Vokasi Unesa Bulletin of Engineering, Technology and Applied Science (VUBETA)
https://journal.unesa.ac.id/index.php/vubeta
Vol. 3, No. 2, 2026, pp. 389~411 DOI: 10.26740/vubeta.v3i2.40146
ISSN: 3064-0768

Machine Learning Models for DDoS Attack Detection: A Systematic Literature Review

Chinyere Chioma Isiekwene[1*], Nureni Ayofe Azeez[2], Solomon A. Akinboro[3], Oladipupo Sennaike[4]
[1,2,3,4]Department of Computer Sciences, Faculty of Science, University of Lagos, Nigeria
[1]Department of Information Technology, Faculty of Computing, MIVA Open University, Nigeria

Article Info

Article history:

Received May 04, 2025
Revised December 03, 2025
Accepted February 15, 2026

Keywords:

Systematic Literature Review
Machine learning models
DDoS
Hybrid models
Networks

ABSTRACT

The study aims to present a detailed analysis of different machine learning models used in the detection of distributed denial of service (DDoS) attacks. The report adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) style to determine the research domain, established a search list, and analyzed all the selected articles from scientific databases such as IEEE, Springer, Elsevier, MDPI, SSRN-JETIR, Wiley online-library, and Google Scholar to meet eligibility criteria. A total of 6560 articles were retrieved, and 75 were deemed eligible for study. The review identified seven subject categories in the literature review, and the results show that 48% of the reviewed papers were from Elsevier (Science Direct), IEEE covered 20%, Springer covered 16%, while MDPI count was 10.67%. 2023 had the highest number of paper sources, followed closely by 2022, then 2024. The study reveals the milestone achieved in the use of machine learning models in detecting distributed denial of service attacks alongside the existing gap in the application of these models.

1. INTRODUCTION

Machine learning studies methods based on data that can replicate, comprehend, analyze, identify, and support genetic and human information-processing tasks. It is a subset of artificial intelligence [46]. Numerous related problems arise, including those involving data collection, sorting, compression, interpretation, and processing. These techniques are frequently used to improve data processing, such as quickly forecasting an event's result, rather than to precisely duplicate human procedures. Subdivided further into Semi-Supervised [59] (Reinforcement), Unsupervised, and Supervised learning respectively, [62] Machine learning has redefined the narrative of data representation, thereby promoting data storytelling. Additionally, Machine learning algorithms may require specific conditions for effectiveness, thereby necessitating hyperparameter tuning [8], [10], [18], [43], [47], [53]. However, despite the use of these models in prediction, the problem of accuracy in detecting DDoS still persists.

The problems manifest through communication and information access, which have been made easier by the widespread use of internet-connected digital devices and the continuous advancement of network technologies. Because of its convenience, cybercriminals are drawn to it, putting customers' access to a wide range of services at risk. Network security follows rules when handling malicious requests, but customers nonetheless encounter several problems, including lost or unavailable resources, sensitive information about them or their organizations, and so forth. Therefore, several strategies, including decision theory, stochastic simulation, and game theory, have been investigated to stop and recognize these attacks [36]. However, these models are not capable of identifying unidentified attacks. Furthermore, the most vulnerable industries are those in healthcare, IT, finance, higher education, telecommunications, energy, and government. Attackers may differ in their intentions for launching an assault, as there are five main reasons they could do so: financial gain, retaliation, ideological conviction, intellectual challenge, and cyber warfare [16]. Given

*Corresponding Author
Email: isiekwenechioma@gmail.com

that these attacks are increasing, it's critical to identify and stop assaults early, before they reach their targets. It's getting harder to determine the Distributed Denial of Service (DDoS) attacks [1].

2. REVIEW METHODOLOGY

This research could be categorized as an identified gap in the attacks known as distributed denial-of-service (DDoS), which is becoming increasingly difficult to detect. Several models have recently been published in the literature to detect them, but due to the large variations in traffic rates and signatures, the problem remains difficult to solve.

2.1 Research question:

The following are the research questions we have defined for our study:
RQ1: What is the state of the art in the OSI model generally in terms of protocols?
RQ2: Which security issues are addressed in each phase of the OSI model and which layer is most vulnerable?
RQ3: Based on the gaps identified, what are the current gaps on which further research focuses?
RQ4: Which layer of the OSI model has been considered for developing a new approach?
RQ5: What algorithms have been selected for developing the new approach?
RQ6: What are the main selected criteria in approach the evaluation [22]? Model development for selecting the model with the highest accuracy?
RQ7: How can the best decision be made among several results produced?

2.2 Research procedures and selection criteria:

The methods for carrying out the selection process, as well as the inclusion/exclusion criteria that define the parameters for the systematic review [44], are covered in this section, and Figure 1 below shows the review structure, which is divided into three phases, namely: planning, conducting, and documentation phases.

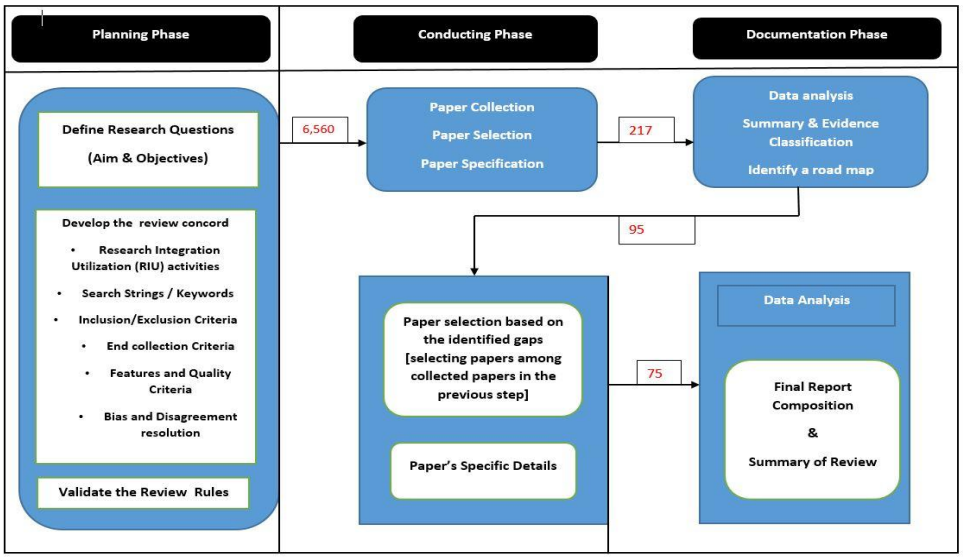

Figure 1. Review Structure Diagram

2.2.1 Inclusion/Exclusion

The following are the inclusion criteria used to choose the papers:

- Publishes conferences, workshops, journals, and peer-reviewed publications that, at one point or another in their life cycle, address any aspect of DDoS, Hyperparameter tuning, Feature Selection, Data Imbalance [5], [41], or Machine Learning in the Application Layer [8], [10], [18], [43], [47], [53], [72].
- Any earlier analyses of the literature in this field.

The following are the exclusion standards:

- Inadequately documented studies that are only accessible as abstracts, presentations, or in other fragmentary forms.
- Several reports on the same investigation. When a study has multiple reports published in different journals, the review includes the most comprehensive version.

2.3 Strategy for searching

Seven scientific databases—MDPI, IEEE, ScienceDirect (Elsevier), Springer, Taylor & Francis, SSRN-JETIR, and Wiley Online Library— were searched. To ensure this review was more thorough, we also scanned the reference lists provided in the publications. Only English-language articles released between 2015 and 2024 were included in the search.

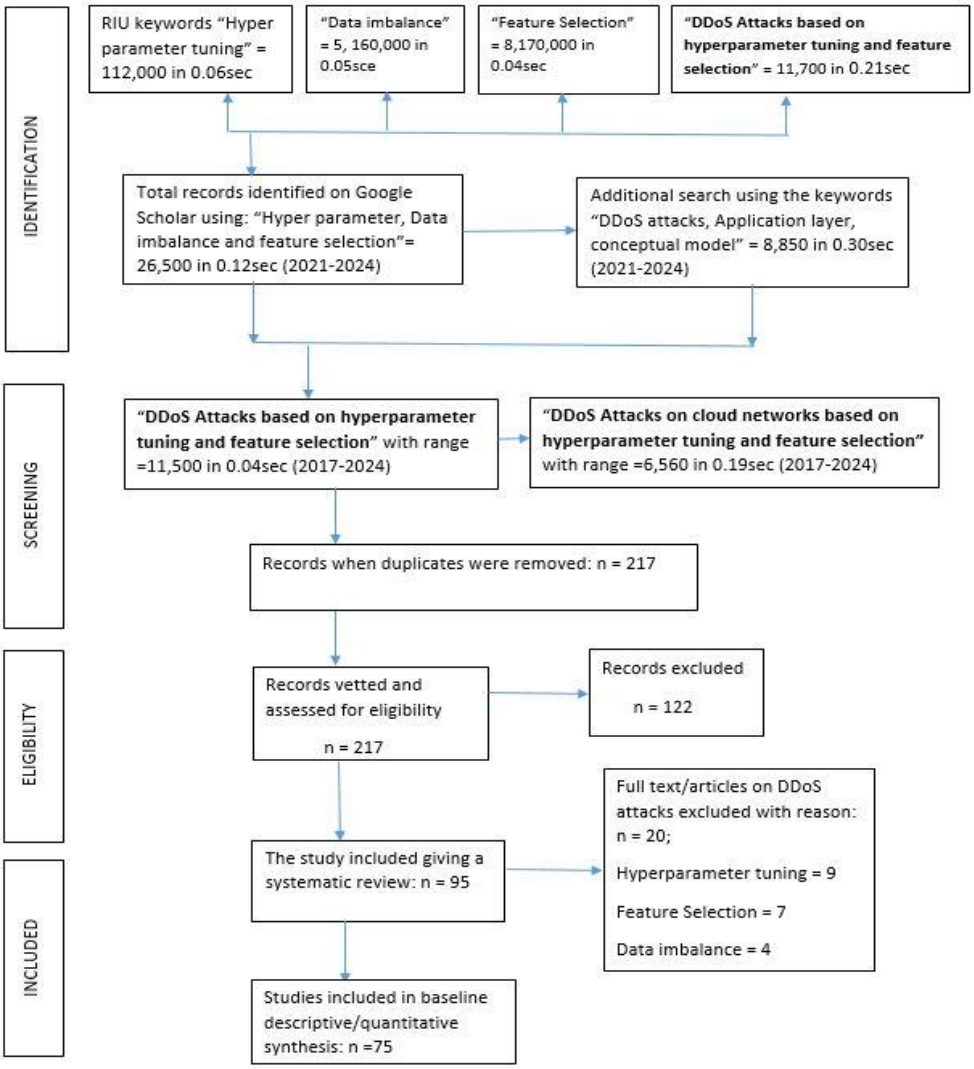

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Diagram.

3. LITERATURE REVIEW

This work thoroughly examines the relevant literature, particularly in machine learning, to detect DDoS attacks. The authors searched the current literature using one academic search engine (Google Scholar) and four well-known digital libraries (IEEE, ACM, ScienceDirect, and Springer). The findings of the SLR are divided into five primary research topics following an assessment of the pertinent literature: The following subjects are covered by the current study:

(i) the different kinds of deep learning techniques for detecting DDoS attacks [2], [31], [32];
(ii) the techniques, advantages, and disadvantages of the deep learning techniques currently used for this purpose [3], [4], [19];
(iii) benchmarked datasets and attack classes in datasets used in the literature;
(iv) preprocessing techniques, hyperparameter values, experimental configurations, and performance metrics;
(v) the research gaps and potential future directions [61], [62], [69].

Halladay et al. [23] provided a thorough analysis that offered an in-depth understanding of the mechanisms underlying the characterization, prevention, detection, and mitigation of these assaults. They also extensively discuss crucial indicators for assessing different solutions and a detailed taxonomy of DDoS attack remedies.

3.1 Data imbalance (under-fitting and overfitting)

A new automatic detection technique is developed by narrowing the feature space, thereby reducing model calculation time and overfitting. The most relevant features are then selected via feature selection, thereby enhancing classification accuracy. Supervised learning techniques, including support vector machines (SVM), gradient boosting (GB), K-nearest neighbor (KNN), decision trees (DT), and logistic regression (LR). The CICDDoS2019 dataset is used to weigh each of these studies. With an accuracy of 99.97%, the experimental findings demonstrate that the GB model performed well compared to cutting-edge techniques [12].

The writer created a reinforcement learning model based on proximal policy optimization (PPO). This study proposed an intrusion detection hyperparameter control system (IDHCS) that tracks and trains a deep neural network (DNN) feature extractor and a k-means clustering module [5]. The CICIDS2017 and UNSW-NB15 datasets were used in experiments to evaluate the system's performance. CICIDS2017 received an F1-score of 0.96552, whilst UNSW-NB15 had an F1-score of 0.94268. The merging of datasets results in a more complex and diverse set of attack kinds and patterns in the experiment. The combined dataset yielded an F1-score of 0.93567, indicating a performance of 97% to 99% compared to CICIDS2017 and UNSW-NB15.

Mazumder [41] 2023 First, information is collected from UNSW-NB15 and CSE-CIC-IDS 2018 datasets. Missing values and unnecessary parameters are removed using null-value handling. In the pre-processing phase, the data were normalized using Min-Max normalization, and to mitigate class imbalance, the Advanced Synthetic Minority Oversampling Technique (ASmoT) was applied after pre-processing. According to the performance metrics, the suggested system used the CSE-CIC-IDS 2018 dataset to achieve 99.89% accuracy and the UNSW-NB15 dataset to achieve 97.535% accuracy.

3.2 Development of novel techniques and detecting malicious packets in real-time [33]

Nafiseh [48] offered a thorough analysis of the security concerns and difficulties that cloud service providers and their customers face when using CC. Additionally, this study proposes a new categorization of distributed denial-of-service (DDoS) attacks and CC attacks, as well as methods for detecting CC DDoS attacks. A qualitative comparison with the latest surveys was also provided. Last but not least, the survey [29] was intended to serve as a manual and reference point for other academics developing new techniques for detecting DDoS attacks within the CC framework.

Machine learning methods like Logistic Regression and Naive Bayes are used to identify assaults and similar scenarios. The CAIDA 2007 Dataset was used in the experimental investigation. Machine learning algorithms are trained and tested on this dataset, and the results verify the algorithms learned behavior [36], [46]. It has also been observed that machine-learning models perform somewhat better than mathematical models. Whereas the machine learning model has 100% accuracy, the mathematical model has 99.75%.

This paper develops a model to detect DDoS attacks on a subset of network traffic packets using a "Long Short-Term Memory (LSTM)" based architecture. The LSTM deep learning technique includes an algorithm for feature selection and extraction. After training, LSTM self-updates to function swiftly and precisely with fewer data points. In this recent work, the proposed LSTM model, trained and tested on the "CICDDoS2019 dataset," achieves up to 98% accuracy. On the CICDDoS2019 dataset, deep learning outperforms machine learning [11], [45].

3.3 Feature selection and engineering

This study suggests using machine learning and feature engineering to identify DDoS attacks in SDNs. After cleaning and normalizing the CSE-CIC-IDS2018 dataset, an improved binary grey wolf optimization method was used to identify the optimal feature subset. [37]. The dataset's features decreased from 79 to 26 following feature extraction, yet despite the fewer features, all of the classifiers performed better across the board. The maximum accuracy of XGBoost was 0.969 on the original dataset. With scores of 0.9913, 0.9843, 0.9992, and 0.9913 in precision, recall, accuracy, and f1_score, respectively, the RF classifier outperformed the others.

To detect DDoS attacks, 8 supervised machine learning approaches are selected, and the model with the highest accuracy, precision, recall, and false alarm rate is identified. The CIC-IDS2017 benchmark dataset is used for training and testing to obtain the experimental results. The K-Fold method is used for cross-validation before preprocessing. K-Fold cross-validation is then used to train and evaluate the eight

models to determine which model is most effective at quickly detecting the DDoS attack. After looking at each parameter, we concluded that Random Forest is the best of the eight models. To identify DDoS attacks as quickly as possible, it has produced 99.885% accuracy, 99.88% precision, 100% recall, and a 0.05% false alarm rate [50].

This study included six distinct types of machine learning algorithms: XGBoost, AdaBoost, Random Forest, Support Vector Machine, Naive Bayes, and Decision Tree. Based on the results, AdaBoost achieves a maximum accuracy of 99.87% in 27.4 seconds. Naive Bayes beats the Support Vector Machine, which has the lowest accuracy (95.73%) and the longest learning time (229 m26s for training and 0.2 s for prediction), with a 94.15% accuracy rate and the fastest computing time (3.2 s). The goal of this research project is to examine and select a set of data to represent DDoS attack occurrences and attack flow statistics. After pre-processing the data to clean and modify it, a machine learning model is built for multi-class categorization. This is done to classify the various types of DDoS attacks [7].

Sayed et al. [61] proposed a Long Short-Term Memory (LSTM) and an Auto-encoder, a technique based on Deep Learning (DL). Accuracy rates for the CICIDS2018 and InSDN datasets are 55.18% and 99.61%, respectively. The model failed to identify any DDoS entities in the InSDN dataset. For the CICIDS2018 and InSDN datasets, the overall accuracies using the RF and IG techniques are 88.21% and 36.22%, and 89.09% and 32.94%, respectively. On the other hand, overfitting is a major issue that substantially impairs the effectiveness of ML/DL models. During training, the model performs quite well; however, it shows no improvement on the test data.

The UNWS-np-15 dataset was obtained from GitHub and used as a simulator in Python. After using the machine learning models, a confusion matrix is used to assess the model's performance. First, the outcomes showed that the Random Forest approach had almost 89% accuracy (PR) and recall (RE). The average Accuracy (AC) of our proposed model is about 89%, which is both good and adequate. The second classification yielded findings indicating that the Precision (PR) and Recall (RE) of the XGBoost method are approximately 90%. The Accuracy (AC) of our suggested model is roughly 90% on average. Comparing our study to earlier research, we found that defect-determination accuracy was significantly higher, at approximately 85% and 79%, respectively. For functional applications, it is essential to provide a deep learning computation alternative that is more user-friendly, faster, and produces better results with less runtime [28].

Developed a fast and efficient detection technique to identify the most recent attacks in real-world scenarios. An effective data pre-processing method combining hybrid feature selection and memory optimization was used to improve the model's universality. The collected characteristics are also used to analyze various weight ranges, hidden neurons, and activation functions to evaluate the attacks using the extreme learning machine (ELM) classifier. For the experiments, they employed the traffic data from CICDDoS-2019. According to the experimental data, the proposed model outperforms previous methods, achieving 99.94% detection accuracy [12].

This study involved a comprehensive analysis and evaluation of fifteen different feature selection (FS) approaches from three primary categories: filter-based, wrapper-based, and embedding, in addition to one ensemble feature selection (EnFS) approach. The experiment demonstrated that the EnFS method beats individual FS and provides a universal optimal feature set for many AI models. Using the recovered feature sets, four DL models, five unsupervised models, and seven ML techniques were trained to evaluate the performance of each FS and EnFS separately. Nevertheless, the findings indicated that optimal performance cannot be achieved with only the optimal feature set [59].

In 2022, [65] proposed a hybrid machine-learning intrusion detection system (IDS). The proposed IDS model outperforms the others on the publicly available NSL-KDD dataset, with fewer data dimensions and feature selection, using a 10-fold cross-validation technique. The hybrid IDS model's performance is validated using a confusion matrix. The parameters of support vector machines (SVMs) are optimized using hybrid Harris Hawks optimization (HHO) and particle swarm optimization (PSO) [66]. The effectiveness of these hybrid algorithms is assessed using performance metrics, including accuracy, precision, sensitivity, specificity, F1 score, and K-nearest neighbors, in comparison with traditional algorithms such as C4.5, K-nearest neighbors, and SVM [55].

This study uses binary and multiclass classification to evaluate the effectiveness of 25 time-based variables for detecting and categorizing 12 types of DDoS attacks. Additionally, they evaluated the performance of one deep learning classifier and eight conventional machine learning classifiers in two distinct scenarios. According to the study, most models identified individual DDoS attacks with approximately 70% accuracy, whereas in the control and time-based tests, they detected DDoS attacks with roughly 99% accuracy. The reduced time-based feature subset, by itself, is advantageous for near-real-time applications that integrate continuous learning, since training on the suggested time-based feature subset has been shown to reduce training time without sacrificing test accuracy [23].

A machine learning-based system for recognizing distributed DOS (DDoS) and DoS threats is offered by Rustam [58] in 2022. A sizable dataset of application-layer network traffic is used for this. To achieve better performance, a novel multi-feature strategy is proposed that integrates features from principal component analysis (PCA) and singular value decomposition (SVD). Several machine learning models are used in extensive trials to determine the viability of the multi-feature method. The efficacy of machine learning models is assessed for each attack class, and the outcomes—including accuracy, recall, F1 score, and other metrics—are discussed in relation to current cutting-edge methodologies. According to experimental data, employing several features improves performance and enables RF to achieve 100% accuracy. Although this study has several limitations, it still shows promising results. The method is first evaluated on a single dataset, and more research is needed to determine how to test it against attacks on additional layers. Secondly, there is no investigation into the effect of dataset size. For deep learning models, larger could mean better outcomes. Third, we have not examined the impact of feature set size, which we plan to do in subsequent research.

3.4 DDoS in the Internet of Things IoTs

Chaudhary et al. [14] presented an attack-detection method for detecting anomalous activity in fog-enabled Internet of Things (IoT) systems. Researchers have previously thoroughly examined the effectiveness of filter-based feature selection algorithms and distinct categories classification algorithms on a prepared dataset containing features specific to Internet of Things networks. For both centralized and fog-enabled IoT network architectures, the classifiers' response times are measured, and established evaluation metrics are used to assess how well the tried-and-true classification technique performs. The experimental results reveal that, in terms of accuracy and latency, the J48 classifier performs better than the other algorithms.

In addition to the NSL-KDD and UNSW-NB15 benchmark datasets, a third dataset—the IoT-Zeek dataset, consisting of Zeek network-based intrusion detection connection logs—was used to validate the proposed methodology. By using deep learning models and ensemble classification to analyze publicly available threat intelligence, both benign and harmful, on Zeek connection logs of IoT devices, they produced the IoT-Zeek dataset. The results demonstrated that the method outperforms the most advanced machine learning models currently available, achieving 97.302% on the IoT-Zeek dataset, 92.904% on the UNSW-NB15 dataset, and 92.092% on the NSL-KDD dataset [15].

Kalakoti [30] This work shows that all machine learning problem formulations applied to the detection of IoT botnet attacks across two datasets achieved detection performance of more than 99% with a constrained number of features. Along with a range of filter and wrapper strategies for feature selection, the research used four main machine-learning techniques to induce the models. They employ filtering techniques, and the selected features are then input into the models.

Utilizing NN, SVM, and SMOA improves our method's ability to mitigate DDoS attacks. When the proposed model was evaluated using modular MVP coding paradigms and the CIC dataset, its accuracy was 99.19%. Our method is well-positioned to achieve even greater efficacy after overcoming the highlighted restrictions, providing a promising future for smart building development [35].

3.5 DDoS detection using Machine Learning on IoT-CIDDS, CIC-DDoS.

The Random Forest (RF) ensemble classifier receives data from Support Vector Machine (SVM), Decision Tree (DT), and Logistic Regression (LR) algorithms. The CICIDS 2017 dataset is used for experiments. The recommended method achieves 98% accuracy, 97% recall, 100% precision, and 98% F-score compared to the individual models. Our model outperforms the individual classifiers in terms of detection rates and accuracy [6].

Maheshwari et al. [39] proposed an ensemble that uses six base classifiers, identified by hyperparameter values two each of SVMs, Random forests, and gradient-boosted machines. The optimal set of weights is found using a novel hybrid metaheuristic optimization method (BHO). Our approach uses a novel dynamic fitness function to eliminate false negatives. The ensemble system was trained with the CIC-DDoS2019 dataset. To assess the performance of the proposed model, a Mininet testbed with a POX controller is used, loaded with multiple datasets. Our model surpasses all existing techniques with low variance and high classification accuracies of 99.4163% and 99.3591% on the CIC-DDoS2019 and CAIDA-2007 datasets, respectively.

Tikhe and Pushpinder [66] proposed an IDS called the wrapper-feature-selection-based hybrid deep learning model (WF-HDL). The DDoS detection model consists of three steps: feature selection, detection, and pre-processing. In this work, two VNF models (a virtual firewall and a virtual router) are trained on the CIC-DDoS2019 dataset. The proposed attack detection model performs well, with an accuracy of 99.69%,

precision of 99.03%, recall of 99.07%, F1-score of 99.05%, and an area under the receiver operating characteristic (ROC) curve of 99.85%.

Dey et al [20] presented a hybrid feature selection scheme that combines statistical test-based filter approaches. The investigation's findings support the notion that the recommended strategy works best with the fewest optimized features—just 13 out of 43 features—and the highest accuracy—99.48 percent.

They used the CICIDS2017 benchmark dataset to develop advanced machine learning techniques. Four cutting-edge machine learning techniques for identifying network attacks were compared. Provide a novel approach, CPRF, to enhance network attack detection performance. They created a distinct feature set by utilizing the recommended CPRF technique. The class probabilities from the network assault dataset that the CPRF approach predicts are used to construct applied machine learning approaches. The results of the exhaustive examination demonstrated that the random forest methodology outperformed state-of-the-art methods, achieving an accuracy of 99.9%. Each applied technique was evaluated for efficacy using a k-fold cross-validation strategy, and its performance was optimized via hyperparameter tuning [57].

Sharif and Beitollahi [63] augmented the system's defenses against evolving attack patterns by integrating Gaussian Mixture Models, Random Forest, and a DDoS expert human. Additionally, they highlighted the use of two datasets, CICIDS2017 and CICDDoS2019, for high-quality data curation, and the integration of the Gaussian Mixture to successfully adapt to shifting data distributions over time. Furthermore, they employed genetic algorithms for automated hyperparameter optimization to ensure successful and effective DDoS detection [70]. The RF classifier scores a remarkable 99.9% accuracy, 100% precision, 99.8% recall, and 99.9% F1 score. Quantitative analysis also shows a significant decrease in false alarms to 0.12% (52 out of 45,149 samples). Our approach handles unknown App-DDoS attacks extremely effectively on all datasets.

By combining Multilayer Perceptron and Convolutional Neural Networks, this model enhances the performance of ML-based DDoS-detection systems in SDN environments. The authors proposed using the Shapley Additive exPlanations (SHAP) feature-selection approach and a Bayesian optimizer to optimize our model. They validated the model on the open-source SDN dataset InSDN to demonstrate that our methodology is also applicable in SDN environments. They used the CICDDoS-2019 dataset to test their method. These results demonstrate that our proposed DDoS-detection model performs effectively in SDN environments compared to existing approaches [62].

Liu and Du [37] tested their method on the Bot-IoT botnet detection dataset and showed that it successfully selects 6 features from the initial 40, achieving an F1-score of 99.63% and a detection accuracy of 99.98%.

3.6 The SMOTE technique

In this research, they describe an enhanced cloud IDS that uses the synthetic minority over-sampling technique (SMOTE) to address imbalanced data. Additionally, a range of attack types is recognized and classified using the random forest (RF) model. The suggested approach was validated using the UNSW-NB15 and Kyoto datasets, yielding accuracies of over 98% and 99% in the multi-class classification scenario, respectively [11].

Dasari and Kaluri [18] proposed a hyperparameter-optimization method for network intrusion classification based on hierarchical machine learning. For this investigation, the CICIDS 2017 dataset was used. Initially, the data was preprocessed using min-max scaling and SMOTE. Each of these algorithms has pre-trained hyperparameters to boost its effectiveness. Metrics including F1-score, recall, accuracy, and precision were used to assess the models' performance. Investigation techniques have shown that the LGBM algorithm can classify DDoS attacks with a 99.77% classification accuracy.

Khedr et al. [34] by preventing infection propagation to the ISP level, efficiently identifying DDoS attacks at both high and low rates, and differentiating between attack traffic and flash crowds, the proposed FMDADM architecture secures both local and remote IoT nodes. The FMDADM outperformed most current state-of-the-art methods across 10 evaluation criteria. Based on the testing results, FMDADM met the following benchmarks: 99.79%, 99.43%, 99.77%, 99.79%, 99.95%, 00.21%, 00.91%, 00.23%, and $2.64.

This study proposes a contractive autoencoder-based deep learning model, grounded in the abnormalities identified by [3]. They trained the model to recognize common traffic patterns using a condensed version of the input data. Then, I applied a stochastic threshold technique to detect the attack. Three well-known intrusion detection system datasets were used in the evaluation process: CIC-IDS2017, NSL-KDD, and CIC-DDoS2019. They also contrasted the model's results with those of a basic autoencoder and other deep learning methods to demonstrate its effectiveness. With accuracy values ranging from 93.41% to 97.58%, our results demonstrate that the recommended technique successfully detected intrusions on the CIC-DDoS2019 dataset.

Rani et al [56] employed a range of machine learning techniques and evaluated their performance through extensive simulation. The results demonstrated that both the Slowloris and CICDDoS2019 datasets achieve higher accuracy when using Random Forest. Among all the ML approaches examined, Random Forest yielded the best results, particularly for detecting and classifying DDoS attacks. F1 scores ranged from 92.4% to 97.3%, accuracy from 99.5% to 99.8%, precision from 95.8% to 98.9%, and recall from 86.7% to 98.2%.

3.7 DDoS detection using hybrid model (DL+ML)

Six learning strategies are the subject of the study: the suggested deep learning (DL)-based hybrid model (i.e., CNN+LSTM) and five machine learning (ML) classifiers (ID3, NB, RF, LR, and AdaBoost). The real-time benchmark dataset CICDDoS2019, which supports both binary and multiclass (14 classes) classification, is used to train and evaluate these learning algorithms. In contrast, the proposed DL-based hybrid model outperforms ID3, NB, RF, LR, and AdaBoost [38].

To further stop the attacker from speculating on the feature set, multiple feature sets are created and used to train baseline machine learning classifiers. The results show that the proposed robust model may improve overall accuracy by at least 13.2% and F1-score by more than 110% against adversarial attacks without the need for adversarial training [60].

Mahmood and Al-Shareeda [40] analyzed some Machine Learning and Deep Learning strategies for detecting and analyzing DDoS attacks. To help decide when to employ which of these techniques, this study also compares and evaluates the important differences between ML and DL techniques.

Diaba and Elmusrati [21] combined the Gated Recurrent Unit and Convolutional Neural Network algorithms, which is the recommended method. Simulation was performed using the benchmark cybersecurity dataset from the Canadian Institute of Cybersecurity Intrusion Detection System. The simulation results demonstrated that the suggested approach outperforms current intrusion detection systems, achieving an inclusive accuracy of 99.7%.

The creation of a hybrid decision table and Naive Bayes approach-based intrusion detection model based on signatures is the primary contribution of the study. Additionally, the suggested method's contribution is assessed by comparing it with the existing body of research on the topic. In the preprocessing phase, only five attack features from the most current CICIDS017 dataset were chosen using Multi-Objective Evolutionary Feature Selection (MOEFS) feature selection. Using five CICIDS2017 traits, the proposed method achieves 96.8% accuracy, surpassing the accuracy reported in the literature [55].

Hossain and Islam [25] combined several decision trees in an ensemble-based method to improve classification accuracy, reduce overfitting, and strengthen the model. Principal component analysis, mutual information, and correlation analysis are the three methods used by the feature selection strategy to determine which traits are most helpful for attack detection.

Pandithurai et al. [54] used a feature selection and Bi-LSTM-based honey badger optimization algorithm to forecast DDoS attacks in a cloud setting. The optimal feature is selected by minimizing the mean square error (MSE) of each feature. The proposed model is examined alongside several existing techniques, including ANNs, DNNs, LSTMs, and DBNs.

Nalayini et al. [50], [51] used a completely new Optimized Dual IDS created and implemented. This one is more accurate than the current IDSs. The recently developed HRDPA data preprocessing technique is part of the proposed IDS. It selects the best features to improve the classifier's prediction accuracy using the RFE approach, hyperparameter tuning, and the Repeated Stratified K-Fold process.

The proposed strategy's efficacy with the CICIDS-2017 data set. The findings show that the proposed technique may reduce unnecessary features by about 51%, increase the modified random forest classifier's detection accuracy to 99.9%, and reduce the model computation time by nearly 50% [67].

Mishra et al. [7], [42], [43] Using the proposed model, standard performance metrics demonstrate significant improvements for both binary and multiclass classification. Overfitting has been addressed using L2 regularization, resulting in a 0.005% reduction in generalization error.

Rani & Ioannou et al. [26], [56] The results of the experiment show that the Enhanced Random Forest algorithm, also known as an ensemble random forest, successfully classifies attacks with an astounding accuracy of 99.98%. It will therefore serve as the initial-stage classifier. Moving forward, the OneClass Support Vector Machine (SVM) algorithm will serve as our second-stage identifier due to its high accuracy, reaching 99.7% in detecting abnormalities.

Azimjonov and Kim [9] opined that the BotIoT-2018 dataset's wasted features likely led to the model's lowest accuracy of 11.31% when trained to use all features. The models with the highest accuracy scores, however, were those trained using feature subsets selected by the backward sequential feature selector; these models achieved 95.66% on KDD CUP 2018, 99.48% on BotIoT 2018, and 99.81% on N-BaIoT 2021.

The suggested model has a flexible, scalable architecture that enables careful examination of network traffic data and the identification of complex patterns indicative of DDoS attacks. Performance measures such as loss rates and detection accuracy demonstrate the adaptability of our method across datasets. Our DNN-based model shows remarkable potential to counter modern DDoS attacks, achieving modest loss rates and detection accuracies of 99.98%, 100%, and 99.99% on the InSDN, CICIDS2018, and Kaggle DDoS datasets, respectively [24].

Zhao [69] They illustrated the benefits of the concept through thorough assessments, showing how it can facilitate the portrayal of similarities, guide model classification/unknown assault identification, optimize defense measures, and expedite filtering reactions. For example, the results demonstrate that, because of its uniform behavioral representation, only 15 distinct attack types can be defended with a single rule. The suggestion is to create a DDoS family to address and overcome these problems. Characterizing traffic patterns, creating attack fingerprints, and executing cross-executed family partitions using community detection are all included in the specified technical roadmap.

Snehi et al. [64] proposed a five-stage defense plan to mitigate Internet-of-Things-based DDoS attacks. The article's solution, based on the behavioral traits of actual devices, is a multi-stage Stack Ensembled structure that provides a universal defense mechanism. The hand-crafted feature selection technique reduced the features by 80%, achieving 99.99% accuracy on benchmark datasets, 98.84% in the simulation environment, and 1.52% collateral damage. The experiment found positive values for significant performance indicators that are often disregarded by researchers. The report also includes a thorough examination of the IoT-DDoS mitigation system's performance data.

To detect DDoS attacks in SDN systems, this study proposes an effective method (BRS + CNN) that combines Balanced Random Sampling (BRS) with Convolutional Neural Networks (CNNs). To mitigate these dangers, we have implemented several strategies, including rate limiting, filtering, and iptables rules to block IP address spoofing. With accuracies of over 99.99% for binary classification and 98.64% for multi-classification, the suggested model performed well in both categories. The suggested DDoS detection solution not only identifies the attack but also emails a specified address with comprehensive contextual information. [49]

This study develops a generalized DL framework for IIoT traffic classification through multiple-domain learning. The WUSTL-IIoT-2021, XIIoTID, and Edge-IIoTSet databases are integrated. The recommended structure consists of two phases. Hyperparameters are adjusted by Bayesian optimization. Experiments conducted on single, combined, and multiple datasets demonstrate the system's efficacy. On the combined dataset, the CNN-GRU model achieved 97.68% accuracy, 97.70% recall, 97.67% precision, and 97.68% F1-score in binary classification. Transfer learning, evaluated on Edge-IIoTSet and trained on XIIoTID, increased accuracy from 50.95% to 97.80% and the F1-score from 48.52% to 97.79%. The weighted averages of recall, precision, and F1-score, as well as the multi-classification accuracy, were roughly 98.85% [10].

The data under observation indicates that the GJOADL-IDSNS system is operating at peak efficiency. Meanwhile, the CNN and LSTM algorithms exhibit inefficient performance. In the meantime, results from the CNN-GRU, CNN-LSTM, MM-WMVEDL, and KELM models are marginally closer. The GJOADL-IDSNS approach's effectiveness, however, is supported by greater precision, recall, accuracy, and F-score values of 99.70%, 98.95%, 98.95%, and 98.95%, respectively [4].

This study presented the DeepDefend framework as a novel approach to identifying and preventing DDoS attacks in cloud computing settings. This system is unique because it combines Deep Learning and powerful Machine Learning algorithms, enabling timely and precise identification of potential assaults. More specifically, compared with CNN-DT with all features and with an optimal selection of features derived using evolutionary algorithms, the AutoCNN-DT model has slashed detection times by 72% and 69%, respectively. Every model exhibited exceptional recall, accuracy, precision, and F1-score values, which ranged from 0.9926 to 0.9997 [52].

Coscia et al. [17] this study proposes a novel approach called Anomaly2Sign. It generates Suricata rules autonomously using a Decision Tree (DT)-based generation method. With classification scores ranging from 99.7% to 99.9%, the proposed method outperforms the compared ML classifiers on the BOUN-DoS and BUET-DDoS datasets.

A Python application is used to classify the traffic into one of the classifications. The (SAE-MLP) obtained the highest accuracy score of 99.75% among all the classifiers that were used. Using the same deep learning approaches, the SDN-DDoS dataset is found to achieve significantly higher traffic categorization accuracy than other publicly available datasets. Additionally, experimental help is provided by the assault detection time of 216.39 s. While malicious traffic includes DDoS assault traffic, such as TCP Syn assault, UDP Flood, and ICMP Flood, TCP, UDP, and ICMP traffic are considered normal [2].

Benmohamed et al. [13] proposed a technique that employs an encoder to extract pertinent features from a preprocessed dataset to accurately detect attacks. Many experiments encompassing a variety of

DDoS attack scenarios were run on the benchmark cybersecurity datasets, CICDS2017 and CICDDoS2019. The experimental results demonstrate that the E-SDNN model outperforms state-of-the-art methods. The proposed E-SDNN model achieved an exceptional overall accuracy of 99.94% and 98.86% on CICIDS2017 and CICDOS2019, respectively.

Table 1. Summary of related works.

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
1.	(Ziyad R. Alashhab et al., 2022)	MDPI	A qualitative comparison with the current surveys was also provided.	The purpose of the survey was to help other academics who are developing novel techniques for detecting DDoS attacks using the context of CC by acting as a reference and guide	Was only based on literature and theory	Statistical methods and algorithms are recommended to enable data storytelling.
2.	(Nalayini and Katiravan, 2022)	SSRN, JETIR	To get the experimental results, training and testing are done on the CIC-IDS2017 benchmark dataset. Before preprocessing, cross-validation using the K-Fold method is carried out.	It generated 99.885% accuracy, 99.88% precision, 100% recall, and a 0.05% false alarm rate to quickly detect DDoS attacks.	The Random Forest model is the best in terms of accuracy, precision, recall, and false alarm rate; therefore, it is insufficient	K-fold alone does not solve the data imbalance. Random Forest shouldn't form the basis for the best result; it should be tested across other models.
3.	(Kumari and Mrunalini, 2022)	Springer	Logistic regression and Naive Bayes are applied. The study made use of the CAIDA 2007 Dataset.	It has been noted that machine learning models performed better than the mathematical models by a small margin. The mathematical model has 99.75% accuracy, while the machine learning model has 100% accuracy.	The model that is being given has a restriction in that it was developed using data from a single dataset. As such, a distributed dataset can be analyzed to offer guidance for future improvements.	More than one dataset should be used to improve the veracity of the model.
4.	(Anupama Mishra et al., 2022).	Springer	This study included six distinct types of machine learning algorithms: XGBoost, AdaBoost, Random Forest, Support Vector Machine, Naive Bayes, and Decision Tree.	AdaBoost computes its maximum accuracy of 99.87% in 27.4 seconds based on the results. Naive Bayes beats the Support Vector Machine, which has the lowest accuracy (95.73%) and the longest learning time (229 m26s for training and 0.2 s for prediction), with a 94.15% accuracy rate and the fastest computing time (3.2 s).	The issue of data imbalance was not considered at the pre-processing stage.	The use of hyperparameter tuning on the machine learning algorithms and exploration of more samples in the datasets.
5.	(Sayed et al., 2022)	IEEE	Proposed a Deep Learning (DL) technique based on Long Short Term Memory (LSTM) and an Auto-encoder to tackle the problem of DDoS attacks in SDNs.	The CICIDS2018 and InSDN datasets have recorded accuracy rates of 99.61% and 55.18%, respectively. For the CICIDS2018 and InSDN datasets, the overall accuracy using the RF technique is 88.21% and 36.22%, whereas the overall accuracy using the IG approach is 89.09% and 32.94%.	Overfitting is a major issue that substantially impairs the effectiveness of ML/DL models. During training, the model can function quite well, however it is unable to show any positive trends with the unknown data.	Data balancing is recommended as well as the use of more datasets for testing and training the model.
6.	(Batchu & Seetha, 2022)	Elsevier	The collected characteristics are used to analyze various weight ranges, hidden neurons, and activation functions to evaluate the attacks using the extreme learning machine (ELM) classifier.	Used the traffic data from CICDDoS-2019 for our experiments. The experimental results show that, with a 99.94% detection accuracy, the proposed model outperforms earlier approaches	Only one type of dataset was used in training and testing.	More testing using various and other types of datasets is recommended as well as feature selection.
7.	(Saha et al., 2022)	MDPI	Thoroughly examined and assessed the effectiveness of fifteen distinct feature selection (FS) techniques from three main categories.	Based on the outcome study, it was concluded that the EnFS technique outperforms the FS method in all three categories (ML, DL, and UL).	Nevertheless, our findings also indicated that the best performance cannot be achieved with just the optimal feature set	Adversarial machine learning (AML) is a study topic that might be incorporated in the future to keep the system safe overall and test

.

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
						using more datasets.
8.	(Halladay, et al., 2022).	IEEE	Investigated the efficacy of 25 time-based variables in identifying and classifying 12 distinct types of DDoS attacks using binary and multiclass classification.	The study found that while most models provided an accuracy of about 70% in identifying specific forms of DDoS attacks, they recognized DDoS attacks with an accuracy of roughly 99% in both the control and time-based tests.	Only examined the performance of one deep learning classifier and eight traditional machine learning classifiers.	More exploration of datasets and machine learning algorithms is recommended
9.	(Rustam, et al., 2022)	MDPI	Proposed a novel multi-feature approach that combines the properties of principal component analysis and singular value decomposition.	According to experimental data, employing several features improves performance and enables RF to achieve 100% accuracy	This method is first evaluated on a single dataset, and secondly, there is no investigation into the effect of dataset size. Third, they have not examined the impact of feature set size.	More research using more datasets is needed to determine how to test it against attacks on the application layer.
10.	(Zivkovic, et al., 2022)	Springer	The optimal architecture for the XGBoost is adaptively found using the FDO. Using the widely used NSL-KDD benchmark dataset, the proposed method is validated.	The experiment's findings demonstrate how much more accurate the suggested FDO-XGBoost method is than the other approaches; on average, its precision and recall have values of 0.82 and 0.77.	Not applicable to any layers of the OSI model, using only one type of hyperparameter.	More algorithms are suggested, and the use of two or more hyperparameters is applied to the Application Layer.
11.	(Prasad & Chandra, 2022)	Springer	Offered a voting architecture with multiple modes to ward off volumetric DDoS (VMFCVD) attacks.	The dataset that has been greatly dimensionally reduced aids FDM in speeding up detection. In the majority of the instances, an accuracy of 99.9% was retained while reducing the dimension for FDM by more than 97%. VMFCVD operates remarkably effectively when a DDoS assault is launched against the server.	Based on the VMFCVD results, it performs better than previous research. However, the application layer is not used in the implementation and limited use of algorithms and data.	Implementation should be channeled toward the networks using more algorithms and datasets.
12.	(Chaudhary , Gupta, & Singh, 2022)	Springer	For both centralized and fog-enabled IoT network architecture, the classifiers' reaction times were computed, and established evaluation metrics were used to gauge how well the tried-and-true classification technique performed.	The experimental results reveal that, concerning accuracy and latency, the J48 classifier performs better than the other algorithms. They showed an attack detection technique to spot unusual activity in the fog-enabled Internet of Things (IoT).	No direct application was done. Used a few algorithms	An attack-based dataset is recommended to be implemented on multiple algorithms and applied to a layer in the OSI model
13.	(Avci and Koca, 2023)	MDPI Electronics	Slime Mold Optimization technique (SMOA) employing an SVM technique for feature selection and an Artificial Neural Network (ANN) estimator.	When the proposed model was evaluated utilizing modular MVP coding paradigms and the CIC dataset, its accuracy was found to be 99.19%	Multiple layers of defense against different methods of cyberattack	Test using more than one dataset, and balance the dataset to resolve the issue of data imbalance. To be applied on one layer first.
14.	(Al-Shareeda et al., 2023)	SSRN, Bulletin of Electrical Engineering and Informatics	Machine Learning (ML) and Deep Learning (DL) strategies	To decide when to employ which of these techniques, this study also compares and evaluates the important differences between ML and DL techniques.	Concluded that both are sufficient. However, the availability of data and its type would determine if ML or DL will be used.	Machine learning is recommended in texts.
15.	(Kumar et al.., 2023)	Elsevier	A model based on long short-term memory (LSTM) is created to identify DDoS attacks on a portion of network traffic packets.	The proposed LSTM model reached up to 98 percent accuracy using the "CICDDoS2019 dataset" for testing and training. Deep learning outperforms machine learning	One dataset does not suffice to base a conclusion on.	The experiment should be tested on more datasets.
16.	(Liu et al.,	MDPI and	The Random Forest,	The findings demonstrate that the	However, no	Advanced feature

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
	2023)	(NIH) National Library of Medicine	Support Vector Machine, Decision Tree, K-Nearest Neighbor, and XGBoost algorithms are utilized.	suggested approach outperformed the others and was able to recognize and classify DDoS attacks in SDNs, offering a fresh perspective and security solution for SDNs.	integration into security technologies to create comprehensive and complete network security solutions.	engineering techniques and more machine learning to improve performance and accuracy.
17.	(Dey, Gupta, & Sahu, 2023)	Elsevier	Merged a non-dominated sorting genetic algorithm (NSGA-II)-based metaheuristic strategy for feature optimization with statistical test-based filter approaches.	The investigation's findings support the notion that the recommended strategy works best with the fewest optimized features—just 13 out of 43 features—and the highest accuracy—99.48 percent.	Out of 43 features, only 13 were optimized. The term heuristic suggests some approaches are not generally acceptable.	More statistical and mathematical models should be used against multiple matrices. More features are to be optimized as well.
18.	(Raza et al., 2023)	IEEE	Utilizing the benchmark dataset CICIDS2017. After comparing four cutting-edge ML techniques for identifying network assaults, they proposed a novel technique known as CPRF. Each applied technique was evaluated for efficacy using a k-fold strategy, and its performance was optimized through hyperparameter tweaking.	The results of the exhaustive examination demonstrated that the random forest methodology outperformed state-of-the-art methods with a high-performance accuracy of 99.9%.	The model architecture's computational complexity was high. The investigation proved that the GNB approach produced computations that were as little as 0.60 seconds.	Reduce the model's computational complexity as much as possible while addressing the problem of data imbalance.
19.	(Liu & Du, 2023)	MDPI	Compared to previous methods that do not select features, this approach has advantages in terms of detection accuracy and training time.	The experiments performed on the Bot-IoT botnet detection dataset show that this method successfully selects 6 characteristics from the initial 40 features, with an F1-score of 99.63% and a detection accuracy of 99.98%.	It chose only 6 features from the initial 40 features.	However, more features could be chosen using other feature selection methods. More datasets could be used on divers matrices.
20.	(Bakro, et al., 2023)	IEEE	They used the synthetic minority over-sampling method (SMOTE) and a hybrid feature selection strategy that includes three techniques—information gain (IG), chi-square (CS), and particle swarm optimization (PSO)—to address the problem of imbalanced data.	The random forest (RF) model helps identify and classify a wide range of attack types. The suggested approach has been validated using the UNSW-NB15 and Kyoto datasets, yielding accuracies of over 98% and 99% in the multi-class classification scenario, respectively.	A variety of attack types are identified and categorized using the random forest (RF) model. However, one model does not suffice.	More algorithms should be used alongside the combination of more data balancing techniques.
21.	(Aktar & Nur, 2023)	Elsevier	A Deep Learning model based on contractive auto encoders is suggested. They trained the model to recognize the common traffic pattern using the condensed version of the input data. Then applied a stochastic threshold technique to detect the attack.	Three well-known intrusion detection system datasets were used in the evaluation process: CIC-IDS2017, NSL-KDD, and CIC-DDoS2019. The results indicate that, with accuracy levels ranging from 93.41% to 97.58%, the recommended technique successfully detected intrusions on the CIC-DDoS2019 dataset. Furthermore, it achieved accuracy of 96.08% and 92.45%, respectively, utilizing the NSL-KDD and CIC-IDS2017 datasets.	Few deep learning approach was used.	They intend to apply the present approach to additional benchmark datasets. Furthermore, they wish to transform the binary classification issue into a multi-class one.
22.	(Sarıkaya et al.., 2023)	Elsevier	In RAIDS, the reconstruction error of an auto-encoder is used to calculate the prediction value of a classifier.	The results show that the proposed robust model may improve overall accuracy by at least 13.2% and F1-score by more than 110% against adversarial attacks without the	Model's ability to avoid adversarial training—a difficult undertaking given the assault area.	Intend to test the suggested paradigm in a practical software-defined network setting later on.

.

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
				need for adversarial training.		
23.	(Diaba & Elmusrati, 2023)	Elsevier	The suggested approach is a mix of the Gated Recurrent Unit and Convolutional Neural Network algorithms. The Canadian Institute of Cybersecurity Intrusion Detection System's benchmark cyber security dataset is used for simulations.	The simulation results demonstrate that the suggested approach outperforms the current intrusion detection systems, with an overall accuracy rate of 99.7%.	Only four models was used in the experiment.	Therefore, the use of more machine learning algorithms and more datasets is hereby recommended.
24.	(Ahmad et al.., 2023)	Elsevier	This study published a paradigm for identifying malicious network traffic. The proposed method is applied to the KDD dataset analysis, and specificity, prediction time, training, and accuracy are assessed as a result.	The system uses two well-known classification-based methods to identify malicious network traffic: a Support Vector Machine (SVM) and a deep neural network model, or Convolutional Neural Network (CNN) coupled with a Gated Recurrent Unit (GRU). The latter is adjusted for improved accuracy of 98.45% and 94.84% using the Slime Mold Algorithm (SMA).	The framework used two classification-based techniques for detection.	More classification techniques is hereby recommended to be used on several datasets.
25.	(Dasari & Kaluri, 2024)	IEEE	The CICIDS 2017 standard dataset was used. SMOTE and min-max scaling were used to preprocess the data. The hierarchical machine learning algorithms, XGboost, LGBM, CatBoost, Random Forest (RF), and Decision Tree (DT), were fed feature selection input using the LASSO technique	The models' performance was evaluated using metrics such as F1-score, recall, precision, and accuracy. Research methods have demonstrated that the LGBM algorithm has a demonstrated 99.77% classification accuracy when it comes to DDoS attack classification	Few algorithms were used on only one dataset.	Aiming to increase forecast accuracy by including more ensemble machine learning and deep learning models as well as different optimization techniques.
26.	(Turukmane & Devendiran, 2024)	Elsevier	In the pre-processing phase, the un-normalized data was modified using Min-Max normalization. Following pre-processing, the Advanced Synthetic Minority Oversampling Technique (ASmoT) is used to lessen the issue of class imbalance	According to the performance metrics, the suggested system used the CSE-CIC-IDS 2018 dataset to achieve 99.89% accuracy, and it also used the UNSW-NB15 dataset to obtain 97.535% accuracy	Only one sampling technique was used.	More sampling techniques alongside more algorithms are recommended
27.	(Hossain & Islam, 2024)	Elsevier	Several decision trees are combined in the ensemble-based method to improve classification accuracy, decrease overfitting, and strengthen the model.	With nearly 100% accuracy, 100% true positive rate, and 0% error rate, the suggested method is a promising one for DDoS attack detection.	Failed to use data balancing technique	An algorithm to automatically detect the highest accuracy is proposed along with the use of data balancing techniques and more machine learning techniques.
28.	(Pandithurai et al., 2024)	Elsevier	The Bi-directional Long short-term Memory (Bi-LSTM) classifier is then fed with the best features to anticipate DDoS attacks. Additionally, the suggested model is	The Bi-LSTM model attained 97% accuracy, 95% sensitivity, 90% specificity, 3% error, 94% precision, and so on when the performance was assessed using the current methodology.	However, the issue of data imbalance is not addressed.	Hyper-parameter tuning, more datasets, data imbalance, and a distinct model to identify the algorithm with the highest accuracy

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
			investigated concerning some of the current methods, such as ANN, DNN, LSTM, and DBN			are recommended.
29.	(Tripathi et al., 2024)	IEEE	The work provided a novel way to optimize the feature selection process in machine learning algorithms. To maximize feature selection and reduction, the suggested strategy focused on features that have a significant influence on the target variable	The effectiveness of the suggested strategy is evaluated using the CICIDS-2017 data set. The ability of the suggested strategy to minimize superfluous features by almost 51% and boost the modified random forest classifier's detection accuracy to 99.9% while cutting down on model computation time by nearly 50% is demonstrated by the results	A 7:3 ratio was taken into consideration for training and testing sets, with a total of 2, 271, 320 recordings for benign traffic and 556, 556 records for attacks taken into consideration. Hence, data was not balanced.	More data balancing techniques is therefore recommended.
30.	(Ioannou, et al., 2024)	Elsevier	The approaches below will be compared for our categorization problem: SVM-support vector machines; KNN-K-Nearest Neighbours; Decisions Tree; Random Forest.; Naive Bayes	The results of the experiment show that the Enhanced Random Forest algorithm, also known as ensemble random forest, successfully classifies attacks with an astounding accuracy rate of 99.98%. It will therefore be our initial stage classifier. Moving forward, the One Class Support Vector Machine (SVM) algorithm will be our second-stage identifier due to its high degree of accuracy, which reaches 99.7% in detecting abnormalities.	All of the strategies under examination perform admirably in terms of accuracy, precision, recall metrics, and the F1 score, according to the results of our simulation. However, few algorithms were used on these matrices.	These topics of future research seek to greatly increase the research's impact and breadth.
31.	(Zhao, et al., 2024)	Elsevier	The suggestion is to create a DDoS family to address and overcome these problems. Characterizing traffic patterns, creating attack fingerprints, and executing cross-executed family partitions using community detection are all included in the specified technical roadmap.	They illustrated the benefits of the concept through thorough assessments, showing how it can facilitate the portrayal of similarities, guide model classification/unknown assault identification, optimize defense measures, and expedite filtering reactions.	For example, the results demonstrate that because of its uniform behavioral representation, only 15 different types of attacks may be defended with just one rule.	A hierarchical approach to provide granularity that is customized for a wide range of applications.
32.	(Snehi, Bhandari, & Verma, 2024)	Elsevier	Provided a five-stage defense strategy. The research community is presented with an aggregated dataset created from the InSDN, BoT-IoT, and UNSW-Sydney datasets as well as a simulated dataset for IoT-DDoS in this article.	With a high accuracy of 99.99% with benchmark datasets, 98.84% accuracy in the simulation environment, and 1.52% collateral damage, the hand-crafted feature selection technique reduced the features by 80%.	The experiment finds positive values for important performance metrics that are frequently overlooked by researchers. However, The experimental investigation was restricted to flooding assaults using TCP, UDP, and HTTP.	suggests expanding the defense solution to include cutting-edge IoT protocols like MQTT, unsupervised deep learning, and blockchain.
33.	(Aljehane, et al., 2024)	Elsevier	The Deep Learning Assisted Intrusion Detection System for Network Security (GJOADL-IDSNS) approach is combined with a new Golden Jackal Optimisation Algorithm.	The GJOADL-IDSNS approach's effectiveness, however, is supported by greater accuracy, precision, recall, and Fscore values of 99.70%, 98.95%, 98.95%, and 98.95%, respectively.	Inefficient performance is presented for the CNN and LSTM algorithms.	Recommended the application of outcomes in specific real-world situations
34.	(Ouhssini et al., 2024)	Elsevier	Presented the DeepDefend framework as a novel approach to identifying and preventing DDoS	More specifically, as compared to CNN-DT with all features and CNN-DT with an optimal selection of features derived using evolutionary algorithms,	No application to real-life scenarios.	To assess the DeepDefend framework's flexibility and durability in a

.

S/No	Author And YEAR	Source	Approach	Strength	Weakness	Recommendation
			attacks in cloud computing settings. This system is unique because it combines deep learning and powerful machine learning algorithms.	the AutoCNN-DT model has slashed detection times by 72% and 69%, respectively. Every model exhibited exceptional precision, recall, accuracy, and F1-score values, which ranged from 0.9926 to 0.9997.		variety of harsh and uncertain scenarios, we intend to deploy it in actual cloud environments shortly. Additionally, we want to develop a specific dataset for time series analysis in cloud systems, which will be very helpful in improving our methodology and advancing more general research in the area.
35.	(Ahuja, Mukhopadhyay, & Singal, 2024)	Springer	The objective was to categorize the network traffic into two categories: harmful and normal, based on features found in the dataset. While TCP, UDP, and ICMP traffic are regarded as normal, malicious traffic includes DDoS assault traffic such as TCP Syn assault, UDP Flood, and ICMP Flood.	Traffic is categorized into one of the classes using a Python programme. Out of all the classifiers that were utilized, (SAE-MLP) achieved the highest accuracy score of 99.75%. When the SDN-DDoS dataset is compared to other publicly available datasets using the same deep learning techniques, it is discovered that the SDN-DDoS dataset has substantially higher traffic categorization accuracy and the assault detection time was 216.39s.	Only five deep-learning algorithms were used on one dataset.	Therefore, more experiments using various types of machine learning algorithms are recommended. As well as more samples of datasets.
36.	(Benmohamed et al., 2024)	Springer	To precisely detect attacks, the suggested solution uses an encoder to extract relevant features from a preprocessed dataset.	On benchmark cybersecurity datasets, CICDS2017 and CICDDoS2019, numerous tests covering a range of DDoS attack scenarios were carried out. The experimental outcomes show how much better the E-SDNN model is than the most advanced techniques.	Two classes of datasets were used on S-DNN, as well as only the SMOTE technique to oversample the minority class.	The combination of more Data balancing techniques like Tomeklink+SMOTE+Bagging is thereby recommended. As well as the use of more machine learning techniques on several types of datasets.

Table 2. Limitations of related works.

S/No		
1.	Dataset dependency and imbalance	The studies that used the Canadian Institute of Cybersecurity (CIC) dataset [2017,2018,2019,2021,2021,2022] produced closely knitted results to the dataset, thereby reducing generalizability.
2.	Static vs Dynamic Feature Selection	Feature selection, extraction, and engineering [2018,2021,2022,2023] in the earlier year seemed rigid. Flexibility was later adapted by adding new features, however, there is a need to combine more features to combat evolving attacks which is not discussed.
3.	Hyper-parameter tuning	Studies in 2021 and 2023 show that Hyperparameter tuning is evolving rapidly, and there is room for improvement in automated hyperparameter optimization.
4.	Specific IDS / DDoS	The use of SNORT in SDN in 2021 and 2022 is dependent on the SNORT-IDS and may not be easily adaptable in other IDS systems or bespoke solutions.
5.	Evaluation Metrics	Although, generally many studies captured high accuracy rates, few discussed other metrics like Precision, recall, AUC, and F1-score in-depth leaving an incomplete evaluation report.
6.	Computational Resources Availability	Studies in 2022 and 2023 present intense computational algorithms making them unsuitable for resource-constrained scenarios. Also, the dataset SDN-DDoS, IoT23 resource might not be optimized enough for large-scale SDN deployments.
7.	Complexities	Security models used in 2021, 2022, and 2023 like D-CAD, SD-VANET's, SDN, 6LoWPAN stack, etc. utilized advanced algorithms like WF-HDL, POA, and DAE-

		CGRU added layers of complexity that could make real-time implementation difficult.
8.	Simulation-Based	Most studies in 2020-2022 relied on simulations rather than real-life networks which might not represent real-world attack scenarios accurately.
9.	Mitigation	Real-time mitigation in [2023] has not been fully harnessed. Fewer works have been done here.
10.	Unique Predictive model	[27]

From 2020 to 2022, the efforts expanded to address data imbalance, feature selection, and mitigation, with an added focus on hyper-parameter control. Techniques such as LSTM-fuzzy deep belief networks, deep neural networks, Machine Learning, IDHCS, D-CAD, and multi-layer perception models were implemented for intrusion detection, often using machine learning and deep learning algorithms with k-fold cross-validation. These studies targeted traffic flow and networks, using datasets such as CIC-DDoS 2019, CIC-DDoS 2017, UNSW-NB15, CAIDA2007, CIC-IDS-2017, and CICIDS 2018.

In 2023 and 2024, the focus shifted to more advanced techniques for handling data imbalance and feature engineering, as well as real-time mitigation and automated hyperparameter optimization. Innovative approaches such as ASmoT (Advanced Synthetic Minority Oversampling Technique), multi-variant models like CON and DISR, LSTM and Bi-LSTM, and algorithms like Pelican optimization (POA), Wrapper feature selection-based hybrid Deep Learning Model (WF-HDL), Deep auto-encoder-convolutional gated recurrent unit (DAE-CGRU), Stack auto-encoder multi-layer Perceptron (SAE-MLP), and Deep Learning based hybrid model were employed. These methods were applied across various domains, including networks, routers, and software applications, using datasets such as UNSW-NB15, CSE-CIC-IDS 2018, CIC-DDoS 2019, and a wide range of IoT-related datasets, such as IoT-CIDDS and NBaIoT-2021.

Despite these advances, several limitations persisted. First, heavy reliance on the CIC datasets reduced the generalizability of results across different environments. Second, early feature selection approaches were rigid, and although flexibility was later introduced, further advancements are needed to keep up with evolving attack scenarios. Third, hyperparameter tuning has made progress, but there is still room for improvement in automation. Fourth, the use of specific intrusion detection systems, such as SNORT, in SDN limited its adaptability to other systems. Additionally, many studies focused on accuracy without thoroughly discussing other evaluation metrics such as precision, recall, and F1-score. Computational resource demands were another challenge, with resource-constrained environments struggling to handle the intensive algorithms required by certain models. Moreover, the reliance on simulation-based studies reduced the accuracy of real-world applications, and real-time mitigation remains underexplored. Lastly, while many studies used a combination of machine learning and deep learning, hybrid models combining ML-DT and DL-CNN appeared only in a 2024 study, as shown in Table 3 below.

Table 3. Summary of the advancement between the techniques, area of detection, and dataset.

Year	Area of improvement	Techniques (methods)	Area of detection	Dataset update
2020 – 2022	Data –imbalance Feature selection Mitigation Hyper-parameter Control	LSTM-fuzzy, Deep Belief Network (DBN), D-CAD, Multi-layer perception, Deep Neural Network (DNN), Intrusion Detection Hyperparameter Control System (IDHCS), Machine Learning, Deep Learning, k-fold cross validation, …	Traffic flow, Networks	– CICDDoS 2019 – CICDDoS 2017 – UNSW-NB15 – CAIDA2007 – CIC-IDS-2017 – CICIDS 2018
2023 – 2024	>Data –imbalance >Feature Selection & Engineering >Mitigation in Real-time >Hyper-parameter Control >Hyper-parameter Adjustment >Automated Hyper-parameter optimization (tuning)	ASmoT (Advanced Synthetic Minority Oversampling Technique) -M-Multi SVM -Machine Learning -Binary Grey Wolf Optimization -Multi-variant approaches like CON and DISR -Wrapper feature selection-based hybrid deep learning model (WF-HDL) -Pelican optimization algorithm (POA) -Deep auto-Encoder-Convolutional Gated Recurrent Unit (DAE-CGRU) -Deep Learning-Based Hybrid Model -Stack auto-encoder multi-layer Perceptron (SAE-MLP).	Networks, Routers, Traffic flow, Software (Apps).	– UNSW-NB15 – CSE-CIC-IDS 2018 – CICDDoS 2019 – SDN-DDoS – IoT-CIDDS – CIC-IDS 2017 – CIC-ToN-IoT – BoT-IoT 2018 – Kyoto – NSL-KDD – IoT23 – NBaIoT-2021 – KDD CUP 2018 – WUSTL-IIoT-2021 – XIIoTID – Edge-IIoT SET – BOUN-DoS – BUET-DDoS

.

4. SUMMARY

4.1 Model architecture based on the review summary

After analyzing the results of all of the existing machine learning models, a hybrid Deep Learning (DL) and Machine Learning (ML) ensemble is proposed:

- CNN + LSTM: CNN layers extract spatial features while LSTM captures temporal dependencies.
- Random Forest (RF) and XGBoost: Serve as baseline classifiers for comparison.

Hybrid Model Equations:

$$F = CNN(x), h_t = LSTM(F_t, h_{t-1}), y^{\wedge} = Softmax(h_T)$$

4.2 Summary of review of relevant papers

This section contains a table highlighting the main ideas from publications that satisfied our most recent inclusion requirements. The 75 publications identified during the search are shown in Table 1. The distribution of these publications across the seven databases examined is shown in Figure 3 and Table 4. As stated in our PRISMA (Figure 2), the referenced complete papers were all indexed in Google Scholar, including the removed duplicates. Elsevier has the most publications on the subject, based on the statistics.

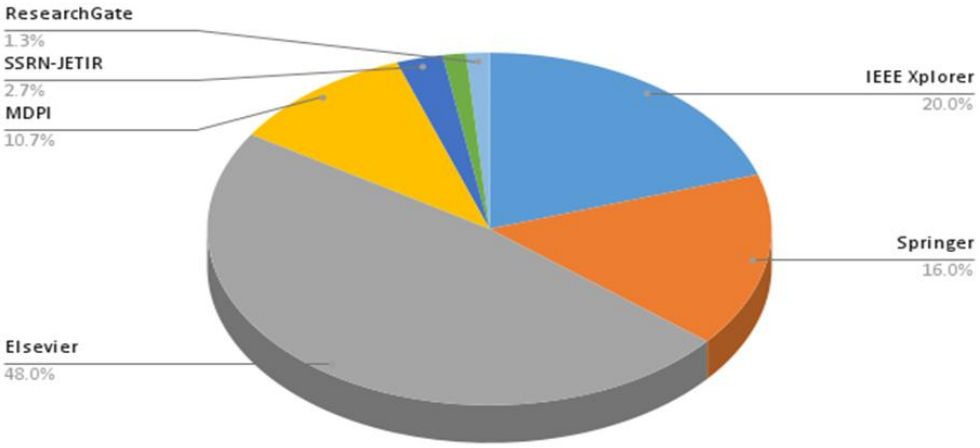

Figure 3. Distribution of included papers by count percentage (%).

Table 4. Distribution of the study's publications based on databases checked following screening.

S/N	Database	URL	Count	% Count
1	IEEE Xplorer	https://ieeexplore.ieee.org	15	20%
2	Springer	https://link.springer.com	12	16%
3	Elsevier (ScienceDirect)	https://sciencedirect.com	36	48%
4	MDPI	https://mdpi.com	8	10.67%
5	SSRN-JETIR	https://www.ssrn.com	2	2.67%
6	Taylor&Francis	https://www.taylorandfrancis.com	1	1.33%
7	WileyOnline Library	https://www.authorservices.wiley.com	1	1.33%

The distribution of the included publications by research year is shown in Figure 4 below. It shows the databases consulted, the number of papers, and the percentage in the overall score for all searches. This study's primary focus was on recent developments in the application of ML/DL to DDoS attack detection across the publications.

4.3 Scientific Contribution of the summary:

The methodology provides a robust, theoretically grounded, and empirically validated approach to DDoS detection. Key contributions include:

i. A hybrid ML/DL architecture capable of processing multi-dimensional, imbalanced data efficiently.
ii. Demonstration of generalizability across diverse datasets.
iii. Integration of interpretable AI via SHAP for model transparency.
iv. Statistically significant improvements in detection accuracy, recall, and F1-score.

This method advances intelligent, scalable, and interpretable intrusion detection systems suitable for real-time cybersecurity applications.

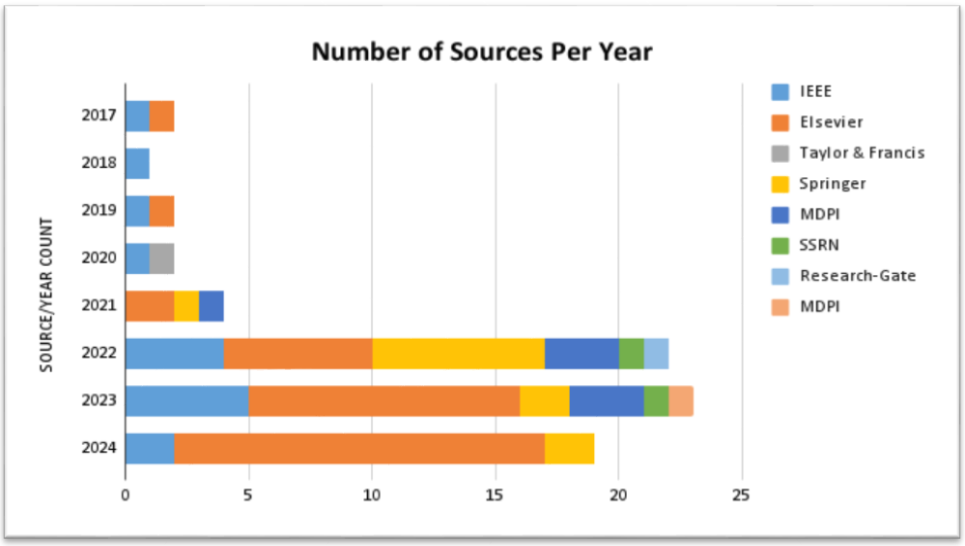

Figure 4. Distribution by year of the included papers.

Figure 5 shows the theme and keyword word clouds of the reviewed articles generated by their word frequencies. This suggests that our research focuses on using feature selection, hyperparameter tuning, and a predictive model to target Distributed Denial of Service (DDoS) attacks at the Application Layer of the OSI Model.

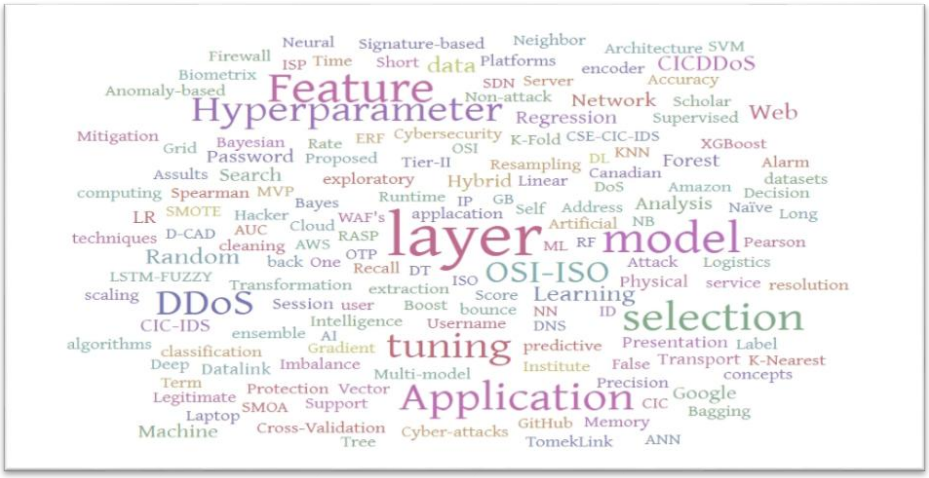

Figure 5. Word cloud of the topic and keywords of reviewed papers.

4.4 Conclusion and Future Work

The most recent advancements in machine learning and deep learning applications for DDoS attack detection and analysis were summarized in the publications listed in the above Table 1 & Table 2. This included details on the methods employed, the datasets used, the features extracted, the hyperparameter adjustments, and the performance metrics reported. Additionally, the authors have investigated how feature selection and hyperparameter tuning methods can be applied to current data to improve the accuracy of automatic DDoS attack detection systems. Machine learning algorithms have demonstrated strong learning and feature-extraction capabilities from datasets, as well as an astounding degree of accuracy in identifying the many forms of DDoS attacks.

DDoS attacks are becoming more varied in form and volume to exhaust the target's resources across the entire network. Therefore, a hybrid model detection system that accurately classifies attacks should be devised to limit the catastrophic impact. To avoid bias toward the conventional machine learning model, we aim to examine the models used between 2022 and 2024 and propose hybridizing machine learning, deep learning, and ensemble learning to predict DDoS attacks. Following the selection of pertinent features using the hybrid feature selection method [6], [7], [14], [15], [20], the implementation stage is used to enhance the model's performance and optimize resource utilization. Ultimately, the ideal hyperparameters and characteristics will be used with a variety of supervised and unsupervised learning techniques to

differentiate DDoS attacks from regular traffic. After that, each of these operation sequences will undergo testing and training across four distinct datasets. The suggested methodology, as shown in Figure 6 below, can therefore be applied as a predictive model to efficiently detect DDoS attacks on any network layer. Furthermore, we plan to apply a balancer on the IDS [41], we would be using up to several matrices on algorithms of classification for the analysis of retrieved datasets using ensemble hybrid model [8], [10], [18], [43], [47], [53], [72] and feature selection [25], [30], [37], [47], [54], [55], [59], [61], [66], [67], [68], to select the model with the highest accuracy. In the future, we will publish the results of further findings.

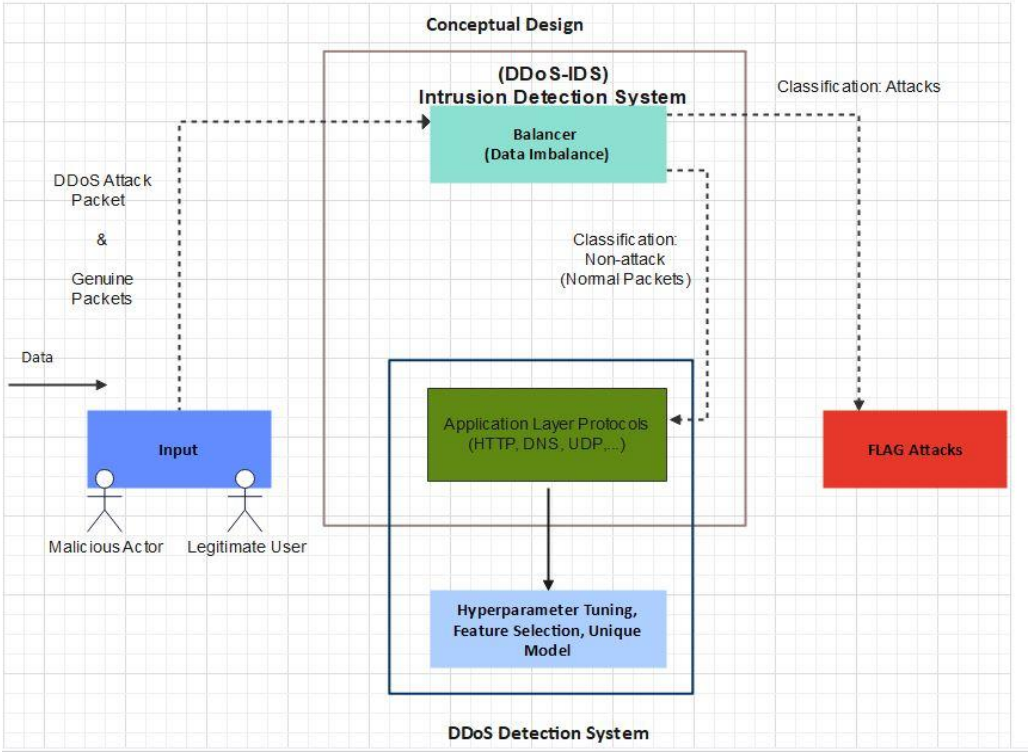

Figure 6. Conceptual Flow Design

List of Abbreviations:

Distributed Denial of Service (DDoS); Internet Protocol (IP) Spoofing; Attacks and Mitigation (A&M); Denial of Service(DoS); Hyperparameter Tuning(HPT); Artificial Intelligence(AI); Gated Recurrent Unit (GRU); Stacked Auto-Encoder Multi-layer Perceptron (SAE-MLP); SHapley Additive exPlanations (SHAP); Intrusion detection hyperparameter control system (IDHCS); neural network (NN); particle swarm optimization (PSO) techniques; principal component analysis (PCA); singular value decomposition (SVD) techniques; multimode voting architecture to defend against volumetric DDoS (VMFCVD); Fast detection mode (FDM); defensive fast detection mode (DFDM); high accuracy mode (HAM); Correlation Feature Selection (CFS); (SVM)Support Vector Machine; deep neural network (DNN); proximal policy optimization (PPO); composite multilayer perceptron (CMP); Linear Discriminant Analysis(LDA); Long Short-Term Memory (LSTM); ensemble feature selection (EnFS) method; hybrid Harris Hawks optimization (HHO); Quadratic Discriminant Analysis(QDA); Bayesian Network(BN); Receiver Operating Characteristic(ROC); Naïve Bayes(NB); True Positive(TP); True Negative(TN); Area Under the ROC Curve(AUC); Multi-Objective Evolutionary Feature Selection (MOEFS); deep belief network (DBN); Supervised Learning(SL); Unsupervised Learning(UL); Primary Data Set(PDS); Secondary Data Set(SDS); Positive Predictive Value(PPV); Negative Predictive Value(NPV); Machine Learning(ML); Deep Learning(DL); Convolutional Neural Network(CNN); Feature Selection(FS); Random Forest(RF); Information Gain (IG); Minimum-Redundancy-Maximum-Relevancy (mRMR); Artificial Neural Network (ANN); Area Under the Receiver Operating Characteristic Curve (AUC-ROC); Rectified Linear Unit(ReLU); Densely-Connected Convolutional Networks(Densenet); Bi-Directional Convolutional Long Short-Term Memory (CLSTM); Logistic Regression(LR); Deep Learning System(DLS); Kappa Coefficient(k); K- Nearest Neighbors(KNN); Decision Tree(DT); Multi-Layer Perceptron Classifier (MLPC); (MLP) Multilayer Perceptron; (ADA) Adaptive Boosting; Precison-(P); Gradient-Boosted Decision Trees(GBDT); Matthew's Correlation Coefficient(MCC); Computer-Aided Design(CAD); Multi-Modal Feature Autoencoder Attention Net(RMANet); Synthetic Minority

Oversampling Technique-SMOTE; Canandian Institute for cybersecurity DDoS – CICDDoS2017; Pearson's Correlation Coefficient (PCC); Mutual Information (MI); Software-Defined Networking (SDN); Physically Unclonable Functions (PUFs); Slime Mold Algorithms (SMA); Deep Convolutional Generative Adversorial Networks(DCGAN); Grey Wolf Optimization (GWO); Whale GWO (WGWO);

REFERENCES

[1] Ahmad, Z. Wan, and A. Ahmad, "A big data analytics for DDoS attack detection using optimized ensemble framework in Internet of Things," *Internet of Things*, vol. 23, no. 100825, Oct. 2023. https://doi.org/10.1016/j.iot.2023.100825

[2] N. Ahuja, D. Mukhopadhyay, and G. Singal, "DDoS attack traffic classification in SDN using deep learning," *Personal and Ubiquitous Computing*, vol. 28, pp. 417–429, 2024. https://doi.org/10.1007/s00779-023-01785-2.

[3] S. Aktar and A. Y. Nur, "Towards DDoS attack detection using a deep learning approach," *Computers & Security*, vol. 129, no. 103251, 2023. https://doi.org/10.1016/j.cose.2023.103251.

[4] M. Aljebreen, H. A. Mengash, M. A. Arasi, S. S. Aljameel, A. S. Salama, and M. A. Hamza, "Enhancing DDoS attack detection using Snake Optimizer with ensemble learning on Internet of Things environment," *IEEE Access*, vol. 11, pp. 104745–104753, 2023. https://doi.org/10.1109/ACCESS.2023.3318316.

[5] N. O. Aljehane, H. A. Mengash, M. M. Eltahir, F. A. Alotaibi, S. S. Aljameel, A. Yafoz, R. Alsini, and M. Assiri, "Golden jackal optimization algorithm with deep learning assisted intrusion detection system for network security," *Alexandria Engineering Journal*, vol. 86, pp. 415–424, 2024. https://doi.org/10.1016/j.aej.2023.11.078.

[6] Analytics Vidhya, "10 techniques to solve imbalanced classes in machine learning," 26, 2023.

[7] B. Anbarasu and I. S. Thaseen, "Anomaly detection using feature selection and ensemble of machine learning models," *Computational Methods and Data Engineering, Lecture Notes on Data Engineering and Communications Technologies*, vol. 139, Singapore: Springer, pp. 215–229, 2023. https://doi.org/10.1007/978-981-19-3015-7_16.

[8] [8] A. Mishra, N. Gupta, and B. B. Gupta, "Defensive mechanism against DDoS attack based on feature selection and multi-classifier algorithms," *Telecommunication Systems*, vol. 82, no. 2, pp. 229–244, 2023. https://doi.org/10.1007/s11235-022-00981-4

[9] AWS, "What is hyperparameter tuning?," Dec. 1, 2023. Accessed: May 12, 2026.

[10] J. Azimjonov and T. Kim, "Designing accurate, lightweight intrusion detection systems for IoT networks using fine-tuned linear SVM and feature selectors," Computers & Security, vol. 137, no. 103598, Feb. 2024. https://doi.org/10.1016/j.cose.2023.103598

[11] B. Babayigit and M. Abubaker, "Towards a generalized hybrid deep learning model with optimized hyperparameters for malicious traffic detection in the Industrial Internet of Things," *Engineering Applications of Artificial Intelligence*, vol. 128, no. 107515, Feb. 2024. https://doi.org/10.1016/j.engappai.2023.107515

[12] M. Bakro, R. R. Kumar, A. Alabrah, Z. Ashraf, M. N. Ahmed, M. Shameem, and A. Abdelsalam, "An improved design for a cloud intrusion detection system using a hybrid feature selection approach with ML classifier," *IEEE Access*, vol. 11, pp. 64228–64247, 2023. https://doi.org/10.1109/ACCESS.2023.3289405

[13] R. K. Batchu and H. Seetha, "On improving the performance of the DDoS attack detection system," *Microprocessors and Microsystems*, vol. 93, no. 104571, 2022. https://doi.org/10.1016/j.micpro.2022.104571

[14] E. Benmohamed, A. Thaljaoui, S. Elkhediri, S. Aladhadh, and M. Alohali, "E-SDNN: Encoder-stacked deep neural networks for DDoS attack detection," *Neural Computing and Applications*, vol. 36, no. 18, pp. 10431–10443, 2024. https://doi.org/10.1007/s00521-024-09622-0

[15] P. Chaudhary, B. Gupta, and A. K. Singh, "Implementing an attack detection system using filter-based feature selection methods for fog-enabled IoT networks," *Telecommunication Systems*, vol. 81, no. 1, pp. 23–39, 2022. https://doi.org/10.1007/s11235-022-00927-w

[16] A. Chohra, P. Shirani, E. B. Karbab, and M. Debbabi, "Chameleon: Optimized feature selection using particle swarm optimization and ensemble methods for network anomaly detection," *Computers & Security*, vol. 117, no. 102684, 2022. https://doi.org/10.1016/j.cose.2022.102684

[17] Cloudflare, "What is a DDoS attack?," Jan. 9, 2024. Accessed: May 12, 2026.

[18] A. Coscia, V. Dentamaro, S. Galantucci, A. Maci, and G. Pirlo, "Automatic decision tree-based NIDPS ruleset generation for DoS/DDoS attacks," *Journal of Information Security and Applications*, vol. 82, no. 103736, 2024. https://doi.org/10.1016/j.jisa.2024.103736

[19] S. Dasari and R. Kaluri, "An effective classification of DDoS attacks in a distributed network by adopting hierarchical machine learning and hyperparameters optimization techniques," *IEEE Access*, vol. 12, pp. 10834–10845, 2024. https://doi.org/10.1109/ACCESS.2024.3352281

[20] 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. https://doi.org/10.1016/j.procs.2023.01.217

[21] A. K. Dey, G. P. Gupta, and S. P. Sahu, "Hybrid meta-heuristic based feature selection mechanism for cyber-attack detection in IoT-enabled networks," *Procedia Computer Science*, vol. 218, pp. 318–327, 2023. https://doi.org/10.1016/j.procs.2023.01.014

[22] S. Y. Diaba and M. Elmusrati, "Proposed algorithm for smart grid DDoS detection based on deep learning," *Neural Networks*, vol. 159, pp. 175–184, 2023. https://doi.org/10.1016/j.neunet.2022.12.011

[23] N. Farhana, A. Firdaus, M. F. Darmawan, and M. F. Ab Razak, "Evaluation of Boruta algorithm in DDoS detection," *Egyptian Informatics Journal*, vol. 24, no. 1, pp. 27–42, 2023. https://doi.org/10.1016/j.eij.2022.10.005

[24] J. Halladay, D. Cullen, N. Briner, J. Warren, K. Fye, R. Basnet, J. Bergen, and T. Doleck, "Detection and characterization of DDoS attacks using time-based features," *IEEE Access*, vol. 10, pp. 49794–49807, 2022. https://doi.org/10.1109/ACCESS.2022.3173319

[25] V. Hnamte, A. A. Najar, H. Nhung-Nguyen, J. Hussain, and S. M. Naik, "DDoS attack detection and mitigation using deep neural network in SDN environment," *Computers & Security*, vol. 138, no. 103661, 2024. https://doi.org/10.1016/j.cose.2023.103661

[26] M. A. Hossain and M. S. Islam, "Enhancing DDoS attack detection with hybrid feature selection and ensemble-based classifier: A promising solution for robust cybersecurity," *Measurement: Sensors*, vol. 32, no. 101037, 2024. https://doi.org/10.1016/j.measen.2024.101037

[27] I. Ioannou, P. Nagaradjane, P. Angin, P. Balasubramanian, K. J. Kavitha, P. Murugan, and V. Vassiliou, "GEMLIDS-MIOT: A green effective machine learning intrusion detection system based on federated learning for medical IoT network security hardening," *Computer Communications*, vol. 218, pp. 209–239, 2024. https://doi.org/10.1016/j.comcom.2024.02.023

[28] İ. Avcı and M. Koca, "Predicting DDoS attacks using machine learning algorithms in building management systems," *Electronics*, vol. 12, no. 19, no. 4142, 2023. https://doi.org/10.3390/electronics12194142

[29] Ismail, M. I. Mohmand, H. Hussain, A. A. Khan, U. Ullah, M. Zakarya, A. Ahmed, M. Raza, I. U. Rahman, and M. Haleem, "A machine learning-based classification and prediction technique for DDoS attacks," *IEEE Access*, vol. 10, pp. 21443–21454, 2022. https://doi.org/10.1109/ACCESS.2022.3152577

[30] M. R. Kadri, A. Abdelli, J. B. Othman, and L. Mokdad, "Survey and classification of DoS and DDoS attack detection and validation approaches for IoT environments," *Internet of Things*, vol. 25, no. 101021, 2024. https://doi.org/10.1016/j.iot.2023.101021

[31] R. Kalakoti, S. Nõmm, and H. Bahsi, "In-depth feature selection for the statistical machine learning-based botnet detection in IoT networks," *IEEE Access*, vol. 10, pp. 94518–94535, 2022. https://doi.org/10.1109/ACCESS.2022.3204001

[32] C. M. Nalayini and J. Katiravan, "Detection of DDoS attack using machine learning algorithms," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 9, no. 7, pp. f223–f232, Jul. 2022.

[33] M. Vishwakarma and N. Kesswani, "DIDS: A deep neural network based real-time intrusion detection system for IoT," *Decision Analytics Journal*, vol. 5, no. 100142, 2022. https://doi.org/10.1016/j.dajour.2022.100142

[34] W. I. Khedr, A. E. Gouda, and E. R. Mohamed, "FMDADM: A multi-layer DDoS attack detection and mitigation framework using machine learning for stateful SDN-based IoT networks," *IEEE Access*, vol. 11, pp. 28934–28954, 2023. https://doi.org/10.1109/ACCESS.2023.3260256

[35] İ. Avcı and M. Koca, "Predicting DDoS attacks using machine learning algorithms in building management systems," *Electronics*, vol. 12, no. 19, no. 4142, 2023. https://doi.org/10.3390/electronics12194142

[36] K. Kumari and M. Mrunalini, "Detecting Denial of Service attacks using machine learning algorithms," *Journal of Big Data*, vol. 9, no. 56, 2022. https://doi.org/10.1186/s40537-022-00616-0

[37] X. Liu and Y. Du, "Towards effective feature selection for IoT botnet attack detection using a genetic algorithm," *Electronics*, vol. 12, no. 5, no. 1260, 2023. https://doi.org/10.3390/electronics12051260

[38] S. S. Mahadik, P. M. Pawar, and R. Muthalagu, "Edge-HetIoT defense against DDoS attack using learning techniques," *Computers & Security*, vol. 132, no. 103347, 2023. https://doi.org/10.1016/j.cose.2023.103347

[39] A. Maheshwari, B. Mehraj, M. S. Khan, and M. S. Idrisi, "An optimized weighted voting-based ensemble model for DDoS attack detection and mitigation in SDN environment," *Microprocessors and Microsystems*, vol. 89, no. 104412, 2022. https://doi.org/10.1016/j.micpro.2021.104412

[40] M. A. Al-Shareeda, S. Manickam, and M. A. Saare, "DDoS attacks detection using machine learning and deep learning techniques: Analysis and comparison," *Bulletin of Electrical Engineering and Informatics*, vol. 12, no. 2, pp. 930–939, 2023. https://doi.org/10.11591/eei.v12i2.4466

[41] S. Mazumder, "5 techniques to handle imbalanced data for a classification problem," *Analytics Vidhya*, 2023.

[42] A. K. Mishra, S. Paliwal, and G. Srivastava, "Anomaly detection using deep convolutional generative adversarial networks in the internet of things," *ISA Transactions*, vol. 145, pp. 493–504, 2024. https://doi.org/10.1016/j.isatra.2023.12.005

[43] D. Mishra, B. Naik, J. Nayak, A. Souri, P. B. Dash, and S. Vimal, "Light gradient boosting machine with optimized hyperparameters for identification of malicious access in IoT network," *Digital Communications and Networks*, vol. 9, no. 1, pp. 125–137, 2023. https://doi.org/10.1016/j.dcan.2022.10.004

[44] M. Mittal, K. Kumar, and S. Behal, "Deep learning approaches for detecting DDoS attacks: A systematic review," *Soft Computing*, vol. 27, no. 18, pp. 13039–13075, 2023. https://doi.org/10.1007/s00500-021-06608-1

[45] M. Mittal, K. Kumar, and S. Behal, "DL-2P-DDoSADF: Deep learning-based two-phase DDoS attack detection framework," *Journal of Information Security and Applications*, vol. 78, no. 103609, 2023. https://doi.org/10.1016/j.jisa.2023.103609

[46] K. Kumari and M. Mrunalini, "Detecting Denial of Service attacks using machine learning algorithms," *Journal of Big Data*, vol. 9, no. 56, 2022. https://doi.org/10.1186/s40537-022-00616-0

[47] M. Türkoğlu, H. Polat, C. Koçak, and O. Polat, "Recognition of DDoS attacks on SD-VANET based on combination of hyperparameter optimization and feature selection," *Expert Systems with Applications*, vol. 203, no. 117500, 2022. https://doi.org/10.1016/j.eswa.2022.117500

[48] N. Soveizi, F. Turkmen, and D. Karastoyanova, "Security and privacy concerns in cloud-based scientific and business workflows: A systematic review," *Future Generation Computer Systems*, vol. 148, pp. 184–200, 2023. https://doi.org/10.1016/j.future.2023.05.015

[49] A. A. Najar and S. M. Naik, "Cyber-Secure SDN: A CNN-based approach for efficient detection and mitigation of DDoS attacks," *Computers & Security*, vol. 139, no. 103716, 2024. https://doi.org/10.1016/j.cose.2024.103716
[50] C. M. Nalayini and J. Katiravan, "Detection of DDoS attack using machine learning algorithms," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 9, no. 7, pp. f223–f232, 2022.
[51] C. M. Nalayini, J. Katiravan, S. Geetha, and J. I. Christy Eunaicy, "A novel dual optimized IDS to detect DDoS attack in SDN using hyper tuned RFE and deep grid network," *Cyber Security and Applications*, vol. 2, no. 100042, 2024. https://doi.org/10.1016/j.csa.2024.100042
[52] M. Ouhssini, K. Afdel, E. Agherrabi, M. Akouhar, and A. Abarda, "DeepDefend: A comprehensive framework for DDoS attack detection and prevention in cloud computing," *Journal of King Saud University - Computer and Information Sciences*, vol. 36, no. 2, 2024. https://doi.org/10.1016/j.jksuci.2024.101938
[53] S. Pandian, "A comprehensive guide on hyperparameter tuning and its techniques," *Analytics Vidhya*, 2022.
[54] O. Pandithurai, C. Venkataiah, S. Tiwari, and N. Ramanjaneyulu, "DDoS attack prediction using a honey badger optimization algorithm-based feature selection and Bi-LSTM in cloud environment," *Expert Systems with Applications*, vol. 241, no. 122544, 2024. https://doi.org/10.1016/j.eswa.2023.122544
[55] R. Panigrahi, S. Borah, M. Pramanik, A. K. Bhoi, P. Barsocchi, S. R. Nayak, and W. Alnumay, "Intrusion detection in cyber–physical environment using hybrid Naïve Bayes—Decision table and multi-objective evolutionary feature selection," *Computer Communications*, vol. 188, pp. 133–144, 2022. https://doi.org/10.1016/j.comcom.2022.03.009
[56] S. J. S. Veluswami, I. Ioannou, P. Nagaradjane, C. Christophorou, V. Vassiliou, S. Charan, and A. Pitsillides, "Detection of DDoS attacks in D2D communications using machine learning approach," *Computer Communications*, vol. 198, pp. 32–51, 2023. https://doi.org/10.1016/j.comcom.2022.11.013
[57] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel class probability features for optimizing network attack detection with machine learning," *IEEE Access*, vol. 11, pp. 98685–98694, 2023. https://doi.org/10.1109/ACCESS.2023.3313596
[58] F. Rustam, M. F. Mushtaq, A. Hamza, M. S. Farooq, A. D. Jurcut, and I. Ashraf, "Denial of service attack classification using machine learning with multi-features," *Electronics*, vol. 11, no. 22, Art. no. 3817, Nov. 2022, doi: 10.3390/electronics11223817.
[59] S. Saha, A. T. Priyoti, A. Sharma, and A. Haque, "Towards an optimized ensemble feature selection for DDoS detection using both supervised and unsupervised method," *Sensors*, vol. 22, no. 23, Art. no. 9144, 2022. https://doi.org/10.3390/s22239144
[60] A. Sarıkaya, B. G. Kılıç, and M. Demirci, "RAIDS: Robust autoencoder-based intrusion detection system model against adversarial attacks," *Computers & Security*, vol. 135, Art. no. 103483, 2023. https://doi.org/10.1016/j.cose.2023.103483
[61] M. S. Elsayed, N.-A. Le-Khac, M. A. Azer, and A. D. Jurcut, "A flow-based anomaly detection approach with feature selection method against DDoS attacks in SDNs," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 4, pp. 1862–1880, 2022. https://doi.org/10.1109/TCCN.2022.3186331
[62] M. A. Setitra, M. Fan, B. L. Y. Agbley, and Z. E. A. Bensalem, "Optimized MLP-CNN model to enhance detecting DDoS attacks in SDN environment," *Network*, vol. 3, no. 4, pp. 538–562, 2023. https://doi.org/10.3390/network3040024
[63] D. M. Sharif and H. Beitollahi, "Detection of application-layer DDoS attacks using machine learning and genetic algorithms," *Computers & Security*, vol. 135, Art. no. 103511, 2023. https://doi.org/10.1016/j.cose.2023.103511
[64] M. Snehi, A. Bhandari, and J. Verma, "Foggier skies, clearer clouds: A real-time IoT-DDoS attack mitigation framework in fog-assisted software-defined cyber-physical systems," *Computers & Security*, vol. 139, Art. no. 103702, 2024. https://doi.org/10.1016/j.cose.2024.103702
[65] S. Sokkalingam and R. Ramakrishnan, "An intelligent intrusion detection system for distributed denial of service attacks: A support vector machine with hybrid optimization algorithm-based approach," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 27, Art. no. e7334, 2022. https://doi.org/10.1002/cpe.7334
[66] G. N. Tikhe and P. S. Patheja, "A wrapper feature selection based hybrid deep learning model for DDoS detection in a network with NFV behaviors," *Wireless Personal Communications*, vol. 133, no. 1, pp. 481–506, 2023. https://doi.org/10.1007/s11277-023-10775-9
[67] G. Tripathi, V. K. Singh, V. Sharma, and M. V. Vinodbhai, "Weighted feature selection for machine learning based accurate intrusion detection in communication networks," *IEEE Access*, vol. 12, pp. 20973–20982, 2024. https://doi.org/10.1109/ACCESS.2024.3362794
[68] A. V. Turukmane and R. Devendiran, "M-MultiSVM: An efficient feature selection assisted network intrusion detection system using machine learning," *Computers & Security*, vol. 137, Art. no. 103587, 2024. https://doi.org/10.1016/j.cose.2023.103587
[69] Z. Zhao, Z. Li, Z. Zhou, J. Yu, Z. Song, X. Xie, and R. Zhang, "DDoS family: A novel perspective for massive types of DDoS attacks," *Computers & Security*, vol. 138, Art. no. 103663, 2024. https://doi.org/10.1016/j.cose.2023.103663
[70] Z. Liu, Y. Wang, F. Feng, Y. Liu, Z. Li, and Y. Shan, "A DDoS detection method based on feature engineering and machine learning in software-defined networks," *Sensors*, vol. 23, no. 13, Art. no. 6176, 2023. https://doi.org/10.3390/s23136176

BIOGRAPHIES OF AUTHORS

Chinyere Chioma Isiekwene obtained her B.Sc. from the University of Ibadan, Oyo State, Nigeria, in 2015, and possess double M.Sc. degrees from the Lagos State University, Ojo-Badagry, Lagos State, in 2020 and the University of Lagos, Akoka, Yaba, Lagos State, Nigeria, in 2023. She is currently pursuing a Ph.D. in Computer Science at the University of Lagos, Nigeria, with research interests encompassing cybersecurity, malware detection, data theft prevention, information security, privacy and trust, data mining techniques for scalable network traffic analysis, anomaly detection, and machine learning. She is a Lecturer in the Faculty of Computing at MIVA Open University, Nigeria, and a member of the Computer Professionals of Nigeria (CPN) and the Nigerian Computer Society (NCS).

Nureni Ayofe Azeez obtained his B.Tech. (Hons.) from the Federal University of Technology, Akure, Nigeria, in 2005, an MSc from the University of Ibadan, Oyo State, Nigeria, in 2008, and a Ph.D. from the University of the Western Cape, South Africa, in 2013, all in Computer Science. His areas of research include Security & Privacy, Trust Management, Access Control, and E-Health. He is a recipient of the Young Scientist Award at the 22nd International CODATA Conference, held in Cape Town, South Africa, in October 2010. He is an associate professor of computer science at the Department of Computer Sciences, University of Lagos, Nigeria.

Akinboro Solomon A. is an Associate Professor from the Department of Computer Science, University of Lagos, Akoka, Nigeria. He holds a B. Tech degree in Computer Engineering from Ladoke Akintola University of Technology, Ogbomosho; an M.Sc. in Computer Science and Engineering; and a PhD in Computer Science from Obafemi Awolowo University, Ile-Ife. Research interests include Data Communication Network, Information Security, Artificial Intelligence and ICT4D. He is a member of the following professional bodies: Nigeria Computer Society, Nigeria Society of Engineers and the Council for the Regulation of Engineering in Nigeria (COREN).