Pattern Augmented Lightweight Convolutional Neural Network for Intrusion Detection System

Yonatan Embiza Tadesse and Young-June Choi *

Posted Date: 26 January 2024

doi: 10.20944/preprints202401.1875.v1

Keywords: Anomaly Detection; Convolutional Neural Networks; Deep Learning; DDoS; DoS; Image Dataset; Intrusion Detection System; Lightweight Model; Machine Learning; Pattern Augmented; Spectrogram

Article

Pattern Augmented Lightweight Convolutional Neural Network for Intrusion Detection System

Yonatan Tadesse and Young-June Choi *

Department of Artificial Intelligence, Ajou University, Suwon, South Korea; enyonathan@ajou.ac.kr
* Correspondence: choiyj@ajou.ac.kr; Tel.: +82-31-219-2634

Abstract: As the world increasingly becomes more interconnected, the demand for safety and security is ever-increasing, particularly for industrial networks. This has prompted numerous researchers to investigate different methodologies and techniques suitable for intrusion detection systems (IDS) requirements. Over the years, many studies have proposed various solutions in this regard including signature-based and machine-learning (ML) based systems. More recently, researchers are considering deep learning (DL) based anomaly detection approaches. Most proposed works in this research field aimed to achieve either one or a combination of high accuracy, considerably low false alarm rates (FARs), high classification specificity and detection sensitivity, achieving lightweight DL models, or other ML and DL-related performance measurement metrics. In this study, we propose a novel method to convert a raw dataset to an image dataset to magnify patterns. Based on this we devise an anomaly detection for IDS using a lightweight convolutional neural network (CNN) that classifies denial of service and distributed denial of service. The proposed methods were evaluated using a modern dataset, CSE-CIC-IDS2018, and a legacy dataset, NSL-KDD. We have also applied a combined dataset to assess the generalization of the proposed model across various datasets. Our experimental results have demonstrated that the proposed methods achieved high accuracy and considerably low FARs with high specificity and sensitivity. The resulting loss and accuracy curves have also demonstrated the excellent generalization of the proposed lightweight CNN model, effectively avoiding overfitting. This holds for both the modern and legacy datasets, including their mixed version.

Keywords: anomaly detection; convolutional neural networks; deep learning; DDoS; DoS; image dataset; intrusion detection system; lightweight model; machine learning; pattern augmented; spectrogram

1. Introduction

Computer networks have revolutionized various aspects of our lives, including communication, knowledge acquisition, and interaction. Numerous sectors, such as healthcare, manufacturing, finance, education, aviation, and entertainment, heavily rely on computer networks, including cloud computing and online gaming. The proliferation of cloud computing has also paved the way for the internet of things (IoT), which has found applications across diverse industries. However, the widespread adoption of computer networks and the Internet has also created opportunities for cyberattacks, prompting cybersecurity concerns. Fortinet, a prominent network security provider, reports that global businesses expended more than $170 billion in 2022 to counter cybercrimes, highlighting the persistent disruption caused by these threats [1]. Malware & ransomware, phishing, distributed denial of service (DDoS), denial of service (DoS), structured query language (SQL) injection, zero-day exploit, domain name system (DNS) tunneling, and Man-in-the-middle are among the most common cyberattacks and network breaches. Particularly, DDoS and malware attacks can be considered the most dangerous attacks depending on the attack scales and the type of the attacked industry.

 doi:10.20944/preprints202401.1875.v1

In this study, we aim to specifically examine DDoS attacks. These attacks have been observed to exhibit a growing level of sophistication, frequency, volume, and efficiency, and possess the capability to cause massive damage against various entities. DDoS attacks are a significant cybersecurity threat, with over 8.4 million attacks recorded in 2019 alone, averaging 23,000 attacks per day or 16 attacks per minute [2]. These attacks can severely damage organizations, disrupting applications and services by overwhelming networks with malicious traffic. Motivations for DDoS attacks range from hacktivism to cybercrime and espionage [3]. Recent instances include attacks on government websites in Ukraine before the Russian invasion [4]. DDoS attacks utilize multiple compromised computers and IoT devices to flood networks, targeting critical assets such as physical locations, data centers, servers, and domains [3]. In cybersecurity and intrusion detection, various mechanisms have been proposed for detecting, identifying, and classifying cyberattacks. DDoS attacks can be categorized based on the network components they target, such as Layer-7 attacks like hypertext transfer protocol (HTTP) flood, Protocol attacks like synchronize (SYN) flood, volumetric attacks such as DNS amplification, and other forms of DDoS attacks with different techniques [5].

Over the years, researchers have proposed diverse techniques for intrusion detection, identification, and mitigation. These range from simple rule-based or signature-based systems to machine learning (ML) and deep neural (DL) network-based approaches. Their pursuit of the optimal technique is driven by objectives such as improving attack detection and classification accuracy, processing time efficiency, resource requirements, real-time applicability, interpretability, and minimizing false alarm rates (FARs). Researchers have explored various ML algorithms and DL methods to develop efficient attack, classifier models. The central focus often revolves around ML and/or DL-based anomaly detection for identifying and classifying malicious traffic flows. Anomaly detection is a broad research field with diverse proposed techniques and approaches. An anomaly can be defined as an observation that appears inconsistent with the rest of a dataset, exhibiting a specific pattern that displays significant changes in a network's normal traffic level [6]. Numerous papers have proposed different anomaly-based models by employing statistical, ML, and DL algorithms, such as Markov processes, statistical moments, multivariate distributions, Bayesian networks, fuzzy logic, decision trees, and neural networks [7-12].

Among those, DL has recently gained popularity, demonstrating to be an effective approach for cyberattack detection, classification, and mitigation across various network environments, including industrial control systems and IoT environments. DL has facilitated the development of anomaly-based detection models that require minimal human intervention and can detect zero-day attacks, unlike signature-based approaches. DL offers advantages such as increased detection rates, robustness to noise, high accuracy, improved system performance, computational efficiency, reduced false alarm rates (FAR), and decreased system complexity [13-25], but none of the prior work successfully achieved high performance with low system complexity.

In this paper, our objective is to develop a lightweight model that achieves higher accuracy, exhibits robust generalization across different datasets, and effectively reduces alarm rates in detecting and classifying DoS and DDoS attacks, while avoiding overfitting. To achieve this, we employ a feature transformation approach to convert the CSE-CIC-IDS2018 and NSL-KDD datasets into spectrogram-based images. The underlying hypothesis is that by transforming the original dataset into an image-based representation, we can enhance the patterns utilized by detector and classifier algorithms. Consequently, we could design a lightweight IDS model that yields improved performance metrics of significant importance.

The paper is structured as follows: Section I introduces the topic, Section II covers related works, Section III presents the proposed methodology and model, Section IV showcases results and includes experimental graphs, discussion, and comparison, and finally, Section V concludes with future directions.

2. Related Work

Numerous studies employed different ML algorithms for anomaly-based intrusion detection systems, including principal component analysis (PCA) based models as in reference [8] that

proposed a novel general form for distance calculation and a new PCA-based detection method for IoT networks. Similarly, reference [9] proposes a robust anomaly detection technique where the training phase is supported using the decision tree algorithm and a hybrid of cuckoo search optimization and k-means is cascaded for the detection. This model produced an improved detection rate and FAR among the other measurement matrices. Similarly, reference [10] proposes ML-based models to overcome the vulnerabilities of cloud computing networks from attacks such as DDoS cyberattacks. In this study, the authors proposed an intrusion detection system that combines fuzzy c-means clustering (FCM) and support vector machines (SVM) to significantly improve timely detection accuracies in a cloud computing environment.

Other ML algorithms are also effective in detecting and classifying several cyberattacks in various environments. Reference [11] presents supervised ML algorithms, which include k-nearest neighbors (KNN), logistic regression (LR), SVM, multilayer perceptron (MLP), decision tree (DT), and random forest (RF), for detection and classification in IoT security. In this study, the authors successfully detected and classified specific attacks including DDoS, DoS, reconnaissance, and information theft in IoT networks. As SVM is among the frequently used ML algorithms in anomaly detection for network security, the study in [12] proposes a one-class SVM for anomaly detection. Despite its computational expensiveness and significant memory requirement, this algorithm is excellent at capturing traffic patterns and malicious anomalies in IoT environments.

Widely used DL approaches for handling complex and high-dimensional data include restricted Boltzmann machines, deep belief networks, feed-forward neural networks, deep neural networks, recurrent neural networks, CNNs, and deep auto-encoders, among others, as identified in various survey papers [8]. DL is suitable for network various environments, such as IoT. The survey in reference [9] details the application of DL approaches in IoT environments with a detailed review of DL models that have been recently proposed for IoT intrusion detection. Based on [9], DL solutions are classified comprehensively based on the application of DL for IoT cybersecurity as effective IoT intrusion detection solutions. Numerous DL-based intrusion detection studies face challenges owing to the shortcomings of publicly available datasets, as highlighted in various survey works [10]. These datasets often suffer from poor representation, outdated information, high data redundancy, unrealistic simulation, limited traffic diversity, and a lack of generalized modern traffic data. The scarcity of high-quality datasets has prompted researchers to develop diverse DL-based solutions specifically designed for intrusion detection systems.

To address zero-day cyberattacks in IoT infrastructure and reduce FARs, [11] proposes a DNN-based IDS using mutual information (MI) for feature dimension reduction. This achieves high accuracy and a low FAR with reduced network complexity. Reference [12] presents a DL-based IDS against BotNet attacks in the IoT, utilizing CNNs to detect popular Botnet attacks and outperforming RNN-based systems. Reference [13] introduces a customized DL approach for detecting and classifying IoT-based cyberattacks, such as DoS, DDoS, data gathering, and data theft, achieving high classification accuracy through feed-forward neural networks with embedding layers and transfer learning. Reference [14] trains and evaluates robust IDS using commonly used datasets KDD Cup 99 and NSL-KDD. This study employs CNN-LSTM neural networks for cross-layer feature fusion, capturing both temporal and global characteristics of intrusion information for enhanced detection capabilities. By connecting CNN and LSTM, the IDS achieves accurate intrusion identification by leveraging comprehensive features extracted from regional and periodic traffic signal characteristics.

Reference [15] introduces a bidirectional long short-term memory (BiLSTM) IDS to address high FARs and low detection accuracies in certain attack classes, specifically user-to-root (U2R) and remote-to-local (R2L) attacks in the NSL-KDD dataset. The proposed solution outperforms LSTM-based IDSs, achieving better accuracies and significantly reduced FARs for U2R and R2L classes. In reference [16], a self-adaptive IDS is proposed using a DL-based model that dynamically adjusts the network structure for different attack types, enabling detection without altering the entire IDS structure. This system, based on an improved genetic algorithm (GA) and deep belief network (DBN), achieves high detection and recognition rates with a compact structure and reduced neural network complexity. Reference [17] presents an anomaly detection-based approach that addresses model

 doi:10.20944/preprints202401.1875.v1

overfitting using temporal CNNs and U-Net networks for attack classification with generalization characteristics. This study evaluates the proposed model on both the old KDD99 dataset and the modern large-scale CSE-CIC-IDS2018 dataset, while mitigating data imbalance challenges using the focal loss function. The results demonstrate the model's generalization ability and effectiveness in handling severe class imbalance.

The utilization of deep learning-based generative models specifically generative adversarial networks (GANs) and variational autoencoders (VAEs), is emphasized in reference [18]. These models demonstrate remarkable performance in generating realistic dataset content for augmenting existing datasets and addressing class imbalance issues. Reference [19] leveraged generative networks to build an anomaly-based IDS with reduced FARs and improved generalization. The proposed weakly supervised model utilizes GANs to generate malicious samples during training, leading to improved detection rates and lower FARs.

We have managed to find only a handful of related works in the research area where a given dataset is transformed into another format to produce better models. One work where the CSE-CIC-IDS2018 as well as the KDD datasets are transformed into their respective image datasets to construct an improved DL-based intrusion detection model for DoS attacks is presented in [18]. In this work, a CNN model is proposed, wherein an input image is prepared by rearranging the dataset's features into a matrix-like representation. For both datasets used in their work, they have generated colored and grayscale image datasets by rearranging the features in both datasets and color coding them to have image pixels of 13x9x1 or 3 (that is, 1 or 3 representing the color channels) and 13x6x1 or 3 for both KDD and CSE-CIC-IDS2018 datasets, respectively. After that several CNN models with different hyperparameters were tested for best performance based on the prepared image dataset. Though the authors reported excellent model accuracies and performance for multi-class classification, generalizing their proposed method to avoid overfitting remains a challenge.

A lightweight IDS was reported in [19] where raw traffic is converted into image data. In this study, the proposed method improves computational efficiency, but the reported experimental results demonstrate that the detection accuracy is considerably low. In contrast, in reference [20] the authors proposed a system that takes grayscale 2D-image datasets as input which are prepared from a few packets of captured raw traffic data. By creating the pattern for the raw traffic data, the authors proposed an IDS model that consists of CNNs and AEs for auto-profiling the traffic patterns and filtering abnormal traffic, and they have reported high classification accuracy and low false alarms in their experimental results. While their unsupervised proposed DL model achieved high accuracy, the FAR is substantially high. Besides the design model consists of layers of CNN cascaded with an auto-profiling auto-encoder. This adds a considerable amount to the complexity of the systems, hence a heavyweight network.

Another study, which is closely related to our method, is published in reference [21]. Expecting to reduce the high FARs observed in many proposed IDSs, the authors of [21] proposed a network IDS framework using a deep CNN that uses network spectrogram images generated using the short-time Fourier transform consuming the CSE-CIC-IDS2017 network dataset. They successfully converted the dataset into images and attempted to reduce the FAR to approximately 1.033% while they managed to achieve an average of 98.758% accuracy for the multiclass classification. They have also used a two-layer CNN and two layers of 128 neurons to build their fully connected module, which results in numerous learnable parameters, which in turn increases system complexity.

In this study, we aim to reduce the FAR and increase the accuracy, specificity, and sensitivity by proposing a light CNN-based IDS model capable of detecting and classifying DoS and DDoS cyberattacks from a normal network flow.

3. Proposed Methodology and Model

Our approach is to transform a publicly available dataset into a more convenient format for DL networks to ensure that a high-performance lightweight design is feasible. Using our proposed methodology, we have converted publicly available datasets into an image format by taking the spectrogram based on the short-term Fourier transform (STFT) technique. In signal analysis, this

 doi:10.20944/preprints202401.1875.v1

method is used for numerous purposes such as observing the frequency and time localizations, mainly when it is represented in a waterfall display. In this study, we used this concept to augment the distinctive pattern of processed network traffic signals. Based on this methodology, we successfully converted the CSE-CIC-IDS2018 dataset and proposed a lightweight CNN system to detect and classify network traffic flows. In addition, we have also acquired the NSL-KDD and converted it into an image dataset to demonstrate the generalization of the proposed methodology when applied to older traffic datasets and in a modern-day dataset.

As shown in Figure 1, we first collect the main dataset, CSE-CIC-IDS2018, from the Canadian Institute for Cybersecurity (CIC) and also acquire the NSL-KDD dataset to verify if our proposed methodology would generalize when tested in different network datasets. After obtaining both datasets, we let the dataset pass through a data preprocessing step for data cleaning, which includes removing unnecessary features, such as removing the time and date of recording of the traffic flow, handling missing, and infinite values. Subsequently, we retained only the normal traffic flows and the different types of DoS and DDoS attacks by excluding any other types of attacks. The retained attack in addition to the normal traffic flows included the Hulk, GoldenEye, Slowloris, and Slowhttptest as DoS attacks and Low Orbit Ion Canon HTTP (LOIC-HTTP), Low Orbit Ion Canon UDP (LOIC-UDP), High Orbit Ion Cannon (HOIC) as DDoS attacks. This dataset as explained in [22] has 80 features extracted using the CICFlowMeter-V3 tool. Similarly, after the preprocessing, 117 features were considered for the NSL-KDD dataset where only DoS attacks and normal flows were considered (DoS types of apache2, back, land, Neptune, mailbomb, pod, processtable, smurf, teardrop, udpstorm, and worm). After data cleaning, we converted each entry in the datasets into its spectrogram representation both in horizontal and vertical display as shown in Figure 2.

After obtaining these spectrogram representations we integrated them to form a pattern-augmented image representation of the dataset entry. The conversion was made possible using Eq. (1), by taking the absolute value of the division of the squared values of the vertical spectrogram points to the horizontal points, where α and β represent the vertical and horizontal spectrogram images in complex number representation, respectively. Because the values of the spectrogram are complex numbers, taking the absolute value is required to specify the color data as numeric or logical values during the image creation.

$$\gamma = \frac{|\alpha^2|}{|\beta^2|} \tag{1}$$

The integration of the two spectrogram image matrixes produces the 129X129 images as in Figure 3.

After obtaining the images for each data entry in all the datasets used, we converted these into grayscale images and stored them as an image dataset file to be used in the proposed DL networks. Since we have converted the more colorful images into less-colored images as shown in Figure 3, we could lower the file size and further decrease the file size by converting them into a grayscale format, as shown in Figure 4 for normal and attack traffic flows. The benefits of our proposed method played a central role in accelerating the detection and classification of our CNN model.

The proposed model is compared with an artificial neural network (ANN) model (Figure 5) to highlight the significant reduction in network weight achieved through our dataset conversion method, as demonstrated by the lightweight CNN model (Figure 6 & 7). As shown from the figures, both models are considerably shallow neural networks, because our converted image dataset is of high quality with an extremely low bias and variance within the data entries. The low bias-variance characteristics of the converted dataset help design the intended lightweight IDS. This is reflected in our experimental results, particularly the accuracy and loss curves during both the testing and training times, presented in the following section. One objective of this study is to propose a lightweight IDS model, and this is achieved owing to the cleanness of the prepared datasets by avoiding the usual requirement for deeper and more complex neural networks.

 doi:10.20944/preprints202401.1875.v1

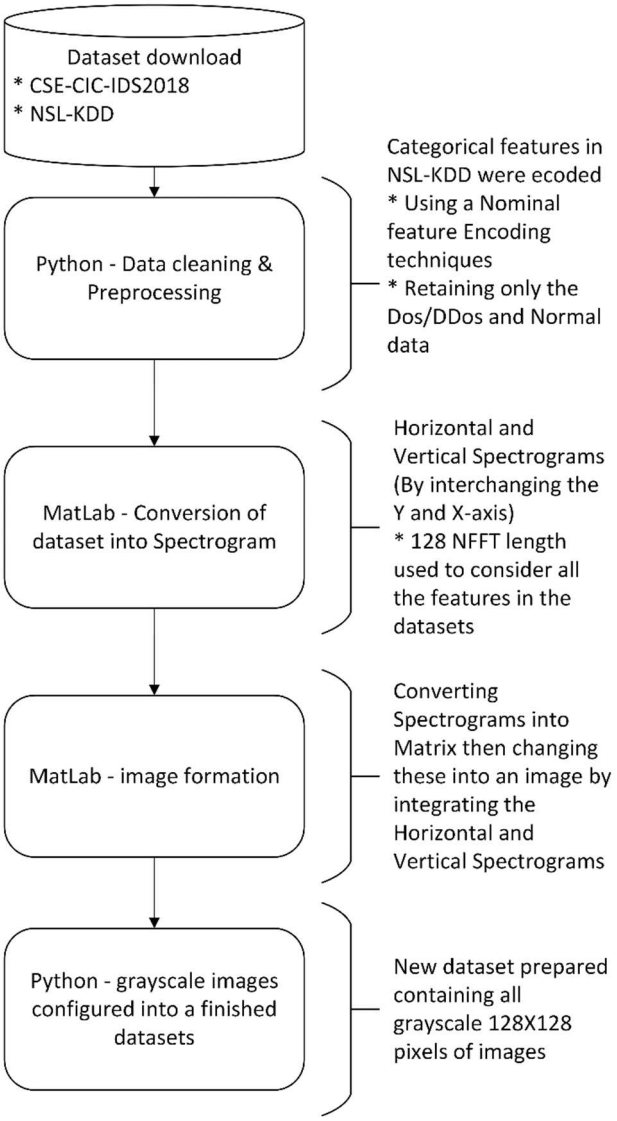

Figure 1. Proposed dataset to image conversion module.

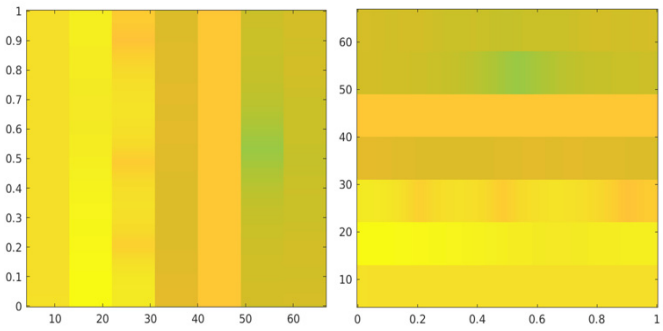

Figure 2. Spectrogram display horizontal, right and vertical, left.

 doi:10.20944/preprints202401.1875.v1

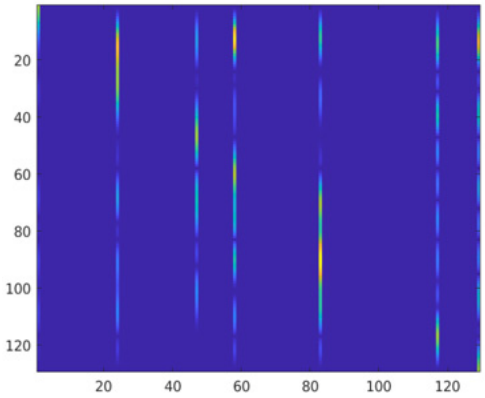

Figure 3. Pattern augmented image.

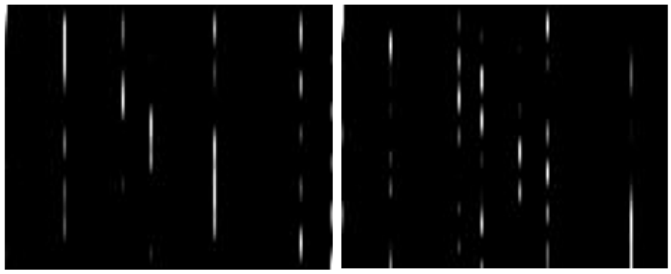

Figure 4. Traffic flow – Normal right and DDoS attack left.

As shown in Figure 5, the ANN model is an extremely shallow model of only a single hidden layer with 15 neurons and consists of input and output layers. In this model, the input is a 1x129x129 grayscale image and the input layer should effectively accept this input. As shown in this Figure 6, the input layer is set only for illustration purposes as it does not contain any learnable parameters, unlike the hidden and output layers. These two layers have numerous learnable parameters as shown in Table 3 (for the binary-class ANN model). In the case of multi-class classification, the output layer of the network consists of eight neurons, representing the assumed seven classes of attacks along with the normal class from the CSE-CIC-IDS2018 dataset. However, when performing binary-class classification in both the ANN and CNN networks, the number of neurons in the output layer is reduced to two.

We have set all the parameters and hyperparameters the same, including mainly the learning rate, the weight initialization method, epoch numbers, batch sizes, the optimizer technique, and types of activation functions. The ReLU activation function was used in all layers other than the output layers where a softmax activation function was used. The CNN binary and multi-class classification models are shown in Figures 6 and 7, respectively.

Both the multi and binary CNN models have almost similar structures and components except that they slightly differ in the hidden and output layers. Similar to the ANN model, the input for the CNN models is 1x129x129 pixels of an image. After the input layer, to process the image input, we utilize two convolutional layers, each consisting of eight filters with a size of 3x3. The input-output relation can be found in Table 4. The fully connected and output layers in the CNN models are similar to those of the corresponding binary-class and multi-class ANN models with similar parameters and hyperparameters. The relation among the inputs and outputs across the layers for the CNN models is depicted in Table 3. Tables 3 and 4 demonstrate the distinct system complexity between the ANN and CNN models, despite their similar performance in detecting and classifying cyberattacks. The ANN networks outperform the CNN networks, but they also generate a high number of learnable parameters when compared with the CNN counterpart models. Hence, the proposed CNN models are useful as lightweight intrusion detection and classification systems.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 January 2024 doi:10.20944/preprints202401.1875.v1

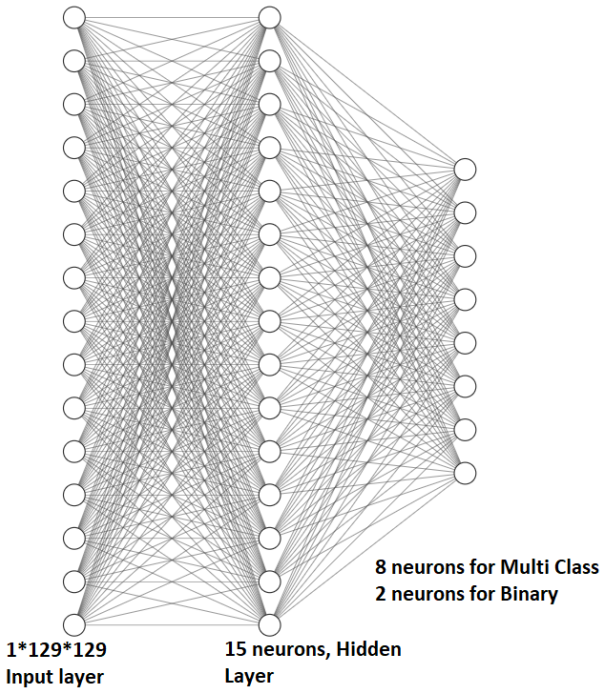

Figure 5. ANN model for binary and multi-class.

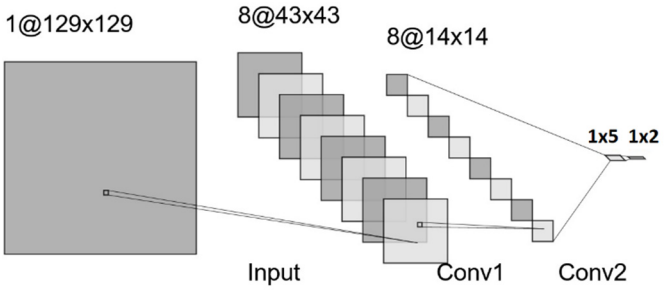

Figure 6. CNN- Binary-Class Model.

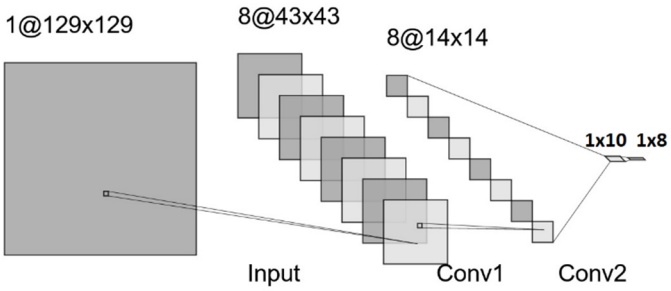

Figure 7. CNN-multi-class model.

4. Results and Discussion

In this section, we present the results and all the performance curves that we obtained during our experiment. To evaluate our proposed models, we have used the CSE-CIC-IDS2018 and NSL-KDD datasets separately as well as in combined form and this is summarized in Table 1.

Table 1. Evaluation Metrics for Binary-class classification.

Dataset	Acc(%)	Pre(%)	FNR	FPR
CIC-IDS2018	**ANN:99.55**	**ANN:99.55**	**FNN:0.0121**	**FNN:0.0027**

 doi:10.20944/preprints202401.1875.v1

NSL-KDD	**CNN:99.42**	**CNN:99.42**	**CNN:0.0089**	**CNN:0.0037**
	ANN:99.87	**ANN:99.87**	**FNN:0.0006**	**FNN:0.0009**
Combined dataset	**CNN:99.82**	**CNN:99.82**	**CNN:0.0008**	**CNN:0.0022**
	ANN:99.63	**ANN:99.63**	**FNN:0.0058**	**FNN:0.0026**
	CNN:99.37	**CNN:99.37**	**CNN:0.0097**	**CNN:0.0034**

[1] Acc: Accuracy, Pre: Precision, FPR: False Positive Rate and FNR: False Negative Rate.

Based on our experimental results, as displayed in Table 2, we observe the efficacy of our dataset conversion method reflected in the high accuracy, precision, recall, and F-1 scores for all the assumed datasets applied to the binary-class models. We generated these evaluation metrics as they are crucial for assessing our model.

Table 2. Sample Dataset.

Dataset	Normal Dataset	Attack	Training Vs Test split
CIC-IDS2018	**200,000**	**140, 000**	**90% Vs 10%**
NSL-KDD	**65, 000**	**110, 000**	**90% Vs 10%**
Combined Dataset	**240, 000**	**170, 000**	**95% Vs 5%**

As observed in Table 2, both models exhibit comparable performance, with the ANN slightly outperforming the CNN architecture. When we interpret this in conjunction with other conditions, such as the model complexity and loss/accuracy curves, we can deduce that CNN is a significantly robust model. By examining Tables 3 and 4, it becomes evident that the CNN model achieves commendable performance despite having a significantly smaller number of learnable parameters when compared to the ANN model. The CNN model only required 8,521 learnable parameters to achieve comparable performance to the ANN model for binary-class classification, while the ANN required approximately a quarter of a million learnable parameters. This shows the manner in which the image dataset prepared using our method aids the CNN model and achieves excellent performance without requiring deeper networks that would result in a large number of parameters.

Table 3. ANN Binary-class Model Summary.

Layer (type)	Output Shape	Parameters
Flatten	(None, 16641)	**0**
Dense	(None, 15)	249630
Dense	(None, 2)	**32**
Total params: 249,662		
Trainable params: 249,662		
Non-trainable params: 0		

Table 4. CNN Binary-class Model Summary.

Layer (type)	Output Shape	Parameters
Conv2D	(None, 43, 43, 8)	80

 doi:10.20944/preprints202401.1875.v1

Conv2D	(None, 14, 14, 8)	584
Flatten	(None, 1568)	0
Dense	(None, 5)	7845
Dense	(None, 2)	12
Total params: 8,521		
Trainable params: 8,521		
Non-trainable params: 0		

In addition, the CNN model in both the binary and multi-class models has achieved excellent performance in avoiding overfitting when compared with the ANN models. For this, we have generated the loss/accuracy curves during the training and evaluation sessions, as shown in Figures 8 and 9.

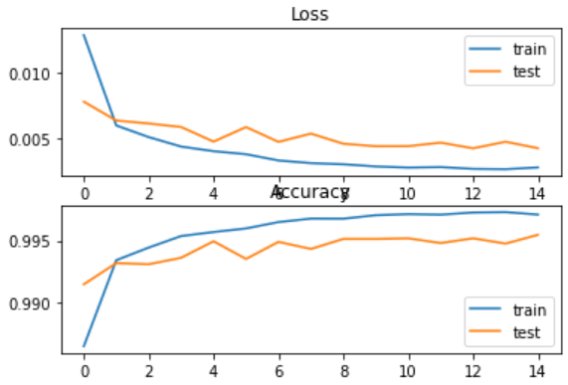

Figure 8. Binary-ANN loss/accuracy during training/testing.

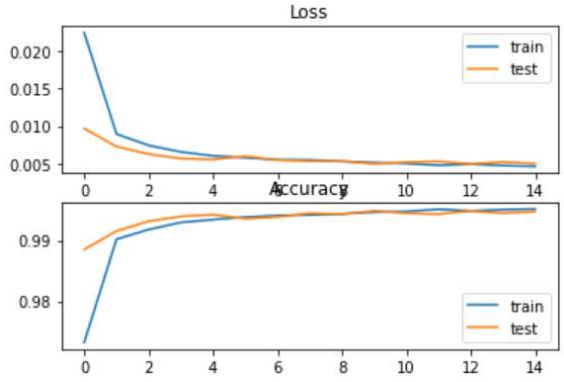

Figure 9. Binary - CNN loss/accuracy during training/testing.

As shown in Figures 8 and 9, the CNN model exhibits better generalization and avoids overfitting more effectively than the ANN model. It has been recorded that both models avoided overfitting, with ANN's accuracy as Train: 0.9975, Test: 0.9955 and CNN's accuracy as Train: 0.9956, Test: 0.9948. From this, it is clear that, based on the bias-variance analysis, the ANN model effectively minimizes the gap toward the maximum possible accuracy of 100%. However, the CNN model performs well in reducing the disparity between training and testing accuracies, which is a crucial indicator of its ability to handle unseen datasets compared to the ANN model. Thus far, the CNN model has achieved remarkable performance while avoiding overfitting and with few learnable parameters, which makes it an extremely lightweight model compared to the ANN model. Excellent performance was also observed from the ROC-AUC-generated curves. From Figures 10 and 11, it is evident that both models are effective in reducing FARs. The lightweight CNN model for the binary class achieved an FPR of 0.0089 and an FNR of 0.0037, while the ANN model for the binary class

 doi:10.20944/preprints202401.1875.v1

achieved an FPR of 0.0121 and an FNR of 0.0027. This shows the excellent specificity and sensitivity of the proposed models, and this is also an indication that the lightweight CNN model will likely outperform other proposed systems for real-time applications.

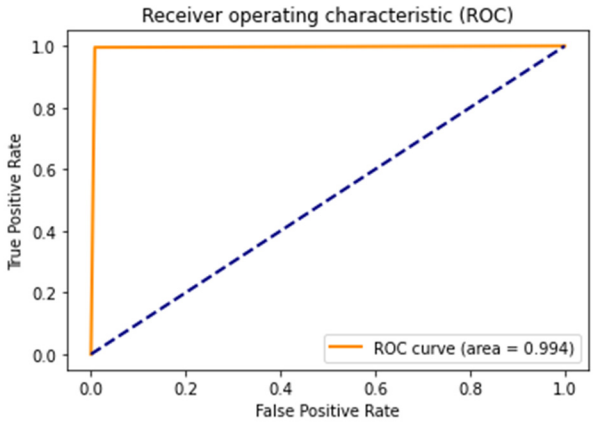

Figure 10. ROC-AUC curve for binary ANN.

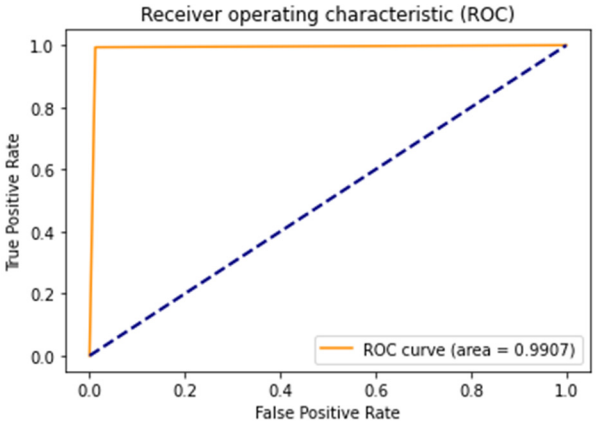

Figure 11. ROC-AUC curve for binary CNN.

To demonstrate the performance and generalization of our proposed approach and models on both legacy and modern network traffic flows, we have generated separate performance curves for the NSL-KDD dataset, which represents an older network dataset, and the CSE-CIC-IDS2018 dataset, which is a more up-to-date network dataset. Furthermore, we compared the proposed approach with different existing study articles; in terms of overall performance with achieved accuracy (Acc.), detection rate (DR), and FAR whenever included in the study articles. As shown in Table 5, there are several proposed techniques similar to ours, such as in terms of the model used and dataset selection. The comparison table shows how our proposed approach outperforms other techniques and approaches. In addition to achieving high accuracy with low FAR, our model presents a simple and lightweight model.

We have also tested our proposed system for multi-class classification. All the results demonstrated excellent performance both for the NSL-KDD and for multi-class classification for the CSE-CIC-IDS2018 datasets. Figure 12 shows the confusion matrix for multi-class classification (that is, classes Normal as normal, Hulk: Attack1, GoldenEye: Attack2, Slowloris: Attack3, Slowhttptest: Attack4, LOIC-HTTP: Attack5, LOIC-UDP: Attack6, and HOIC: Attack7) using the CNN lightweight model for multi-class classification. In addition, we have also evaluated our system's performance by testing it with a mixture of these two datasets, which belong to different network versions and time instances, serving as inputs to our proposed system. Subsequently, our system could effectively classify the mixed normal and attack signals without difficulties as observed in Figure 13 (Loss/accuracy) for binary-class classification (achieving Train: 0.9955, Test: 0.9937 accuracy).

 doi:10.20944/preprints202401.1875.v1

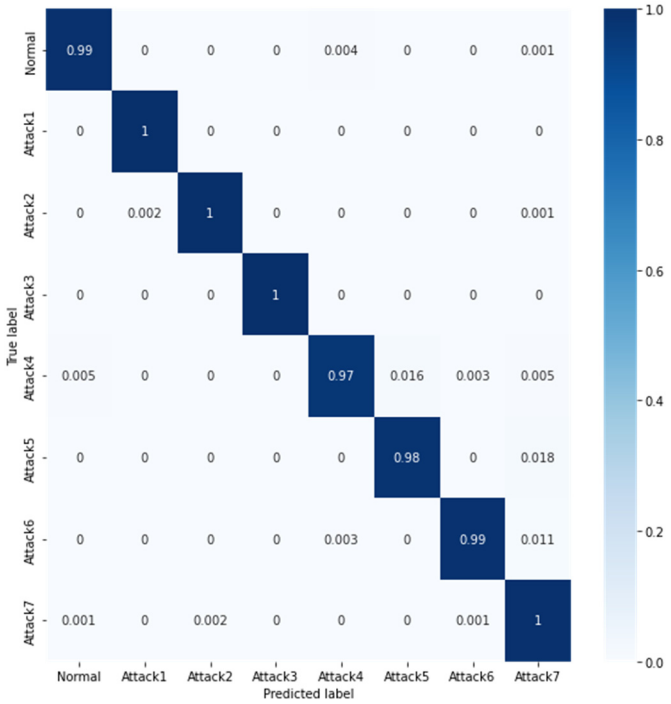

Figure 12. Confusion matrix for the multi-class CNN model.

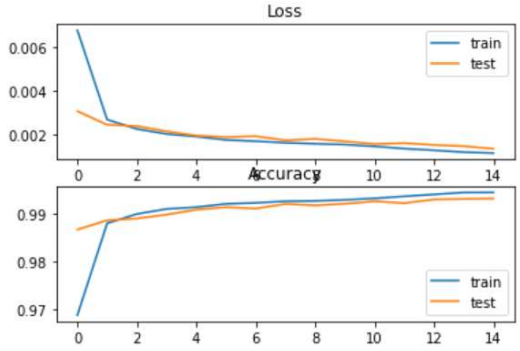

Figure 13. NSL-KDD+CSE-CIC-IDS2018 CNN binary classification loss/accuracy.

5. Conclusion

From our study, we have observed the effectiveness of applications of signal analysis techniques, such as STFT and other analysis tools, if used for image creation and pattern augmentation. We successfully converted the datasets CSE-CIC-IDS2018 and NSL-KDD into their corresponding image datasets for better performance in designing anomaly-based intrusion detection systems. As observed from the results of our experiment, the conversion of the datasets into their corresponding image datasets enabled us to design and propose an extremely lightweight intrusion detection model. In addition to being lightweight, the model exhibits excellent system performance, demonstrated by reduced FARs, high accuracy, and exceptional sensitivity and specificity measures. Furthermore, it avoids the overfitting problem, which is an undesirable phenomenon in DL methodologies.

For future work, we aim to extend our work to different types of network datasets with better models for IDS. Our next goal is to reduce the size of the images in the converted dataset without degrading its quality (that is, to lower the size of 129x129 pixels of the grayscale image). Furthermore, we plan to assess different ML and/or DL algorithms and networks for greater performance based on the prepared dataset. Furthermore, inspired by the promising results of our experiment, we envision applying this method to other application areas such as modulation identification. By leveraging the insights gained from this study, we can explore new avenues for improving performance and expanding the application of our approach.

 doi:10.20944/preprints202401.1875.v1

Table 5. Comparison table.

Article	Used dataset	Model	Evaluation Criteria
Novaes et al. [23]	CICDDoS2019	GANs	Acc: 94.38
Olaimat et al. [24]	CICIDS2017	GANs	Acc: 93.20
Huang et al. [25]	NSLKDD	IGANs	Acc: 84.45
Dlamini et al. [26]	NSLKDD	CGANs	F1 Score: 73.00
Han et al. [27]	KDD99	GANs Ensemble	Precision: 96.70
Ezeme et al. [28]	KDD99	cGANs	Acc: 85.63
Imtiaz U. et al. [29]	KDD99	cGANs	Precision: 99.05
E. min et al. [30]	NSLKDD, CICIDS2017	Autoencoder	DR: 99.00
W.Wang [31]	DARPA 1998 & ISCX 2012	CNN+LSTM	DR: 99.00, FAR: 0.02
M. Al-Qatf [32]	KDD99	Autoencoder + SVM	DR:95.00
Shone et al. [33]	KDD99 & NSLKDD	Asymmetric Autoencoder	Acc: 97.90, FAR: 2.10
Ludwig SA [34]	NSLKDD	Ensemble combining AE, DBN, DNN & ELM Algorithms	Acc: 92.49, FAR: 0.147
Yin et al. [35]	NSLKDD	RNN network and comparison with machine learning	Acc: 83.28, FAR: 0.07
A.Diro et al. [36]	NSLKDD	DNN with 4 hidden layers	Acc: 99.20
D. Aksu et al. [37]	CICIDS2017	DNN with 7 hidden layers	Acc: 98.00
T. Tang et al. [38]	NSLKDD	DNN with 3 hidden layers	Acc: 75.75
Andresini et al. [39]	CICIDS2017	Autoencoder and 1D CNN	Acc: 97.00
Roopak et al. [40]	CICIDS2017	CNN+LSTM	Acc: 96.20
A.S. Khan [21]	CICIDS2017	SDCNN	Acc: 98.76
Atefinia & Ahmadi [41]	CICIDS2018	Modular DNN	Acc: 100
Basnet et al. [10]	CICIDS2018	MLP	Acc: 99
Catillo et al. [42]	CICIDS2018	Deep Autoencoder	Acc: 99.20
Kim et al. [18]	CICIDS2018	CNN	Acc: 99.99
Lin et al. [43]	CICIDS2018	LSTM	Acc: 96.20
Our system	CICIDS2018, NSLKDD, & mix of these two	Lightweight CNN	Acc: 99.37, Pre: 99.37, FNR: 0.0034, FPR: 0.0097

Funding: This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2023-RS-2023-00255968) grant funded by the Korea government(MSIT).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: CSE-CIC-IDS. (2018). Datasets Research from Canadian Institute for Cybersecurity j UNB, 2018. [Online]. Available: https://www.unb.ca/cic/datasets/ids-2018.html. NSL-KDD Dataset. Available online: https://www.unb.ca/cic/datasets/nsl.html (accessed on 5 March 2023).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. "cybersecurity-statistics @ www.fortinet.com." [Online]. Available: https://www.fortinet.com/resources/cyberglossary/cybersecurity-statistics
2. C. Munroe, "IDC MarketScape IDC MarketScape : Worldwide Service Providers 2018 Vendor Assessment," no. January, pp. 1–6, 2018.
3. C. Cisco, "5 Steps to Protecting Your Organization from a DDoS Attack".
4. "tracking-cyber-operations-and-actors-russia-ukraine-war @ www.cfr.org."
5. "index @ www.cloudflare.com." [Online]. Available: https://www.cloudflare.com/
6. G. Fernandes, J. J. P. C. Rodrigues, L. F. Carvalho, J. F. Al-Muhtadi, and M. L. Proença, "A comprehensive survey on network anomaly detection," *Telecommun. Syst.*, vol. 70, no. 3, pp. 447–489, 2019, doi: 10.1007/s11235-018-0475-8.
7. D. Protic and M. Stankovic, "A hybrid model for anomaly-based intrusion detection in complex computer networks," *Proc. - 2020 21st Int. Arab Conf. Inf. Technol. ACIT 2020*, pp. 2160–2167, 2020, doi: 10.1109/ACIT50332.2020.9299965.
8. M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," *J. Inf. Secur. Appl.*, vol. 50, p. 102419, 2020, doi: 10.1016/j.jisa.2019.102419.
9. S. Tsimenidis, T. Lagkas, and K. Rantos, *Deep Learning in IoT Intrusion Detection*, vol. 30, no. 1. Springer US, 2022. doi: 10.1007/s10922-021-09621-9.
10. R. B. Basnet, R. Shash, C. Johnson, L. Walgren, and T. Doleck, "Towards detecting and classifying network intrusion traffic using deep learning frameworks," *J. Internet Serv. Inf. Secur.*, vol. 9, no. 4, pp. 1–17, 2019, doi: 10.22667/JISIS.2019.11.30.001.
11. Z. Ahmad *et al.*, "Anomaly detection using deep neural network for iot architecture," *Appl. Sci.*, vol. 11, no. 15, 2021, doi: 10.3390/app11157050.
12. I. Idrissi, M. Boukabous, M. Azizi, O. Moussaoui, and H. El Fadili, "Toward a deep learning-based intrusion detection system for iot against botnet attacks," *IAES Int. J. Artif. Intell.*, vol. 10, no. 1, pp. 110–120, 2021, doi: 10.11591/ijai.v10.i1.pp110-120.
13. M. Ge, N. F. Syed, X. Fu, Z. Baig, and A. Robles-Kelly, "Towards a deep learning-driven intrusion detection approach for Internet of Things," *Comput. Networks*, vol. 186, no. August 2020, p. 107784, 2021, doi: 10.1016/j.comnet.2020.107784.
14. R. Yao, N. Wang, Z. Liu, P. Chen, and X. Sheng, "Intrusion detection system in the advanced metering infrastructure: A cross-layer feature-fusion CNN-LSTM-based approach," *Sensors (Switzerland)*, vol. 21, no. 2, pp. 1–17, 2021, doi: 10.3390/s21020626.
15. Y. Imrana, Y. Xiang, L. Ali, and Z. Abdul-Rauf, "A bidirectional LSTM deep learning approach for intrusion detection," *Expert Syst. Appl.*, vol. 185, no. June, p. 115524, 2021, doi: 10.1016/j.eswa.2021.115524.
16. Y. Zhang, P. Li, and X. Wang, "Intrusion Detection for IoT Based on Improved Genetic Algorithm and Deep Belief Network," *IEEE Access*, vol. 7, pp. 31711–31722, 2019, doi: 10.1109/ACCESS.2019.2903723.
17. A. Mezina, R. Burget, and C. M. Travieso-Gonzalez, "Network Anomaly Detection With Temporal Convolutional Network and U-Net Model," *IEEE Access*, vol. 9, pp. 143608–143622, 2021, doi: 10.1109/access.2021.3121998.
18. J. Kim, J. Kim, H. Kim, M. Shim, and E. Choi, "CNN-based network intrusion detection against denial-of-service attacks," *Electron.*, vol. 9, no. 6, pp. 1–21, 2020, doi: 10.3390/electronics9060916.
19. V. Pham, E. Seo, and T. M. Chung, "Lightweight convolutional neural network based intrusion detection system," *J. Commun.*, vol. 15, no. 11, pp. 808–817, 2020, doi: 10.12720/jcm.15.11.808-817.
20. R. H. Hwang, M. C. Peng, C. W. Huang, P. C. Lin, and V. L. Nguyen, "An Unsupervised Deep Learning Model for Early Network Traffic Anomaly Detection," *IEEE Access*, vol. 8, pp. 30387–30399, 2020, doi: 10.1109/ACCESS.2020.2973023.
21. A. S. Khan, Z. Ahmad, J. Abdullah, and F. Ahmad, "A Spectrogram Image-Based Network Anomaly Detection System Using Deep Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 87079–87093, 2021, doi: 10.1109/ACCESS.2021.3088149.
22. "ef23092713b1e5491cfcc5bc918d5322c5751c28 @ registry.opendata.aws." [Online]. Available: https://registry.opendata.aws/cse-cic-ids2018/
23. M. P. Novaes, L. F. Carvalho, J. Lloret, and M. L. Proença, "Adversarial Deep Learning Approach Detection and Defense against DDoS Attacks in SDN Environments," *Futur. Gener. Comput. Syst.*, vol. 125, no. C, pp.

156–167, Dec. 2021, doi: 10.1016/j.future.2021.06.047.
24. M. Al Olaimat, D. Lee, Y. Kim, J. Kim, and J. Kim, "A Learning-based Data Augmentation for Network Anomaly Detection," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, 2020, pp. 1–10. doi: 10.1109/ICCCN49398.2020.9209598.
25. S. Huang and K. Lei, "IGAN-IDS: An imbalanced generative adversarial network towards intrusion detection system in ad-hoc networks," *Ad Hoc Networks*, vol. 105, p. 102177, 2020, doi: https://doi.org/10.1016/j.adhoc.2020.102177.
26. G. Dlamini and M. Fahim, "DGM: a data generative model to improve minority class presence in anomaly detection domain," *Neural Comput. Appl.*, vol. 33, no. 20, pp. 13635–13646, 2021, doi: 10.1007/s00521-021-05993-w.
27. X. Han, X. Chen, and L. Liu, "GAN Ensemble for Anomaly Detection," 2018.
28. O. M. Ezeme, Q. H. Mahmoud, and A. Azim, "Design and development of AD-CGAN: Conditional generative adversarial networks for anomaly detection," *IEEE Access*, vol. 8, pp. 177667–177681, 2020, doi: 10.1109/ACCESS.2020.3025530.
29. I. Ullah and Q. H. Mahmoud, "A Framework for Anomaly Detection in IoT Networks Using Conditional Generative Adversarial Networks," *IEEE Access*, vol. 9, pp. 165907–165931, 2021, doi: 10.1109/ACCESS.2021.3132127.
30. E. Min, J. Long, Q. Liu, J. Cui, Z. Cai, and J. Ma, "SU-IDS: A Semi-supervised and Unsupervised Framework for Network Intrusion Detection BT - Cloud Computing and Security," 2018, pp. 322–334.
31. W. Wang *et al.*, "HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection," *IEEE Access*, vol. 6, pp. 1792–1806, 2018.
32. M. Al-Qatf, Y. Lasheng, M. Al-habib, and K. Al-Sabahi, "Deep Learning Approach Combining Sparse Autoencoder With SVM for Network Intrusion Detection," *IEEE Access*, vol. 6, pp. 52843–52856, 2018.
33. N. Shone, T. N. Ngọc, V. D. Phai, and Q. Shi, "A Deep Learning Approach to Network Intrusion Detection," *IEEE Trans. Emerg. Top. Comput. Intell.*, vol. 2, pp. 41–50, 2018.
34. S. A. Ludwig, "Intrusion Detection of Multiple Attack Classes using a Deep Neural Net Ensemble," 2017.
35. C. Yin, Y. Zhu, J. Fei, and X.-Z. He, "A Deep Learning Approach for Intrusion Detection Using Recurrent Neural Networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.
36. A. A. Diro and N. K. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 82, pp. 761–768, 2017.
37. D. Aksu and M. A. Aydın, "Detecting Port Scan Attempts with Comparative Analysis of Deep Learning and Support Vector Machine Algorithms," *2018 Int. Congr. Big Data, Deep Learn. Fight. Cyber Terror.*, pp. 77–80, 2018.
38. T. A. Tang, L. Mhamdi, D. C. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep Recurrent Neural Network for Intrusion Detection in SDN-based Networks," *2018 4th IEEE Conf. Netw. Softwarization Work.*, pp. 202–206, 2018.
39. G. Andresini, A. Appice, N. Di Mauro, C. Loglisci, and D. Malerba, "Multi-Channel Deep Feature Learning for Intrusion Detection," *IEEE Access*, vol. 8, pp. 53346–53359, 2020.
40. M. Roopak, G.-Y. Tian, and J. A. Chambers, "Deep Learning Models for Cyber Security in IoT Networks," *2019 IEEE 9th Annu. Comput. Commun. Work. Conf.*, pp. 452–457, 2019.
41. R. Atefinia and M. Ahmadi, "Network intrusion detection using multi-architectural modular deep neural network," *Journal of Supercomputing*, vol. 77, no. 4. pp. 3571–3593, 2021. doi: 10.1007/s11227-020-03410-y.
42. M. Catillo, M. Rak, and U. Villano, "2L-ZED-IDS: A Two-Level Anomaly Detector for Multiple Attack Classes BT - Web, Artificial Intelligence and Network Applications," 2020, pp. 687–696.
43. P. Lin, K. Ye, and C. Xu, "Dynamic Network Anomaly Detection System by Using Deep Learning Techniques," 2019.