sciendo INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Special Issue: IIoT Sensors in Energy Management | DOI: 10.2478/ijssis-2024-0040 Issue 1 | Vol. 17 (2024)

Advancing DDoS attack detection with hybrid deep learning: integrating convolutional neural networks, PCA, and vision transformers

Jahangir Shaikh[1], Toqeer Ali Syed[2], Syed Aziz Shah[3], Salman Jan[4], Qurat Ul Ain[5], Pradeep Kumar Singh[6]*

[1]Department of Computing and Engineering, Abasyn University, Islamabad, Pakistan

[2]Department of Information Technology, Faculty of Computer and Information System, Islamic University of Madinah, Madinah, Saudi Arabia

[3]Department of Health and Life Sciences, Coventry Center for Intelligent Healthcare, Coventry University, Coventry, UK

[4]Department of Information Technology, Alburaimi University College, Pakistan

[5]Department of Computer Science, School of Computer Science and Creative Technologies, University of the West of England, Bristal, UK

[6]Department of Computer Science and Engineering, Central University of Jammu, Jammu, India

*E-mail: pradeep.cse@cujammu.ac.in

Received for publication June 12, 2024.

Abstract

Distributed denial of service (DDoS) attacks pose a significant security risk, particularly with the increasing reliance on cloud computing and information technology (IT). These attacks not only allow unauthorized users to access services but also deny legitimate users the ability to utilize them. Traditional antivirus solutions and firewalls prove insufficient in detecting DDoS attacks within large networks. Intrusion detection systems (IDS) are essential for detecting unauthorized or malicious activities and ensuring the confidentiality, integrity, and availability of services. However, traditional IDS often rely on predefined signatures and patterns, making them susceptible to evasion tactics. In response, this research introduces a deep learning (DL)-based IDS that integrates convolutional neural networks (CNN) with principal component analysis (PCA) and explores the application of vision transformers (ViT). The proposed hybrid model was tested on the CICDDoS2019 dataset, achieving a notable improvement in detection accuracy. Specifically, the CNN-based model initially identified DDoS attacks with an accuracy of 99.72%. Upon integrating ViT, the model's accuracy further improved to 99.99%. This innovative approach signifies a considerable advancement in the detection capabilities for DDoS attacks and underscores the potential for integrating more sophisticated DL models into cybersecurity defenses.

Keywords

DDOS, Hybrid Deep Learning, CNN, Visions Transformer

I. Introduction

The internet has transformed our way of communication, conducting, interacting, and carrying out day-to-day business operations. Internet services have been integrated into traditional sectors like research, education, banking, defense, medicine, and entertainment. This reliance on the internet by government, military, and commercial entities for daily operations has increased the importance of addressing cybersecurity issues [1]. Data volume transferred on the internet continues to grow, and the importance of cyber security has become more evident. Data are now considered a more valuable resource than oil [2]. According to Macas and Wu [3] there will likely be more IP-connected devices by 2023, which can generate a huge volume of internet protocol (IP) traffic, posing major security concerns. Both the public and private sectors have significant investments in the field of information technology (IT), which has raised the demand for data security. With the rise of cyber threats, reliable and trustworthy cybersecurity

systems need to employ a range of precautionary measures. This necessitates the development of new techniques and measures in cybersecurity at a comparable pace [4]. To establish effective cyber security policies, it is crucial to be flexible in recognizing prospects that may enhance security and adjust to evolving threats. They offer real-time visibility into cyber security as well as enable it to take proactive steps in preventing threats from turning into significant problems. This major effect steered the growth of the IT industry and e-commerce [5]. There are several security risks on the internet, but distributed denial of service (DDoS) attacks are critical along with widespread cyber threats [6]. The aim of carrying out a DDoS attack is to drain the victim's resources. It allows unauthorized users to access services while preventing authorized users from accessing them. The intruder sends a high volume of traffic to the target. As a result, there is unavailability of services for a certain period, hence they cannot be provided to real users. It is one of the most prevalent and aggravating challenges. In recent years various well-known companies, like Amazon elastic compute cloud (EC2) and Rackspace, have faced DDoS attacks [7], resulting in financial losses. Unfortunately, cybercriminals successfully breached and obtained data from Sina Weibo in March 2020, compromising the personal information of 538 million users [8]. Organizations face crippling financial consequences as a result of ransomware attacks, data breaches, and intellectual property theft, which cost billions of dollars a year. DDoS attacks are widely increasing and hazardous due to their prolonged duration and huge amount of traffic. Detecting these attacks is challenging due to their scattered nature and difficulty in distinguishing between normal and attack traffic. DDoS attack is regarded to be the largest threat to the IT industry [9], and a substantial increase is recorded every year. To identify these attacks and ensure network protection, traditional antivirus software and firewalls are insufficient and need efficient intrusion detection systems (IDS). IDS plays a vital role in network security, since Dorothy Denning gave the model of intrusion detection in 1985 [10]. An IDS is made to identify online threats while maintaining service availability and privacy. It gathers and examines information to find any dangers, giving security experts insightful information. To solve the issues faced by cybercrimes, multidisciplinary techniques such as machine learning (ML), deep learning (DL), and statistics are employed. These methods can aid in the efficient analysis and reduction of the hazards connected to cybercrimes. ML- and DL-based approaches are widely utilized as a part of artificial intelligence (AI) applications as described in Figure 1.

Because of the increasing threat landscape, studies have been conducted to create strong cybersecurity measures, which has led to the investigation of cutting-edge technologies like DL. The convolutional neural network (CNN) approach is a well-suited strategy for IDS among various DL algorithms. These algorithms have shown to be effective in a variety of ensemble methodologies, indicating that they have the potential to improve intrusion detection capabilities. Big data can be processed via DL techniques, which are thought to be more sophisticated. They are made to automatically extract pertinent data, which is necessary for creating reliable systems. Network attack analysis may be done quickly and efficiently using the automation of DL approaches. They have shown to be

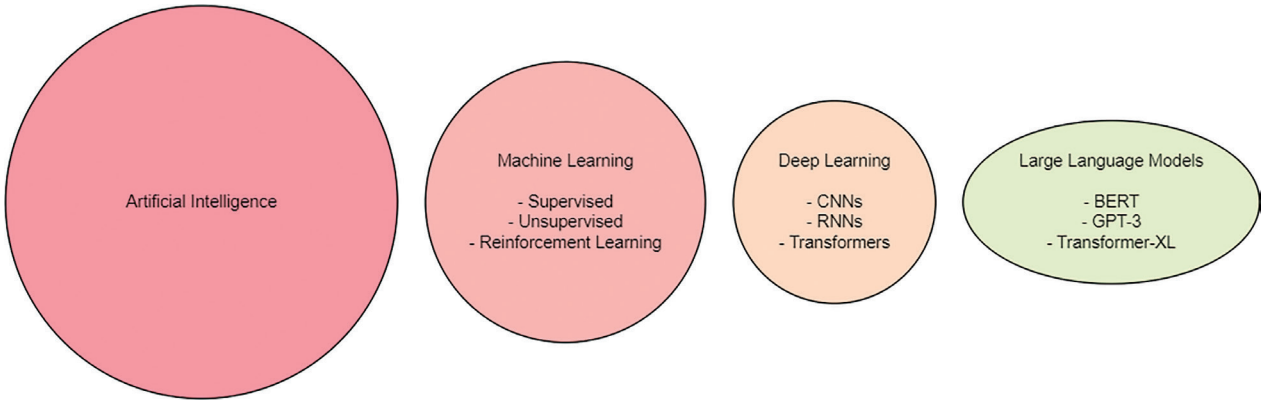

Figure 1: Overview of the AI ecosystem highlighting the roles and interconnections between ML, DL, and LLM. AI, artificial intelligence; CNNs, convolutional neural network; DL, deep learning; LLM, large language models; ML, machine learning.

useful in tackling several cyber security issues, such as DDoS attack detection. The prime focus of this research is to delve into DL techniques for IDS, aiming to improve detection against DDoS attacks.

There can never be a completely foolproof system due to diversity and frequently changing intrusion traffic patterns; however, we can use better tools to make it more difficult for attackers [11]. Limited availability of dataset and traditional datasets do not contain the latest attack patterns and hence undergo numerous problems that result in the absence of diversity of attack which may make model training and evaluation ineffective [12]. For the maximum number of features selected in DL-based approaches, insignificant features must be removed as they increase time and space complexity [13]. Therefore, there is a need to develop an efficient DL-based IDS for DDoS attacks with adequate preprocessing, balanced classes, and comprehensive feature extraction to reduce computing requirement and enhance detection efficacy. The proposed model employs DL methods, primarily CNN, which is used for extracting features from input [14]. In addition to CNN, a vision transformer (ViT) was also explored to assess its capability in enhancing the detection of DDoS attacks by capturing complex patterns in the data. The latest DDoS-specific datasets, CICDDoS2019, which is publicly available, were used to train and evaluate the proposed models [15]. Initially, the dataset was imbalanced, containing very few data records of normal traffic. The Synthetic Minority Oversampling Technique (SMOTE) method has been used to resolve the class imbalance issue, and principal component analysis (PCA) is employed for dimensionality reduction to increase the efficacy and sensitivity of the models. This approach makes this research unique compared with previous studies on DL-based models. The primary objectives of this research are as follows:

- To examine and eliminate null or ambiguous values and redundancies from the dataset that could influence prediction outcomes;
- To develop a hybrid DL-based model that can precisely detect DDoS attacks;
- To improve the accuracy of the model by addressing the problem of computation needs using PCA for dimension reduction and SMOTE for class imbalance;
- To evaluate the model's accuracy against other state-of-the-art approaches.

The further organization of the paper is as follows: Section II presents the background; Section III outlines the literature review; Section IV describes the methodology of the proposed model; evaluation is explained in Section V; whereas Section VI includes the conclusion and future work.

II. Background

a. DDoS attacks

In a DDoS attack, various compromised systems are used to exploit a targeted system, service, or network with a massive volume of incoming traffic, overwhelming its computational resources and making it inaccessible to actual users. DDoS attacks aim to interfere with a specific online service or website, making it unavailable to authorized users and harming the victim organization's finances and reputation. DDoS attacks are usually organized by a network of seized devices or bots, generating a coordinated and dispersed attempt to overwhelm the target with traffic, making it challenging to stop the attack and identify its origin. DDoS attacks can be carried out for several purposes, such as extortion, retaliation, gaining an advantage over competitors, or ideological convictions. With advancements in technology, DDoS attacks have become easier to execute using automated platforms. Attackers can use quick automated tools to attack devices with weak protection. Notable examples include a massive DDoS attack on GitHub in early 2018, utilizing self-learning algorithms based on AI, which caused significant disruption with a capacity of 1.35 Tbps [16]. Additionally, an attack on Amazon Web Services (AWS) in 2019 resulted in an 8-hr outage, impacting the accessibility of cloud services, websites, and applications for numerous customers [17]. In February 2020, the most extensive DDoS attack to date affected Amazon cloud services, leading to 3 days of heightened threat. Another example is a data breach that allowed hackers to obtain the data of more than 533 million Facebook users, found by a white hat security organization in 2021 [18].

b. Traditional way to mitigate DDoS attacks

When attempting to counteract a DDoS attack, the most crucial thing to keep in mind is the difference between genuine and malicious traffic. There are various forms of DDoS attacks on the Internet. A multi-vector DDoS attack uses many attack vectors to overload a target in different ways, which might divert attention away from mitigation efforts focused on a single trajectory. It takes a range of strategies to

oppose various trajectories in a multi-vector DDoS attack. DDoS attack prevention can be challenging, but there are numerous effective techniques as follows.

- *Network breakdown*. Networks can be divided into smaller, easier-to-manage segments to lessen the effect of DDoS attacks. Virtual local area networks (VLANs) may be constructed to do this, and firewalls can prevent an attack from spreading. The optimum approach is zero-trust microsegmentation. Adding a firewall at the device level and hiding devices outside of the operating system continue to be the most dependable ways to prevent DDoS attacks.
- *Monitoring of network traffic*. Monitoring application traffic is crucial. With enough monitoring, we can identify the majority of attacks. DDoS attacks are often done with large amounts of traffic, however, they may also be launched with just one weak HTTP endpoint. When traffic reaches a predetermined level, you ought to receive an alert or warning. Having your monitoring tool alerts configured correctly is the best practice. It assists you in minimizing damage and quickly identifying the DDoS attack.
- *Load balancing of devices*. By dividing the traffic among several servers, a DDoS attack cannot overwhelm one particular server or resource. It is possible to do load balancing with hardware or software methods.
- *Blocking of IP address*. DDoS traffic cannot reach its target server if known or suspected malicious IP sources are blocked.
- *Rate limiting*. Attackers can contact the server indefinitely. It can make it impossible for regular users to access some resources. A rate limit is the most calls or requests that a user is permitted to make in a certain period. When this limit is reached, provide the HTTP error code and temporarily block access.

c. DL-based approach for DDoS detection

DDoS attacks attempt to render a service unusable by saturating it with traffic from numerous sources. The ability to detect complex and sophisticated DDoS attacks is constrained by traditional detection techniques including signature-based and statistical-based approaches. Due to their capacity to automatically identify patterns and features from big datasets, to identify DDoS attacks DL-based techniques are efficacious. Following are a few steps to implement the proposed model:

- *Data collection*. A large dataset of network traffic that includes both benign and DDoS attack traffic should be gathered. The dataset has to be varied and includes a range of DDoS attacks.
- *Data preprocessing*. Preprocessing steps involve the reduction of dimensionality, removing noise, and normalizing the data. This process is very crucial to ensure data quality and lowering the DL model's computational cost.
- *Feature extraction*. In this process, features extracted from the preprocessed dataset can be utilized as input of the DL model. Packet size, frequency, and packet destination IP address are the most common features used in the detection of DDoS attacks.
- *Model training*. To train a DL model such as CNN using preprocessed and feature-extracted data. The model will be trained on CICDDoS2019 dataset traffic to ensure it can accurately detect DDoS attacks.
- *Model evaluation*. To evaluate the model a test dataset was used which consists of DDoS and benign data. Accuracy, precision, recall, and F1 score parameters are used to evaluate the performance of the model.
- *Deployment*. To deploy the DL model which has been trained as a DDoS detection system that can track network traffic in real-time. When a DDoS attack is discovered, the system must be able to notify network managers so they can take the necessary precautions to diminish the attack.

It is important to remember that the performance of DL-based DDoS attack detection is dependent on the data and model's architecture. Thus, it is crucial to adhere to best practices at every stage of the procedure and continually assess and improve the system.

d. Overview of IDS

Anderson J.P. first introduced an IDS in 1980. He described an intrusion or threat as any effort to access information without authorization or change data in a way that renders the system inaccessible or unreliable. The intrusion detection expert system (IDES) was an IDS prototype that Denning [10] suggested in 1985. It is either a hardware or a software used to identify harmful activity on computer networks. An IDS's main goal is to protect computer systems by identifying and recognizing various unsafe

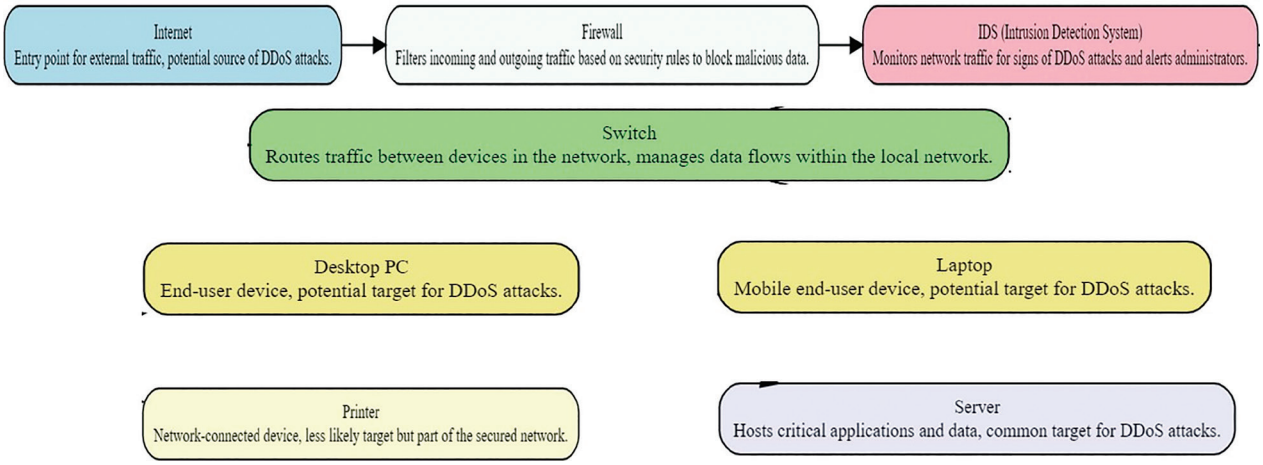

Figure 2: Network diagram depicting components involved in DDoS defense, including internet access points, firewalls, switches, IDS, and end-user devices. DDoS, distributed denial of service; IDS, intrusion detection systems.

network traffic. To keep network confidentiality, integrity, availability, and functioning intact, it is essential to implement IDS into network security [3]. The general architecture of IDS in any computer network is shown in Figure 2.

e. Classification of an IDS

There are several categories in which intrusion detection methods can be classified, primarily based on network infrastructure. An IDS can be classified on the basis of deployment into network and host-based [19]. Whereas on the basis of detection, it is classified as anomaly and signature-based. Figure 3 provides an overview of an IDS.

- *Anomaly based*. Detection of anomalies in any network involves creating a standard profile that represents normal network behavior. It identifies any notable deviations from this established behavior, indicating potential anomalies. However, the disadvantage of utilizing an IDS based on anomalies is the possibility of generating a high number of false positive (FP) detections [19].
- *Signature-based*. It is also called misuse detection and utilizes pattern- matching methods to identify known attacks. This method involves analyzing network activities by comparing them against a database of established signatures or patterns of previous intrusions. When an activity matches a signature stored in the database, it triggers an alarm to alert the system [21].
- *Network-based*. In the modern network environment, network-based IDS (NIDS) is a vital security tool that plays a vital role in detecting and classifying all the traffic flowing between network devices. Since intrusions often exhibit asymmetrical patterns, NIDS analyzes network traffic to differentiate between regular traffic and abnormal activities [19].
- *Host-based*. Host-based intrusion detection system (HIDS) has the objective of gathering security-related information from a particular system or host. It concentrates on detecting intrusions by analyzing the internal components of the host system, such as the operating system and file systems. To accomplish this, HIDS utilizes sensors or agents deployed on vulnerable machines to identify potential intrusions [21].

f. Convolutional neural network

A CNN is an artificial neural network (ANN) with fully connected, pooling, and convolutional layers. These layers work together to turn the raw data into visually appealing representations. In order to find important patterns in the data, the convolutional and pooling layers in Figure 4 analyze the input first and produce various feature maps. The output of these layers is received by a fully connected layer, which classifies it. During training, the convolutional and fully connected layers' weight parameters are optimized by the use of gradient descent. CNN provides the advantage of automated feature extraction, which has gained popularity in recent research. When using CNN in the context of 2D networks, additional preprocessing could be required to make the 1D input compatible with the network's 2D structure, such as photos. For example,

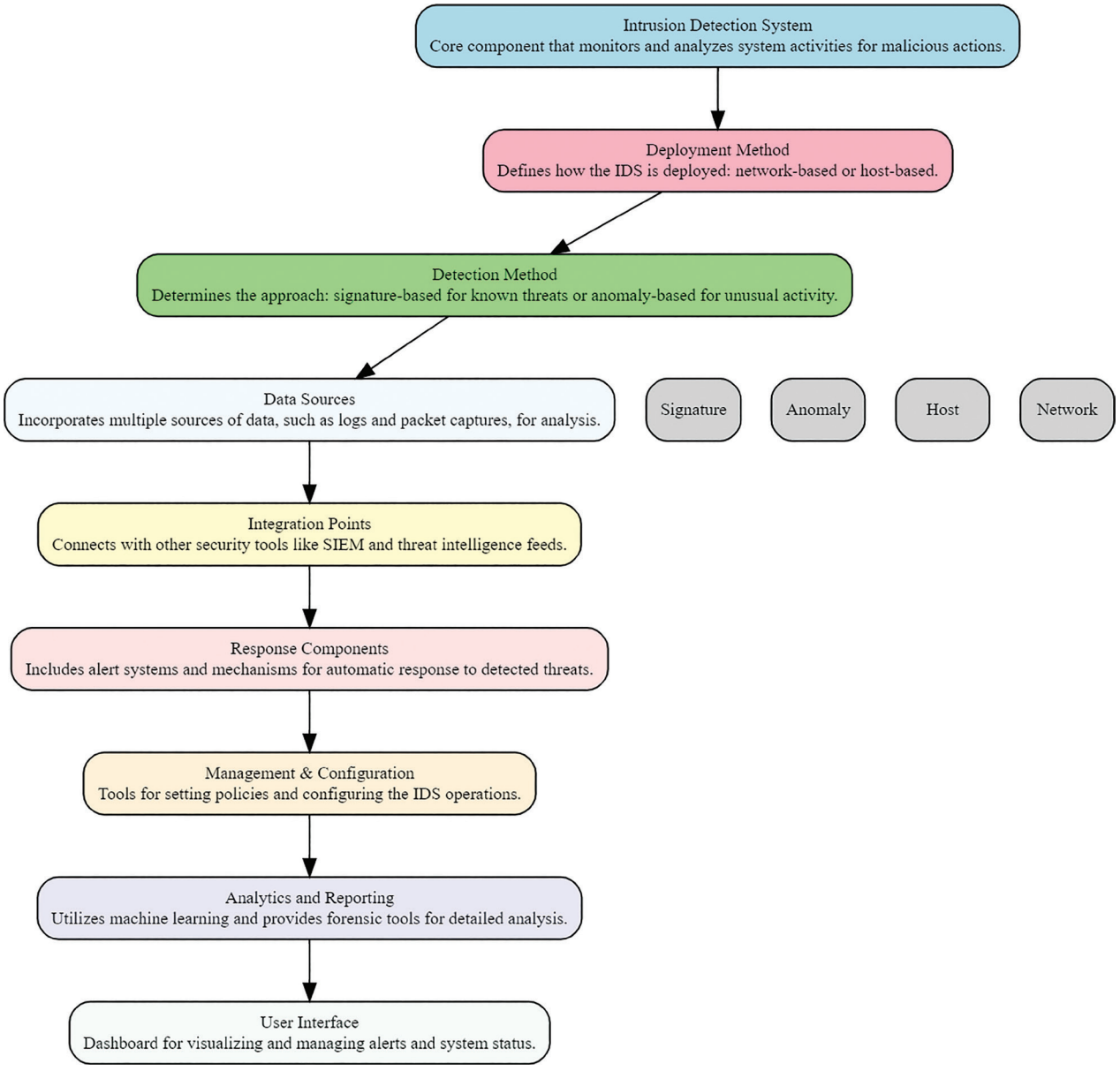

Figure 3: Architectural overview of an IDS, detailing its core components, detection methodologies, data integration points, and response mechanisms. IDS, intrusion detection systems.

Jia et al. [20] presented a unique CNN architecture created especially for DDoS detection on datasets like KDD99, private datasets, and CICDoS2019.

- *Convolutional layer*. In each convolutional layer there are sets of filters that convolve with input to produce feature maps. Each element (i, j) in the feature map is computed by applying the activation function to the sum of element-wise products of W and input x, plus the bias: $h(i, j) = \text{ReLU}(b + ((W \times x)))$.
- *Activation function*. ReLU, sigmoid, softmax, and tanh are examples of commonly used activation functions. The model's activation function adds nonlinearity and aids in the capture of complicated patterns and representations. ReLU and sigmoid activation functions have been used in this research.
- *Pooling layer*. By combining nearby values, pooling layers reduce the spatial dimensions of the feature maps. Average and max pooling are two frequently used pooling operations. The pooling

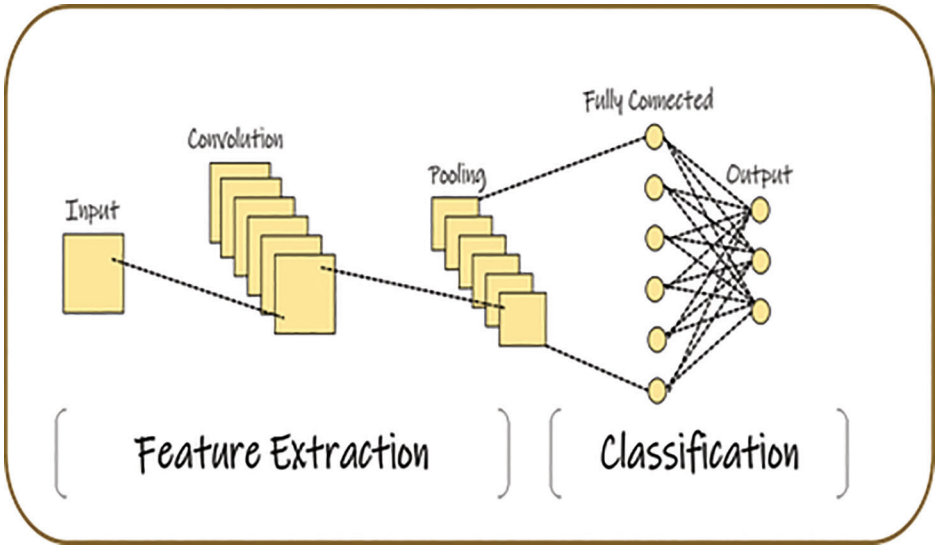

Figure 4: CNN. CNN, convolutional neural network.

function calculates an average or maximum summary statistic for a certain local area of the feature map. $h(i, j) = \text{MaxPooling}(\Sigma(\Sigma(x)))$.

- *Fully connected layer*. This takes the flattened feature map (x) and applies a set of weights (W) and bias terms (b) to produce an output. The activation function is used to introduce nonlinearity to the model: $h = \text{Sigmoid}(W \times x + b)$.

g. PCA

The curse of dimensionality is common and large datasets are hard to interpret. PCA is used for dimensionality reduction of the dataset by creating new uncorrelated variables having maximum variance [1]. In PCA, the original dataset is projected in the direction of high variance, which is orthogonal to each other. We can reduce the computational costs by reducing the number of features. After standardizing the data, eigenvectors and eigenvalues are extracted through the covariance matrix. The number of steps for PCA is as follows : standardize the data, calculate the covariance matrix, find eigenvalues and eigenvectors, take the first K-eigenvectors having the highest eigenvalues, and in the end, project the dataset into k eigenvectors.

h. Class imbalance

Class imbalance issues occur when a model is trained on an unbalanced dataset; it is a serious concern that causes biased outputs with a high false positive rate (FPRs) in ML and DL models. Techniques such as undersampling and oversampling are frequently utilized to address these issues. The SMOTE method is used in this research to address the class imbalance issues. Yizhen Jia et al. initially suggested SMOTE in Ref. [20]. It is very helpful when working with datasets where the majority class (DDoS attacks) outnumbers the minority class (normal traffic). The primary goal of SMOTE is to create synthetic samples for the minority class in order to balance the dataset and provide the model more representative training data. SMOTE is a useful tactic for handling imbalance class sizes.

III. Literature review

Various techniques have been discussed in this chapter for the detection of DDoS attacks in a computer network environment. Most recent research articles have been selected as the latest arising issues and their solutions are addressed in it. To understand how researchers are utilizing ML and DL techniques in detecting DDoS attacks by using an IDS, we have analyzed cutting-edge articles in this section. The use of CNN in IDS has been investigated in numerous papers. For example, DL algorithms like CNN and long short term memory (LSTM) were used to classify benign and malicious traffic. When evaluated on the KDDCup99 dataset, the model obtains an accuracy of 99.99%, while on the network security laboratory knowledge discovery and data mining dataset (NSL-KDD) dataset, it achieves a 99.95% accuracy. On the KDDCup-'99 dataset, a CNN-based IDS uses the LeNet-5 CNN architecture with two convolutional layers. Ten test datasets used in the study had

5,000 randomly chosen instances. Average detection rates were 97% and average false alarm rates (FAS) were 9.9% for the technique. However, it had trouble detecting U2R class samples, whereas Probe and R2L attacks had an average accuracy of 56.26 and 61.47, respectively. A CNN technique was introduced to the KDDCup-'99 dataset by Xiao et al. [21] in another DL model. To initially satisfy the constraints of the CNN model, they used reduction techniques like PCA and autoencoder. Hussain et al. [22] developed a unified framework using deep CNNs to detect DDoS attacks orchestrated by malicious devices promptly. Their framework successfully identified normal and under-attack cells with an accuracy of over 91%. Lucid, a portable deep-learning method for DDoS detection, was introduced by Doriguzzi-Corin et al. [23]. To categorize traffic as malicious or benign, they used a CNN model. Their method established its appropriateness for resource-constrained contexts and produced efficient DDoS detection with little processing overhead. DL and information entropy analysis were incorporated in Wang and Liu [24] suggested the DDoS attack detection method. For evaluating suspicious traffic, their solution used information entropy detection, which was followed by fine-grained packet-based detection using the CNN model. The findings demonstrated a high accuracy rate of 98.98%, showing the efficacy of the model in identifying DDoS attacks in an software-defined networking (SDN) environment.

For DDoS detection, Ma et al. [25] introduced a novel DL method. A CNN model with feature fusion and symmetric logarithmic loss function was introduced. The NSL-KDD dataset was used, and its performance was shown to be superior to that of earlier methods. Using the KDD99 dataset, Kim et al. [26] created an IDS based on LSTM and recurrent neural network (RNN). The model was trained with 100-time steps, 50 batches, and 500 epochs, and during testing and validation, it achieved excellent accuracy results of 98.93%. Naseer et al. [27] propose three deep neural network algorithms (RNN, CNN, and autoencoder (AE)) for developing an IDS with the NSL-KDD dataset. These models were contrasted with outdated ML models like SVM, decision trees, and K-nearest neighbors (KNN). The NSLKDDTest+ and NSLKDDTest+ datasets were used for the training. The models' accuracy ranged from 82% to 89%, with LSTM achieving the best accuracy at 89%. Using the CICDDoS2019 dataset, Aslan [28] provided a strategy for identifying DDoS attacks. This methodology centered on applying various methods to reduce dimensionality reduction. Gain Ratio and random forest (RF) were used in the study to attain 99% accuracy, whereas other classifiers like KNN and AdaBoost had lower accuracy rates. Gaur and Kumar [29] used the multilayer long short-term memory (M-LSTM) model for DDoS attack detection and achieved Precision, Recall, and F1 scores of 98.75%, 97.5%, and 98%, respectively, by using the CICDDoS2019 dataset. Kumar et al. [30] created an LSTM model to identify DDoS attacks using the CICDDoS2019 dataset and achieved an accuracy score of 98%. Using PCA and AutoEncoder, the authors of [31, 32] reduced the number of features. After that, they employed Lenet5 CNN to detect intrusions on the KDDCup99 dataset and found that the CNN was more effective at doing so. The CNN-IDS model performs better in access control (AC), FAR, and timeliness than conventional algorithms, based on testing findings. Many research studies have relied on outdated datasets, which can create a performance gap between frequently occurring and less frequently occurring attack classes. A useful dataset needs to include both old and new attacks in a balanced proportion.

IV. Proposed solution

The effectiveness of a DL-based model, particularly those employed in cybersecurity depend on the quality of the dataset. This section delves into the crucial steps of acquiring and preprocessing the dataset, ensuring that the data are appropriately prepared for subsequent model training and evaluation which addresses the difficulties with handling missing values, class imbalance, and standardizing data values. The dataset must be in the proper format for training in order to get optimal performance. The dataset has been prepared where preprocessing involves several steps.

a. Acquisition of dataset

The well-known DDoS-specific CICDoS2019 dataset, released by the Canadian Institute for Cybersecurity and publicly available on their website [33], is being utilized in this research. The CICDDoS2019 dataset detail is shown in Table 1 [34], which includes benign and wide-ranging DDoS attacks. Intrusion detection techniques in DL heavily rely on the availability of datasets. However, the scarcity of such datasets is due to privacy and regulatory concerns regarding the sensitive information present in network traffic. To overcome this challenge, researchers often resort to creating simulated data. However, these simulated datasets often lack completeness and do not

Table 1: CCICDDoS2019 dataset

Type	Records
Benign	56,863
LDAP	2,179,930
TFTP	20,082,580
DNS	5,071,011
NetBIOS	4,093,279
MSSQL	4,522,492
SSDP	2,610,611
NTP	1,202,642
SNMP	5,159,870
UDP	3,134,645
SYN	1,582,289
UDP-Lag	366,461
WebDDoS	439
Total	50,063,112

LDAP, lightweight directory access protocol; MSSQL, microsoft SQL server; NTP, network time protocol; SNMP, simple network management protocol; SSDP, simple service discovery protocol; SYN, synchronize; TFTP, trivial file transfer protocol; UDP-Lag, UDP-based lag attack; WebDDoS, web-based distributed denial of service.

adequately cover the range of application behaviors. Several public domain datasets have been widely used for IDS to overcome this issue.

There are two types of DDoS attacks. Response-based authentication that adheres to a standard protocol worldwide is called reflection-based authentication. It consists of domain name system (DNS), TFTP, and NetBIOS types. Another type of attack known as "exploitation-based" attack occurs when a hacker tries to harm oneself by taking advantage of a vulnerable system. It consists of SYN, UDP, and UDPLag. This dataset was selected mainly due to its high potential for future research applications and the fact that it is one of the most recent in the IDS field for detecting DDoS attacks. The dataset contains around 50 million data points; 2.2 million records (2 lac from each class) were selected due to the limitations of computational resources.

b. Dataset preprocessing

The removal of null values and outliers, normalization, dimensionality reduction, and resolving class imbalance issues are all included in data preprocessing. Certain features, such as SimilarHTTP, Flow ID, Source Port, Unnamed: 0, Source IP, Destination IP, Timestamp, Destination Port, and Label, were found redundant and removed from the dataset in order to prepare it for model training. The following standard procedures were used to preprocess the dataset.

- *Data cleaning*. Removing redundant, NaN, or unnecessary data points from the dataset is one of the most crucial phases in lowering noise and raising the overall quality of the data. Moreover, the extremely low variance was reduced for 12 features: Idle Std, Max, Mean, Min, Inbound, min seg size forward, act data pkt fwd, Active Std, Min, Mean, Max, and Init Win bytes backward. Removing these data improves the model's efficiency and reduces computational complexity.
- *Label encoding*. Using the Label Encoder Sklearn package, the label encoding method is utilized to transform categorical input into numerical data. In order to perform binary classification, the model was trained to distinguish between normal and malicious traffic. The labels for attack and normal traffic were encoded as binary values of 0 and 1, respectively.
- *Dimensionality reduction*. The original dataset is projected using the unsupervised PCA dimensionality reduction approach in an orthogonal direction to the direction of high variance. The steps include:
  1. Calculate the covariance matrix.
  2. Find eigenvalues and eigenvectors.
  3. Take first K-eigenvectors having the highest eigenvalues.
  4. Project dataset into k eigenvectors.

Based on the variance and eigenvalues, the dataset's dimensions were trimmed down to 20 features as per cumulated explained variance shown in Figure 5. PCA is used to reduce dimensionality and select the top 20 features. Models with reduced dimensionality perform better in terms of accuracy and sensitivity rates while consuming less processing power.

- *Dataset split*. Because it might impact the outcomes, keeping the dataset balanced is essential for both training and testing the model. The Sklearn library is used to partition data: 70% of the data is utilized for training, while 30% is used for testing and validation.

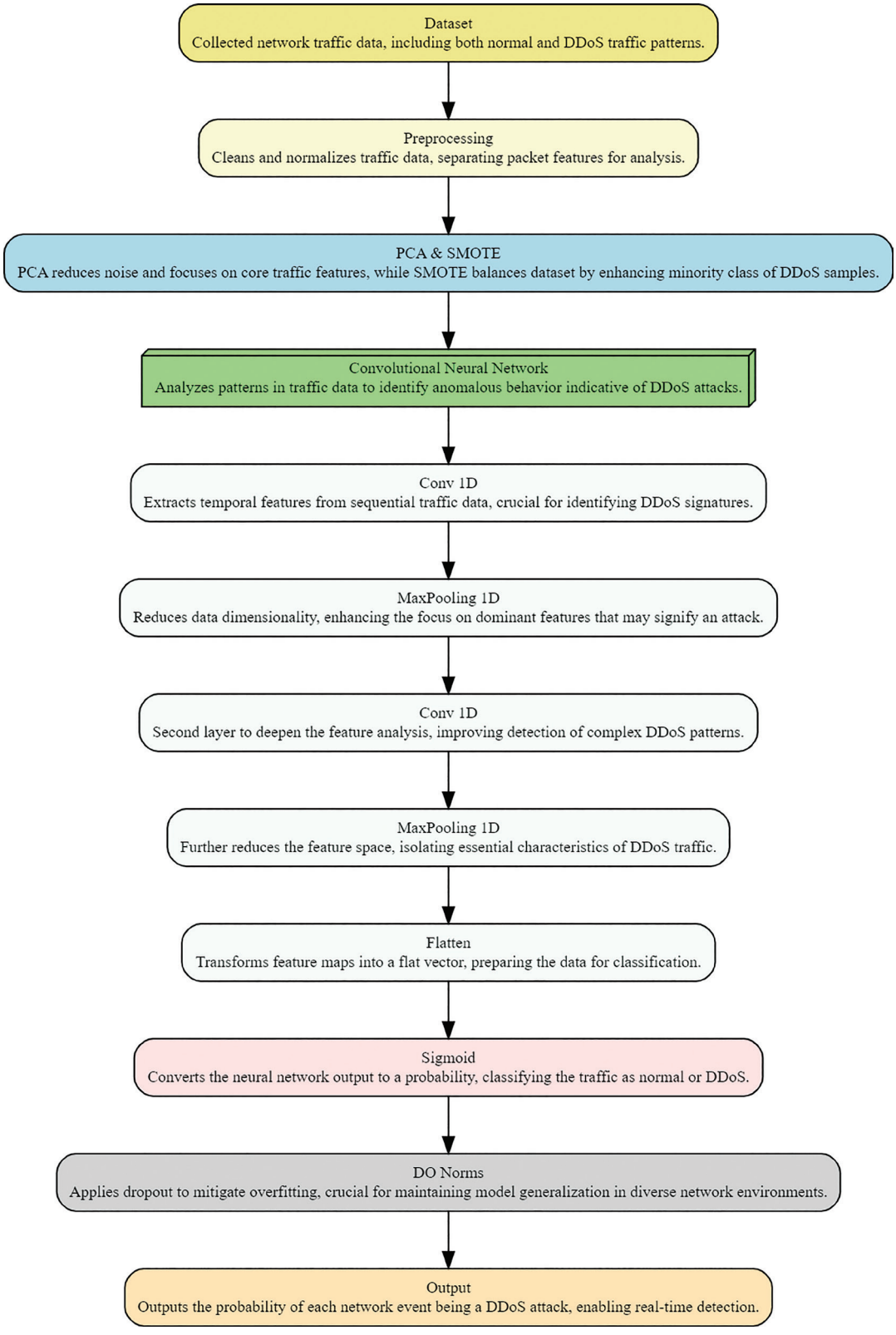

Figure 5: Proposed architecture processing network traffic data through CNN for DDoS attack detection, incorporating PCA and SMOTE for data preparation. CNN, convolutional neural network; DDoS, distributed denial of service; PCA, principal component analysis; SMOTE, Synthetic Minority Oversampling Technique.

c. Methodology

In order to detect DDoS attacks using the CICDDoS2019 dataset for binary class classification, this research presents a DL-based CNN model. Initially, the CNN layer receives the data. Filters in the convolution layer extract the most significant features, producing a feature map. This is followed by a max pooling layer to reduce dimensionality. To prevent overfitting, a dropout layer is employed, followed by a sigmoid activation function to distinguish between DDoS and normal traffic. The proposed DL-based CNN model, shown in Figure 5, consists of two convolutional layers with 32 and 64 filters followed by max pooling layers for downsampling. Flattening and dense layers are added with L2 regularization to avoid overfitting. The final output layer uses a sigmoid activation function for binary classification shown in Figure 5. The hyperparameter values of the model are as follows:

- Activation function: Relu, Sigmoid
- Loss function: binary crossentropy
- Optimizer: RMSprop
- Learning rate: 0.0001
- Epochs: 5
- Dropout: 0.2
- Regularization: L2
- Dimensionality: PCA
- Class imbalance: SMOTE
- Batch size: 512
- Dataset split: 70/15/15

In the study, a dual-path architecture was proposed to enhance the detection of DDoS attacks, utilizing both CNN and ViT. Initially, the network traffic data, characteristic of DDoS and benign activities, underwent a rigorous preprocessing phase in which features were normalized and encoded into a format suitable for image-like data processing. This preprocessed data was then fed simultaneously into two distinct model pathways: the CNN and ViT. The CNN pathway focused on leveraging spatial hierarchies through convolutional layers, which have been the primary focus of this paper. Concurrently, the ViT pathway processed the same data through its transformer layers, designed to capture complex patterns via self-attention mechanisms. Although the primary focus of the research remained on the CNN model, ViT was introduced to compare its performance against traditional CNN architectures. The comparative analysis revealed that ViT not only

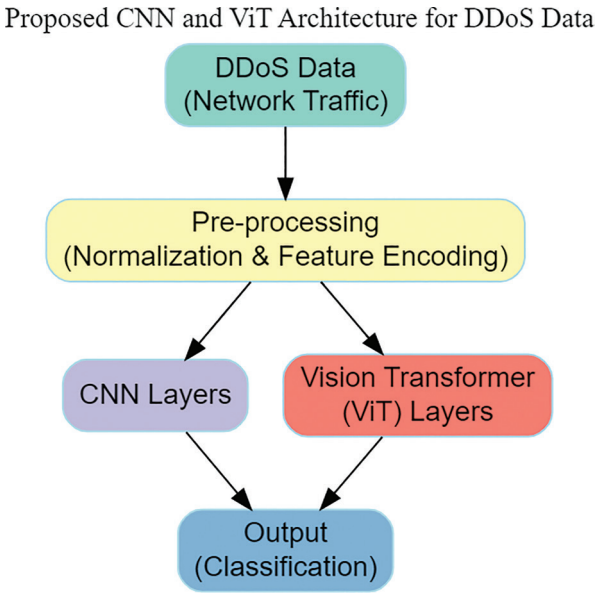

Figure 6: Proposed architecture diagram illustrating the dual-path approach for DDoS attack detection, featuring both CNN and ViT layers. CNN, convolutional neural network; DDoS, distributed denial of service; ViT, vision transformers.

complemented CNN's capabilities but also demonstrated superior performance in recognizing intricate patterns indicative of DDoS attacks, thereby affirming its potential as a powerful tool in cybersecurity defenses. The process of the methodology is mentioned in Figure 6.

d. Experimental environment for CNN and ViT

The Kaggle open-source platform was utilized for CNN experiment in this research to evaluate the proposed hybrid DL-based model with the CICDDoS2019 dataset. Kaggle was chosen for its open-source design, cloud-based Jupyter Notebook environment, and free availability. On Kaggle, users have the opportunity to collaborate, access, and exchange datasets, utilize notebooks with graphics processing unit (GPU) support, and participate in competitions with fellow data scientists to tackle various data science challenges.

The Hugging Face Transformers library was utilized to apply a ViT model to the DDoS dataset, which had been previously transformed into an image-like format. The selected ViT model, initially

pretrained on large-scale image datasets, was fine-tuned on the synthetic images derived from network traffic features characteristic of DDoS and benign traffic. This fine-tuning process involved adjusting the model's top layers to suit the classification tasks, specifically to distinguish between DDoS attacks and normal activities. During the training phase, the model's deeper layers were fine-tuned while the initial layers were kept frozen to leverage the generic features learned from natural images. The training was optimized through meticulous adjustments of learning rates and batch sizes, tailored to the specific patterns and anomalies present in network data. Upon completion, ViT achieved an improved accuracy rate, demonstrating the model's efficacy in capturing complex, nuanced patterns in DDoS attack behaviors, which markedly enhanced the detection capabilities.

V. Evaluation of model

a. Evaluation criteria

This study employs a number of criteria, including accuracy, precision, F1 score, sensitivity or recall, and duration, to assess the model's efficacy. Table 2 offers a thorough synopsis of the classification findings along with a full explanation of the confusion matrix. The evaluation metrics are as follows.

- True positive (TP): Represents the correctly detected attacks.
- FP: Represents the incorrectly detected normal traffic as attacks.
- TN: Represents the correctly detected normal traffic.
- FN: Represents the incorrectly detected attacks as normal traffic.

The following metrics are calculated using these mathematical equations:

Table 2: Confusion matrix

	Actual positive	Actual negative
Positive predict	TP	FP
Negative predict	FN	TN

FN, false negative; FP, false positive; TN, true negative; TP, true positive.

- Precision: It measures the ratio of correctly predicted attacks by all data categorized as attacks:

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

- Accuracy: It measures the ratio of correct predictions to the total predictions:

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

- F1-Score: It is the combination of precision and recall and is calculated by the harmonic mean of precision & recall:

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

- Sensitivity or Recall: Also called TP rate, it measures the ratio of correctly predicted attacks by all actual attack instances:

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

- Training time: Represents the time taken by the model during training.

b. Experimental results

This section includes the results obtained by the CNN model and comparison with other state-of-the-art models. The proposed CNN model achieves superior results in classifying DDoS attacks. The CNN design consists of multiple layers, including pooling layers, convolution layers, and fully connected layers. The Sigmoid and Relu activation functions are used to classify objects by assigning probabilistic values. Each layer has a specific purpose and is responsible for extracting abstract features that are not easily visible to humans. The performance of the model was assessed using PCA as a dimensionality reduction method on the open-source Kaggle platform, using 30, 20, and 10 features; 2.2 million data points are selected from the dataset. Since the data were initially unbalanced, we employed the SMOTE approach, which produces 4.4 million data points by balancing the dataset's normal and DDoS traffic.

As shown in Figure 7, the model's highest accuracy was attained with 20 features using PCA and binary classification at 99.72%. The graph presents a detailed comparison of three CNN model

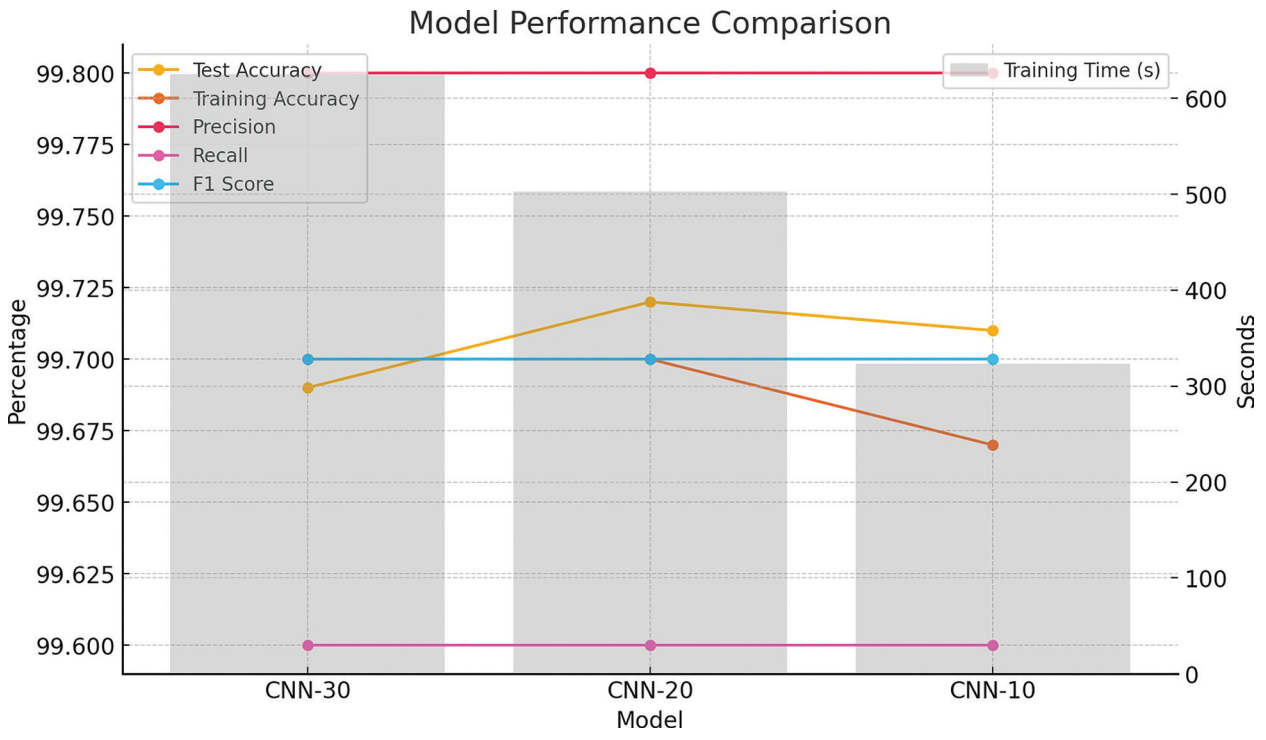

Figure 7: Performance and efficiency of CNN models: This graph shows key performance metrics—Test Accuracy, Training Accuracy, Precision, Recall, and F1 Score—across three CNN configurations (CNN-30, CNN-20, CNN-10). It also illustrates the reduction in training time for each model, emphasizing the balance between model accuracy and computational efficiency. CNN, convolutional neural networks.

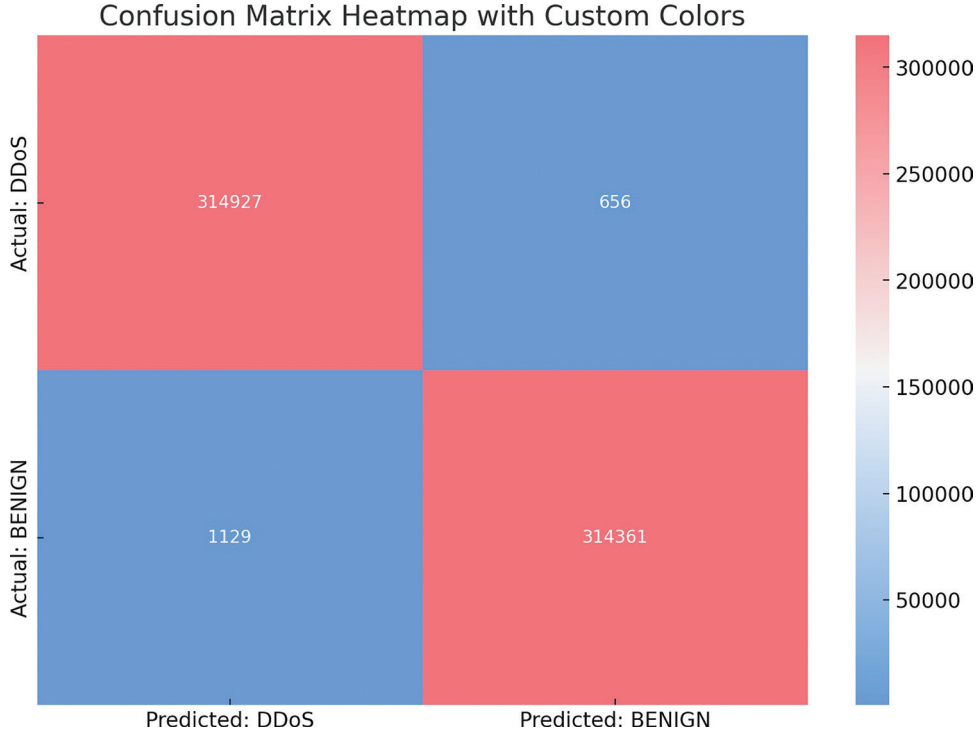

Figure 8: Confusion matrix. DDoS, distributed denial of service.

configurations, CNN-30, CNN-20, and CNN-10, highlighting their performance on various metrics and training times. All models show a high level of accuracy, precision, recall, and F1 score, maintaining values around 99.6%–99.8%, indicating robust performance across configurations. The CNN-20 model slightly outperforms the others in test accuracy at 99.72%, although the differences are minimal, suggesting that all models are similarly effective. Notably, the training time decreases significantly from CNN-30 to CNN-10, from over 600 s to just about 323 s, illustrating a marked improvement in efficiency with simpler or reduced models. This decrease in training time without a significant drop in performance metrics suggests that less complex models may provide a more efficient solution while maintaining high accuracy and precision, making them suitable for environments where computational resources or time are constraints. Figure 8 also shows the confusion matrix.

c. ViT

ViT represented a shift from CNN by applying the transformer architecture, originally designed for natural language processing, to image analysis. Instead of processing the entire image at once, ViTs divided the image into patches and treated each patch as a token similar to words in a sentence. This allowed the model to learn spatial hierarchies and apply self-attention mechanisms across patches, facilitating a deeper understanding of the relationships and contextual relevance within the image.

For applying a ViT to a DDoS dataset, the data had been transformed into a format that mimicked visual data. This involved encoding network traffic features – such as packet size, frequency, type, source/destination IPs — into image-like grid formats where each “pixel” or patch represented different aspects or combinations of network features. These images then served as input for ViT, allowing it to learn and identify patterns indicative of DDoS activity. The self-attention mechanism of ViT was particularly advantageous in this context, as it could highlight unusual patterns or anomalies in network traffic that were characteristic of DDoS attacks.

ViT outperforms the CNN models across several key performance metrics. It achieves a test accuracy of 99.99%, which is significantly higher compared with the highest among the CNN models (CNN-20) at 99.72%. This improvement in accuracy indicates ViT’s superior capability in handling complex pattern recognition tasks, potentially due to its attention mechanisms which allow it to focus more effectively on important features within the data. Furthermore, ViT also shows a notable improvement in Recall, at 99.95%, which is crucial for minimizing false negatives (FNs) — a critical aspect in DDoS detection where missing an attack can be costly. The F1 Score of ViT is also higher at 99.85%, suggesting a better balance between precision and recall compared with the CNN models. This efficiency, combined with superior performance, underscores ViT’s potential as a powerful tool in network security contexts, particularly in detecting and analyzing DDoS attacks. Further details on the implementation and technical specifications of ViT are beyond the scope of this paper and will be explored in-depth in a subsequent publication.

VI. Conclusion and future work

DDoS attack detection has become increasingly challenging due to evolving network behaviors and attack patterns, particularly with traditional intrusion detection methods. To address this challenge, our research developed a unique approach utilizing a DL-based CNN model enhanced by PCA. The DDoS-specific CICDDoS2019 dataset was employed in this study. A one-dimensional CNN-based intrusion detection model was presented, incorporating three primary phases: data preprocessing, classification, and hyper-parameter optimization. Extensive experimentation determined the optimal combination of parameters, notably, a configuration of 32–64 convolutional kernels with a learning rate of 0.0001 yielded the best results. Normalization of incoming traffic attributes and feature reduction from 86 to 20 attributes through PCA demonstrated that the CNN model performs robustly, achieving an accuracy of 99.72% and requiring the least amount of training time. This model surpasses existing benchmarks and prior studies, offering high accuracy and a low FAR for DDoS attacks, making it suitable for application across other datasets. Future work will explore additional DL classifiers, class imbalance, dimensionality reduction techniques, and hyper-parameter optimization on various datasets. Although the implementation details and full exploration of ViT, which shows promise for further enhancing detection capabilities, are beyond the scope of this paper and will be addressed in a subsequent publication, the research un-derscores that ongoing developments in DDoS threat mitigation remain crucial, and every piece of research contributes significantly to these efforts. Finally, the implications of the findings and real-world applications of the research and future work are provided hereunder.

Implications of findings: This research elaborates on how the significant improvement in detection accuracy (up to 99.99% with the integration of ViT) can enhance the effectiveness of IDS in mitigating DDoS attacks. We discussed the broader impact of our findings on the cybersecurity field, particularly in the context of increasingly sophisticated cyber threats.

Real-world applications: This research provides examples of how the proposed hybrid model could be applied in real-world scenarios, such as in cloud computing environments, financial institutions, and government agencies. We also highlighted the potential for integrating this model into existing cybersecurity infrastructures to improve the detection and prevention of DDoS attacks.

Future directions: This research outlines potential future research directions, including exploring other DL architectures and further optimizing the model for deployment in various network environments. We also mentioned the importance of adapting the model to address emerging threats as cyberattack strategies continue to evolve.

Acknowledgments

The authors thank ABASYN University Islamabad (www.abasynisb. edu.pk) Islamabad, Pakistan.

Conflicts of interest

The authors have no conflicts of interest to disclose.

References

[1] F. Laghrissi, S. Douzi, K. Douzi, and B. Hssina, "Intrusion detection systems using long short-term memory (LSTM)," *Journal of Big Data*, vol. 8, no. 1, 2021.

[2] "The world's most valuable resource is no longer oil, but data," *The Economist*, [Accessed: 05-Mar-2023].

[3] M. Macas and C. Wu, "Review: Deep learning methods for cyber security and intrusion detection systems," in *2020 IEEE Latin-American Conference on Communications (LATINCOM)*.

[4] A. Rudskoy, A. Borovkov, P. Romanov, and O. Kolosova, "Reducing global risks in the process of transition to the digital economy," *IOP Conference Series: Materials Science and Engineering*, vol. 497, p. 012088, 2019.

[5] A. Kushairi, R. Singh, and M. Ong-Abdullah, "The oil palm industry in malaysia: Thriving with transformative technologies," *Journal of Oil Palm Research*, vol. 29, no. 4, pp. 431–439, 2017.

[6] A. Alsirhani, S. Sampalli, and P. Bodorik, "Ddos attack detection system: Utilizing classification algorithms with apache spark," in *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, 2018.

[7] O. A. Wahab, J. Bentahar, H. Otrok, and A. Mourad, "Optimal load distribution for the detection of vm-based ddos attacks in the cloud," *IEEE Transactions on Services Computing*, vol. 13, no. 1, pp. 114–129, 2020.

[8] A. Yao, G. Li, X. Li, F. Jiang, J. Xu, and X. Liu, "Differential privacy in edge computing-based smart city applications: Security issues, solutions, and future directions," *Array*, 2023.

[9] R. V. Deshmukh and K. K. Devadkar, "Understanding ddos attack & its effect in cloud environment," *Procedia Computer Science*, vol. 49, pp. 202–210, 2015.

[10] D. E. Denning, "An intrusion-detection model," *IEEE Transactions on Software Engineering*, vol. SE-13, no. 2, pp. 222–232, 1987.

[11] J. A. Perez-Diaz, I. A. Valdovinos, K.-K. R. Choo, and D. Zhu, "A flexible sdn-based architecture for identifying and mitigating low-rate DDoS attacks using machine learning," *IEEE Access*, vol. 8, pp. 155 859–155 872, 2020.

[12] M. S. Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, "Ddosnet: A deep-learning model for detecting network attacks," in *2020 IEEE 21st International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, 2020.

[13] P. Verma, S. Tapaswi, and W. W. Godfrey, "An adaptive threshold-based attribute selection to classify requests under ddos attack in cloud-based systems," *Arabian Journal for Science and Engineering*, vol. 45, no. 4, pp. 2813–2834, 2020.

[14] Rusyaidi, S. Jaf, and Z. Ibrahim, "Detecting distributed denial of service in network traffic with deep learning," *International Journal of Advanced Computer Science and Applications*, vol. 13, pp. 34–41, 2022.

[15] "Ddos evaluation dataset (cicddos2019)," 2019, [Accessed: 03-05-2023]. [Online]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html

[16] "What was the largest ddos attack of all time?" [Accessed: 05-03-2023]. [Online]. Available: https://www.cloudflare.com/learning/ddos/famous-ddos-attacks

[17] "Eight-hour DDoS attack struck AWS customers," 2019, [Accessed: 25-Mar-2023]. [Online]. Available:

https://www.darkreading.com/cloud/ eight-hour-ddos-attack-struckaws-customers/d/d-id/1336165

[18] M. Alazab, S.-H. Hong, and J. Ng, "Louder bark with no bite: Privacy protection through the regulation of mandatory data breach notification in australia," *Future Generation Computer Systems*, vol. 116, pp. 22–29, 2021.

[19] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," 2019.

[20] Y. Jia, F. Zhong, A. Alrawais, B. Gong, and X. Cheng, "Flow-guard: an intelligent edge defense mechanism against ioi ddos attacks," *IEEE Internet of Things Journal*, no. 1, 2020.

[21] Y. Xiao, C. Xing, T. Zhang, and Z. Zhao, "An intrusion detection model based on feature reduction and convolutional neural networks," *IEEE Access*, vol. 7, pp. 42 210–42 219, 2019.

[22] B. Hussain, Q. Du, B. Sun, and Z. Han, "Deep learning-based ddos-attack detection for cyber–physical system over 5g network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 860–870, 2021.

[23] R. Doriguzzi-Corin, S. Millar, S. Scott-Hayward, J. Martinezdel Rincon, and D. Siracusa, "Lucid: A practical, lightweight deep learning solution for ddos attack detection," *IEEE Transactions on Network and Service Management*, vol. 17, no. 2, pp. 876–889, 2020.

[24] L. Wang and Y. Liu, "A ddos attack detection method based on information entropy and deep learning in sdn," in *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2020.

[25] L. Ma, Y. Chai, L. Cui, D. Ma, Y. Fu, and A. Xiao, "A deep learning-based ddos detection framework for internet of things," in *2020 IEEE International Conference on Communications (ICC)*, 2020.

[26] J. Kim, J. Kim, H. L. Thu, and H. Kim, "Long short-term memory recurrent neural network classifier for intrusion detection," in *2016 International Conference on Platform Technology and Service (PlatCon)*, 2016.

[27] S. Naseer *et al.*, "Enhanced network anomaly detection based on deep neural networks," *IEEE Access*, vol. 6, pp. 48 231–48 246, 2018.

[28] Ö. Aslan, "A methodology to detect distributed denial of service attacks," *Informatics Tech Journal*, vol. 15, no. 2, pp. 149–158, 2022.

[29] M. V. Gaur and R. Kumar, "MLSTM: Multiclass long short-term memory based approach for detection of ddos attacks," *Philippine Statistician*, vol. 71, no. 3s2, pp. 1375–1394, 2022.

[30] 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.

[31] Sowmya and T. M. Anita, "An intelligent hybrid gapi feature selection technique for network intrusion detection systems," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 7s, pp. 718–731, 2023.

[32] L. Y. Ahmed, M. M. Hamdy, and H. Mahmoud, "Improved ddos detection utilizing deep neural networks and feedforward neural networks as auto encoder," *Future Internet*, vol. 14, no. 8, pp. 240–248, 2022.

[33] "Ddos evaluation dataset (cicddos2019)," 2019, [Accessed: 25-03-2023]. [Online]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html

[34] M. A. Ferrag, L. Shu, H. Djallel, and K.-K. R. Choo, "Deep learning-based intrusion detection for distributed denial of service attack in agriculture 4.0," *Electronics (Basel)*, vol. 10, no. 11, p. 1257, 2021.