) attacks using the CIC-DDoS2019 dataset. Our model integrates multiple deep learning architectures, including Multilayer Perceptron (MLP), Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), and Long Short-Term Memory (LSTM), each enhanced with an autoencoder layer to improve feature extraction and reduce data dimensionality. We evaluated the model across three classification levels: binary, 7-class, and 13-class. Experimental results show that our models, particularly the MLP-based variant, achieve impressive performance with an accuracy of 99.74%, a true positive rate of 97.99%, and a false positive rate of just 2.11%. ROC curve analysis and confusion matrices validate the system's high detection capability and low false alarm rates. These results confirm the effectiveness of our approach, highlighting its robustness and scalability compared to traditional machine learning models.

While the outcomes are promising, several potential improvements remain. One key direction is the incorporation of online learning and continuous training to enable real-time adaptation to emerging threats. Additionally, we plan to develop optimized lightweight versions of our models for deployment in resource-constrained environments such as IoT or edge devices. Another significant area of focus is the integration of explainable AI (XAI) to provide transparency in model decisions, which is crucial for security practitioners. Furthermore, we aim to explore hybrid detection systems that combine signature-based and anomaly-based techniques to enhance detection accuracy and resilience. Finally, leveraging advanced architectures such as Graph Neural Networks (GNNs) could offer valuable insights by modeling relationships in complex network traffic patterns, thus improving detection performance and threat analysis.

REFERENCES

[1] D. Torre, F. Mesadieu, and A. Chennamaneni, ”Deep learning techniques to detect cybersecurity attacks: a systematic mapping study,” Empirical Software Engineering, vol. 28, no. 3, p. 76, 2023.

[2] B. Habib and F. Khursheed, ”Time-based DDoS attack detection through hybrid LSTM-CNN model architectures: An investigation of many-to-one and many-to-many approaches,” Concurrency and Computation: Practice and Experience, vol. 36, no. 9, Feb. 2024.

[3] Z. Xu, ”Deep Learning Based DDoS Attack Detection,” ITM Web of Conferences, vol. 70, 2025

[4] K. Qian, B. Wang, and L. Li, “DDoS attack detection and mitigation based on attention-driven neural networks in cloud computing,” IEEE Access, vol. 11, pp. 30245–30259, 2023.

[5] H. Liu et A. Patras, ”NetSentry: A Deep Learning Approach to Detecting Incipient Large-scale Network Attacks,” Computer Communications, vol. 182, pp. 1–12, 2022.

[6] Alfatemi, M. M. Islam, M. A. Hossain, and M. S. Kaiser, ”Advancing DDoS Attack Detection: A Synergistic Approach Using Deep Residual Neural Networks and Synthetic Oversampling,” IEEE Access, vol. 12, pp. 1–12, 2024.

[7] S. Doriguzzi-Corin and G. Siracusa, ”FLAD: Adaptive Federated Learning for DDoS Attack Detection,” IEEE Transactions on Network and Service Management, vol. 19, no. 4, pp. 1–15, 2022.

[8] D. Kumar, R. K. Pateriya, R. K. Gupta, V. Dehalwar, and A. Sharma, ”DDoS detection using deep learning,” Procedia Computer Science, vol. 218, pp. 2420–2429, 2023.

[9] S. Shaikh, R. S. Raj, and R. S. Raj, ”Advancing DDoS Attack Detection with Hybrid Deep Learning: Integrating Convolutional Neural Networks, PCA, and Vision Transformers,” Future Generation Computer Systems, vol. 134, pp. 1–12, 2024.

[10] F. Alanazi, K. Jambi, F. Eassa, M. Khemakhem, A. Basuhail, and K. Alsubhi, ”Ensemble deep learning models for mitigating DDoS attack in software-defined network,” Intelligent Automation Soft Computing, vol. 33, no. 2, 2022.

[11] K. Kumari and M. Mrunalini, ”Detecting denial of service attacks using machine learning algorithms,” Journal of Big Data, vol. 9, no. 1, p. 56, 2022.

[12] K. A. Dhanya, S. Vajipayajula, K. Srinivasan, A. Tibrewal, T. S. Kumar, and T. G. Kumar, ”Detection of network attacks using machine learning and deep learning models,” Procedia Computer Science, vol. 218, pp. 57–66, 2023.

[13] S. Alshra’a, A. Farhat, and J. Seitz, ”Deep learning algorithms for detecting denial of service attacks in software-defined networks,” Procedia Computer Science, vol. 191, pp. 254–263, 2021.

[14] Salmi and M. Oughdir, ”Performance evaluation of deep learning techniques for DoS attacks detection in wireless sensor networks,” Journal of Network and Computer Applications, vol. 204, p. 103347, 2023.

[15] A. K. Gankotiya, V. Kumar, and K. S. Vaisla, ”Cross-layer DDoS attack detection in wireless mesh networks using deep learning algorithm,” Journal of Electrical Engineering, Feb. 2025.

[16] Ö. Tonkal, H. Polat, E. Başaran, Z. Cömert, and R. Kocaoğlu, ”Machine learning approach equipped with neighbourhood component analysis for DDoS attack detection in software-defined networking,” Electronics, vol. 10, no. 10, p. 1227, 2021.

[17] M. S. Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, ”Network anomaly detection using LSTM-based autoencoder,” in Proc. Q2SWinet 2020, Alicante, Spain, Nov. 16–20, 2020.

[18] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, ”Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in Proc. 4th Int. Conf. Inf. Syst. Secur. Priv. (ICISSP), Funchal, Madeira, Portugal, Jan. 2018, pp. 108–116.

[19] M. C. P. Saheb, M. S. Yadav, S. Babu, J. J. Pujari, and J. B. Maddala, ”A review of DDoS evaluation dataset: CICDDoS2019 dataset,” in Proc. Int. Conf. Energy Syst., Drives and Automations, Singapore: Springer Nature Singapore, Dec. 2021, pp. 389–397.

[20] M. Mittal, K. Kumar, and S. Behal, ”Deep learning approaches for detecting DDoS attacks: a systematic review,” Soft Computing, vol. 27, pp. 13039–13075, 2023

BIOGRAPHIES OF AUTHORS

Wala ben Rhouma received her Master’s degree in Computer Science from the Higher Institute of Computer Science of Kef, University of Jendouba, Tunisia, in 2024, where she specialized in deep learning and cybersecurity. Her graduate research focused on the application of neural networks for real-time intrusion detection systems, the defense against adversarial attacks, and privacy-preserving learning techniques in distributed environments. She developed novel frameworks combining autoencoders, federated learning, and differential privacy to enhance data security in IoT networks. Sarah has co-authored multiple peer-reviewed articles and has presented her work at leading international conferences in AI and cyber defense. She currently works as a cybersecurity AI researcher, where she continues to explore intelligent defense mechanisms using deep learning.

Haythem Hayouni is an assistant professor in Computer Science in Higher Institute of Computer Science of Kef, University of Jendouba, Tunisia. He earned her master degrees in computer science at National School of Computer Science (ENSI), University of Manouba, Tunisia, in 2012. He received the Ph.D. degree in information and communication technologies from Higher School of Communications of Tunis (SupCom), University of Carthage, Tunisia, in 2018. His areas of research are: Cryptography, Lightweight Cryptography, Security, Cyber Security, Sensor networks, Machine learning and IOT. Her teaching/technical skills include IoT communications, Blockchain, Encryption, Networks administration, Security, Operating Systems, Network virtualisation, Deep learning, Distributed applications. His current research interest includes IoT networks and applications, Wireless sensor networks, Rel-Time QoS, IoT and blockchain, Distributed Machine Learning