# Deep Learning Based Network Intrusion Detection System for Resource-Constrained Environments

Syed Rizvi[1], Mark Scanlon[2], Jimmy McGibney[1], and John Sheppard[1(✉)]

[1] South East Technological University, Waterford, Ireland
Syed.Rizvi@postgrad.wit.ie, {Jimmy.McGibney,John.Sheppard}@setu.ie
[2] School of Computer Science, University College Dublin, Dublin D04 V1W8, Ireland
mark.scanlon@ucd.ie

**Abstract.** Network intrusion detection systems (IDS) examine network packets and alert system administrators and investigators to low-level security violations. In large networks, these reports become unmanageable. To create flexible and effective intrusion detection systems for unpredictable attacks, there are several challenges to overcome. Much work has been done on the use of deep learning techniques in IDS; however, substantial computational resources and processing time are often required. In this paper, a 1D-Dilated Causal Neural Network (1D-DCNN) based IDS for binary classification is employed. The dilated convolution with a dilation rate of 2 is introduced to compensate the max pooling layer, preventing the information loss imposed by pooling and downsampling. The dilated convolution can also expand its receptive field to gather additional contextual data. To assess the efficacy of the suggested solution, experiments were conducted on two popular publicly available datasets, namely CIC-IDS2017 and CSE-CIC-IDS2018. Simulation outcomes show that the 1D-DCNN based method outperforms some existing deep learning approaches in terms of accuracy. The proposed model attained a high precision with malicious attack detection rate accuracy of 99.7% for CIC-IDS2017 and 99.98% for CSE-CIC-IDS2018.

**Keywords:** Intrusion Detection Systems · Dilated Causal Neural Network · Network Investigation

## 1 Introduction

Communication and networking systems are vulnerable to numerous intrusion threats due to the number of applications that operate on modern networks and their increasing size, complexity, and vulnerability. The investigation of modern networks results in massive volumes of information to be analyzed and classified by investigators [27], and this volume is set to be further compounded by the growing prevalence of Internet of Things (IoT) devices. To address these growing vulnerabilities, modern network systems must be capable of detecting and investigating network breaches in a more intelligent and effective way [10].


S. Goel et al. (Eds.): ICDF2C 2022, LNICST 508, pp. 355–367, 2023.
https://doi.org/10.1007/978-3-031-36574-4_21

Network-based intrusion detection systems (IDS) are an attack detection method that offers protection by continuously scanning network traffic for illegal and suspicious activity [15]. Network IDSs can be considered in two categories:

1. *Signature-based approach* – identifies attacks based on known signatures.
2. *Anomaly-based approach* – detects anomalous attacks based on artificial intelligence (AI) techniques.

Anomaly detection offers the benefit of detecting unknown attacks rather than relying on the signature profiles of known attacks. For this reason, much effort has been devoted to the development of anomaly detection IDSs based on machine learning and deep learning algorithms [17,28].

Machine learning techniques have relatively straightforward structures, whereas deep learning relies on an artificial neural network (ANN). Deep learning outperforms typical machine learning techniques when engaging with large sets of data [7]. Moreover, machine learning algorithms require human involvement for feature extraction to achieve better results. Manual feature engineering is unrealistic with multidimensional and large-scale data due to the fast growth in transmitted traffic [6]. Deep learning algorithms can acquire feature representations from datasets without human interaction to generate more effective results. Traditional machine learning models are often referred to as shallow models, due to their simple structure. Deep structure, which has numerous hidden layers, is one distinguishing feature of deep learning. However, due to the complex architecture, multi-layer deep learning models require a significant amount of computation and processing time [7].

To overcome this limitation, lightweight algorithms with minimal computational costs that can effectively address complex problems are being developed for resource-constrained environments. Lightweight deep learning techniques, such as 1D-Dilated Causal Neural Network (1D-DCNN), have been demonstrated to be effective for both classification and regression problems [19,22].

### 1.1 Contribution of This Work

In this paper, a 1D-DCNN based intrusion detection model has been employed to perform binary classification on two popular datasets used in the literature, namely, CIC-IDS2017 [23] and CSE-CIC-IDS2018 [4]. To the best of the authors' knowledge, this work is the first time that a 1D-DCNN model has been applied to network intrusion detection. The suggested approach identifies attacks with high accuracy through less complex architecture. Furthermore, numerous experiments have been performed to evaluate the behavior of our proposed model on recent datasets. The suggested model achieved 99.7% and 99.98% accuracy on the CIC-IDS2017 and CSE-CIC-IDS2018 datasets respectively.

## 2 Related Work

Deep learning has been used in network intrusion detection due to its scalability and automated feature development [7]. It has the capacity to extract better

representations from data and to analyze sophisticated traffic patterns from massive amounts of network traffic data.

Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) were employed to extract features and classify network data on CIC-IDS2017 by [24]. LSTM was used to find temporal information, while a CNN was employed to extract spatial characteristics. The authors additionally optimized weights on the training dataset to address the class imbalance problem. Compared to other machine learning methods, 1D-CNN is growing in popularity because of its superior feature extraction skills. The authors of [3] used a 1D-CNN on time-series data with 42 features for supervised learning. To reduce computational complexity, increase feature count, and improve output size, a max pooling layer is combined with the CNN layer. The performance of the proposed 1D-CNN is compared against Support Vector Machine (SVM), Random Forest (RF), and the hybrid architecture of 1D-CNN and LSTM for both balanced and unbalanced training datasets. Random oversampling was employed to solve the problem of imbalanced data after a comprehensive analysis were carried out utilizing the publicly available dataset.

Kim et al. [14], presented a CNN model for the CSE-CIC-IDS2018, by transforming data into images and evaluating its effectiveness with a Recurrent Neural Network (RNN) model. The experimental outcomes illustrate that CNN outperforms the RNN model on CSE-CIC-IDS2018. Lin et al. [16], suggested a method to increase network efficiency for anomaly detection by merging LSTM with Attention Mechanism (AM). The presented model was developed using the CSE-CIC-IDS2018 dataset, and analysis revealed that accuracy was 96.22%, detection rate was 15%, and recall rate was 96%. Kim et al. [14] compared the efficacy of a CNN model to an RNN model on the CSE-CIC-IDS2018 dataset. The experimental results show that CNN outperforms the RNN model.

## 3 Methodology

### 3.1 Dilated Causal Neural Network Architecture

Dilated convolutions, also known as atrous convolutions, are convolutions in which the filter is applied across a region that is longer than its length by omitting input values at a specific phase. By diluting the original filter with zeros, it becomes comparable to a convolution with a bigger filter, and in spite of this fact, it is far more effective. The architecture for dilated convolution is shown in Fig. 1. Using `x[n]` as the input, `D` as the dilation factor, `N` as the length, and `p_k` as the parameters, a dilated causal convolution is performed. `y[n]` is the resulting receptive field, that is:

$$y[n] = \sum_{k=0}^{N-1} p_k \cdot x[n - D \cdot k] \tag{1}$$

Even though `y` has only `N` parameters, it has a size of `D (N1) + 1`. Higher receptive fields require a minimal number of parameters while retaining causality,

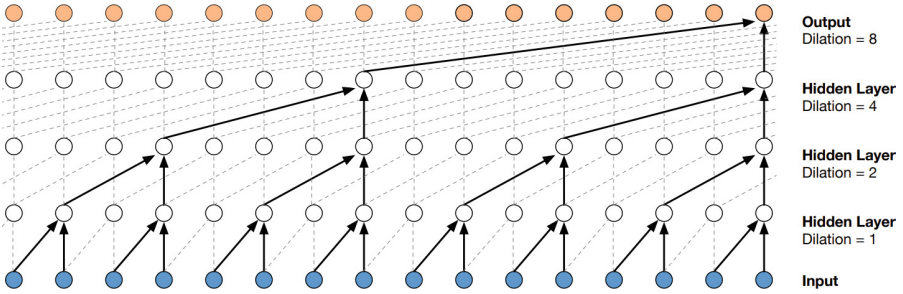



**Fig. 1.** Visualization of Dilated Causal Neural Network (Figure by [20]).

sampling rate, and using all inputs by stacking dilated causal convolutions with increasing dilation factors [11].

The manner by which causal convolution operates preserves the input's time sequence order, such that the model's projected outcomes are independent of upcoming time steps. Every convolutional layer in the dilated convolution has a set number of steps – where part of input values are skipped. The primary characteristic is that the network's receptive field does not require excessively large filters or convolutional layers, which significantly, or often exponentially, decrease the network's size.

Training such a model has the objective of minimizing the cost function by determining the weights at each layer of the network. Based on the error function, the optimizer progressively changes the weights. The output form of each layer's convolution is created by utilizing an activation function, which enables the model to identify a non-linear representation of the data.

### 3.2 Dataset Description

**CIC-IDS2017.** The Canadian Institute for Cybersecurity created the CIC-IDS2017 dataset in 2017 [23]. An experimental setup with malicious activity and victim networks was put up in order to construct the dataset. The dataset was collected over a 5-day period (Monday to Friday) using real-world data and is supplied as packet captures files and CSV files, with each row including generic information paired with characteristics derived by CICFlowMeter. An agent based on the `java-B-profile` system was built, which produced 80.32% benign traffic. Furthermore, several different types of network attacks are contained in the dataset including Heartbleed, botnet, SSH brute-forcing, web login brute-forcing, denial of service (DoS), distributed denial of service (DDoS), SQL injection, infiltration, and cross-site scripting. Through the experiment, more than 80 characteristics and 15 classes were recorded. It distinguishes from other network datasets in that it generates ultra-realistic network data and attack data based on actual users using distinct network profiles. This makes the CIC-IDS2017 a contemporary solution for supplementing intrusion detection systems. However, class imbalance is one of the dataset's major drawbacks.

**CSE-CIC-IDS2018.** The Canadian Communications Security Establishment (CSE) and CIC collaborated on an IDS dataset in 2018, the CSE-CIC-IDS2018 dataset [4]. The data set includes system logs and network traffic. It comprises 10 days of sub-datasets acquired on different days through executing 16 distinct sorts of attacks. The CICFlowMeter-V3 utility was utilized to build this dataset, which comprises around 80 different types of characteristics. This dataset contains numerous attack profiles that may be applied to network topologies and protocols in a comprehensive way in the field of intelligent security. The attacker's infrastructure consists of 50 computers, as opposed to the victim organization, which has 5 departments, 420 machines, and 30 servers. This dataset contains seven separate attack scenarios, including brute-force, DoS, DDoS, Heartbleed, botnet, web attacks, and penetrating an organization from within.

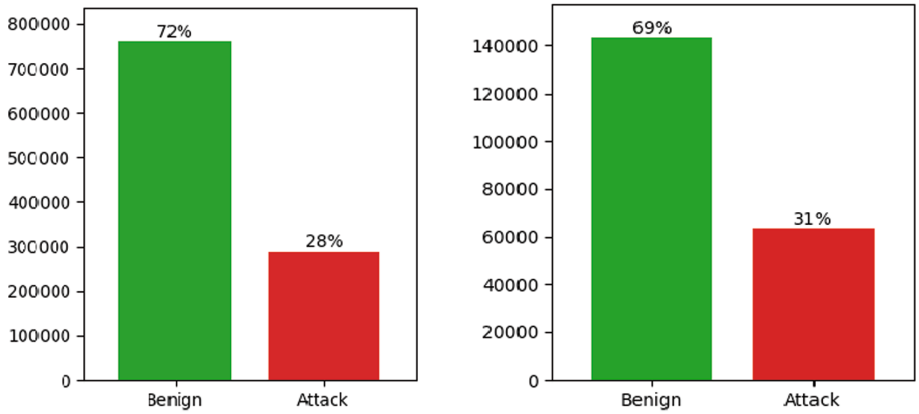
## 4 Experiments and Results

### 4.1 Experiments

**Simulation Environment.** The simulation environment was based on the Windows operating system. Its components comprised of an Intel Core i7-1165G7 with 8GB RAM. Python 3.9.7 was used for the development of the proposed DC-CNN model together with Keras, Tensorflow, and scikit-learn and data preprocessing was achieved with the `pandas` library.

**Exploratory Data Analysis.** Firstly, different days from the datasets have been concatenated to perform Exploratory Data Analysis (EDA). EDA is an essential process to achieve the desired output from an AI model. The dataset consists of 80 features, out of which 7 different irrelevant features, i.e., "Dst Port, Timestamp, FwdPSH Flags, Bwd PSH Flags, Fwd URG Flags, Bwd URG Flags, Flow Byts/s, Flow Pkts/s", have been removed to train our proposed model. The dataset required some pre-processing to deal with missing and duplicate values and to convert the labels to binary form (benign as 0 and attacks as 1), which could be utilized to perform binary classification. The outcomes of EDA on the CIC-IDS2017 and CSE-CIC-IDS2018 datasets are shown in Figs. 2a and 2b. After EDA, the datasets were split into the same ratio to train the model and evaluate the results.

**Dilated Causal Neural Network.** 1D-DCNN is capable of identifying and learning proper characteristics from a dataset. The proposed model is based on an input layer, hidden layers, and an output layer, each with its own set of parameters. Initially, investigation with number of filters: 8, 16, 64, and 128 with filter size: 4 and 8 along with incremental dilation rates with padding causal has been executed, but the results were insignificant. Other hyperparameters, including but not limited to activation functions, optimizer, and loss function, were crucial to achieving significant improvement. Initial experiments show that the number




(a) Distribution on CIC-IDS2017 (b) Distribution on CSE-CIC-IDS2018

**Fig. 2.** Benign and Attacks Distribution for Model Training.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 103294, 64)        19264

 conv1d_1 (Conv1D)           (None, 103294, 64)        16448

 conv1d_2 (Conv1D)           (None, 103294, 64)        16448

 conv1d_3 (Conv1D)           (None, 103294, 64)        16448

 conv1d_4 (Conv1D)           (None, 103294, 64)        16448

 conv1d_5 (Conv1D)           (None, 103294, 64)        16448

 conv1d_6 (Conv1D)           (None, 103294, 64)        16448

 dropout (Dropout)           (None, 103294, 64)        0

 dense (Dense)               (None, 103294, 1)         65

=================================================================
Total params: 118,017
Trainable params: 118,017
Non-trainable params: 0
_________________________________________________________________
```

**Fig. 3.** 1D-dilated casual neural network model summary.

of filters $> 64$ leads to lower accuracy and longer training time. Our optimum convolution operation parameters are: number of filters $= 64$, filter size $= 4$, and batch size $= 32$ using a blend of two distinct activation functions, namely *tanh* and *selu*. The model summary is shown in Fig. 3. The summary is descriptive

and provides details about the model's layers and their arrangement: each layer's output form and number of parameters (weights) and the total number of model parameters (weights). The dense layer with activation functions *sigmoid* has been utilized with a dropout value of 0.1. The performance of the model was also examined with several optimizers such as Adam, Stochastic Gradient Descent (SGD), and RMSProp along with different loss functions at distinct learning rates of 0.01, 0.001, and 0.0001. The learning curve was reviewed intensely during training to determine the hyperparameters of the optimizer. *Binary cross entropy* has been selected as the loss function. Besides, to identify the optimal performance, various momentum values for SGD have been evaluated.

Following experimentation with tuning various hyperparameters, the following parameters have been set to achieve the best performance: number of filters set to 64, filter size set to 4, and SGD as an optimizer with a learning rate of 0.01, momentum = 0.9, loss function = binary cross entropy with epochs set to 1800. The 1D-DCNN IDS model has been evaluated using the two datasets CIC-IDS2017 and CSE-CIC-ISD2018. It has been employed for binary classification to distinguish between attacks and normal traffic. All benign traffic was considered as 0, whereas all types of attacks are counted as 1. In the training phase, an early stopping technique was employed. After specific iterations, if the validation accuracy did not increase, model training was terminated and the hyperparameters were modified. This procedure persisted until the model training hyperparameters were all established.

### 4.2 Experimental Results

It was noticed that the proposed 1D-DCNN model performed considerably better on both datasets in terms of accuracy and that the model's training time is reasonably low – in the 1,000 to 1,800 epoch range. The outcomes of some extensive experiments using different hyperparameters are represented in Table 1. According to the results, the proposed model performed best with a combination of two different activation functions, namely *selu* and *tanh*, along with a specific number of filters and filter sizes, whereas the dilation rate was incremental in each experiment. The highest accuracy score achieved during the experiment on the CSE-CIC-IDS2018 dataset was 99.98%, as shown in Fig. 4.

Moreover, the performance of our proposed model on another IDS dataset, CIC-IDS2017, is shown in Fig. 5. The model achieved 99.7% accuracy on this dataset. Table 2 shows the comparison between the proposed 1D-DCNN based IDS and other existing work in terms of accuracy.

### 4.3 Discussion

Deep learning models have been a challenge to use on resource-constrained IoT devices because they require significant computational resources including memory, processor, and storage. As IoT devices are often easier to infiltrate than traditional computing systems, they are increasingly being targeted by malware-based attackers. This is caused by a variety of factors, including but not limited

**Table 1.** Results of hyperparameters tuning of the proposed model

| Exp No | No of filters | Filter size | Activation function | Loss | Accuracy | Wall Time |
|---|---|---|---|---|---|---|
| 1 | **8** | 4 | relu | 0.305 | 66.91 | 6 min 35 s |
| 2 | **16** | 4 | relu | 0.307 | 69.27 | 9 min 9 s |
| 3 | **32** | 4 | relu | 0.307 | 69.36 | 15 min 56 s |
| 4 | **64** | 4 | relu | 0.307 | 69.27 | 39 min 7 s |
| 5 | **8** | 8 | relu | 0.32 | 68.90 | 7 min 55 s |
| 6 | **16** | 8 | relu | 0.307 | 69.45 | 6 min 29 s |
| 7 | **32** | 8 | relu | 0.306 | 69.26 | 12 min 3 s |
| 8 | **64** | 8 | relu | 0.307 | 69.27 | 26 min 37 s |
| 9 | **16** | 4 | relu/selu | 0.307 | 69.28 | 5 min 21 s |
| 10 | **32** | 4 | relu/selu | 0.306 | 69.40 | 16 min 40 s |
| 11 | **64** | 4 | relu/selu | 0.307 | 69.28 | 18 min 32 s |
| 12 | **16** | 8 | relu/selu | 0.307 | 69.27 | 7 min 9 s |
| 13 | **32** | 8 | relu/selu | 0.307 | 69.27 | 12 min 20 s |
| 14 | **8** | 4 | tanh/selu | 0.305 | 69.50 | 4 min 48 s |
| 15 | **16** | 4 | tanh/selu | 0.075 | 93.76 | 6 min 54 s |
| 16 | **32** | 4 | tanh/selu | 0.073 | 94.02 | 16 min 48 s |
| 17 | **64** | 4 | tanh/selu | 0.067 | 94.25 | 29 min 48 s |
| 18 | **16** | 8 | tanh/selu | 0.303 | 69.72 | 11 min 48 s |
| 19 | **32** | 8 | tanh/selu | 0.297 | 70.38 | 20 min 50 s |
| 20 | **64** | 8 | tanh/selu | 0.109 | 92.61 | 39 min 24 s |

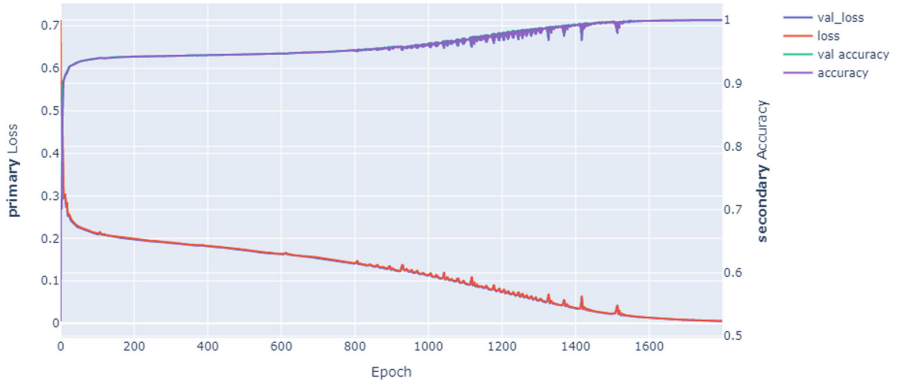



**Fig. 4.** Accuracy Progression During Learning on CSE-CIC-IDS2018 dataset

to the presence of legacy devices with no security upgrades, a low emphasis assigned to security throughout the development cycle, and insufficient login credentials (Table 3).

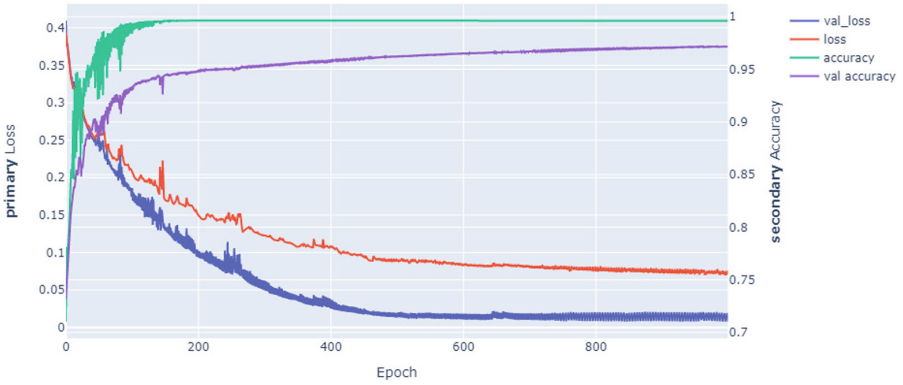



**Fig. 5.** Accuracy Progression During Learning on CIC-IDS2017 dataset

**Table 2.** Comparison of our proposed model with existing approaches

| Dataset | Reference | Accuracy |
|---|---|---|
| CSE-CIC-IDS2018 | Farhan et al. [9] | 90.25% |
| | Peng et al. [26] | 95% |
| | Lin et al. [16] | 95% |
| | Kim et al. [14] | 96% |
| | Khan [12] | 97.75% |
| | Emeç and Özcanhan [8] | 98.78% |
| | **Our model** | 99.98% |
| CIC-IDS2017 | Kim et al. [13] | 93% |
| | Varanasi and Razia [25] | 99% |
| | Asad et al. [2] | 98% |
| | Pekta and Acarman [21] | 99.09% |
| | Ma et al. [18] | 99.62% |
| | Alsyaibani et al. [1] | 98.34% |
| | **Our model** | 99.70% |

**Table 3.** Confusion Matrix: 1D-DCNN on CSE-CIC-IDS2018

| | Predicted Normal | Predicted Attack |
|---|---|---|
| Actual Normal | 71,550 | 0 |
| Actual Attack | 16 | 31,728 |

IoT devices are interesting to attackers for a myriad of reasons such as creating a botnet for DDoS attacks, mining cryptocurrencies, stealing private information, sabotage, or as a springboard for potential attacks and lateral movement through the network. According to a Kaspersky report [5], over 1.5 billion IoT

device breaches occurred in the first half of 2021, which was more than double the equivalent 2020 figure.

The following are key tangible benefits of the proposed model over existing classification models:

- In the convolutional layers of the network, a dilated convolution with dilation rate of 2 was introduced to compensate the max pooling layer, preventing the information loss imposed by pooling and down-sampling. The dilated convolution can also expand its receptive field to gather additional contextual data.
- The architecture of 1D-DCNN facilitates the gathering of additional contextual data, which helps to reduce the false alarm rate.
- The depthwise separable convolution used to decrease computational complexity and increase computational efficiency, allowing the model to successfully learn the representative features of datasets.
- 1D-DCNN are often more efficient and less costly to train compared to RNN and CNN models.

The approach outlined as part of this paper also has some limitations:

- The suggested IDS approach only investigated binary classification, i.e., normal and attack traffic.
- The lower threshold of computational power needed to successfully run the developed models on resource constrained environments has not yet been evaluated.
- The 1D convolutional network, without dilated convolutions, is a natural predictor. The prediction should not be dismissed with a peek into the future that occurs when convolutions are not causal. Measuring the proposed models' applicability and effectiveness versus other novel techniques is outside the scope of this paper.

While considering the above benefits and limitations, the application of the proposed technique still has significant merit due to its lightweight nature for resource constrained environments. As a result of these benefits, the lightweight 1D-DCNN outlined as part of this paper can be deployed on IoT/edge devices for intrusion detection with higher accuracy and precision.

## 5 Conclusion

This paper introduces an IDS model based on a 1D-dilated convolutional neural network approach for network attack detection and investigation. A dilated convolution, as compared to standard convolution, can enhance the receptive field without changing network parameters or reducing network capacity. 1D-DCNN has already proven its effectiveness in other application areas. The CIC-IDS2017 and CSE-CIC-IDS2018 datasets were used to train and test the suggested approach. The effectiveness of DCNNs to extract discriminative and efficient features

from the data has been demonstrated through experimental studies. The method suggested here is more reliable for IDS when compared to other state-of-the-art techniques. The depthwise separable convolution was used to decrease computational complexity and increase computational efficiency, allowing the model to successfully learn the representative features of the datasets.

The increasing prevalence of attackers targeting IoT/edge devices and networks is correlated with the increasing adoption of this new technology. The lightweight nature of 1D-DCNN facilitates its deployment on IoT/edge devices. Enabling the devices themselves to categorize their network traffic locally can contribute to the protection of IoT networks from sophisticated attacks.

### 5.1 Future Work

Our experiments achieved promising outcomes on both the CIC-IDS2017 and CSE-CIC-IDS2018 datasets in terms of accuracy, i.e., 99.70% and 99.98% respectively, with high precision. However, there is potential for improvement. The effectiveness of the model in terms of computational cost and performance needs to be compared with other state-of-the-art deep learning approaches, such as CNN and RNN, and with traditional machine learning algorithms such as Random Forest (RF) and Support Vector Machine (SVM). Hyperparameter tuning combined with dimensionality reduction and/or attention mechanism techniques may be able to further improve results and is worthy of future exploration. The model also needs to explore multi-classification, to investigate the detection rate of different attacks instead of the detection of anomalies. Moreover, other commonly used datasets in the literature, e.g., NSL-KDD and UNSW-NB15, can also be considered to further assess the performance of the proposed model against the state-of-the-art.

## References

1. Alsyaibani, O.M.A., Utami, E., Hartanto, A.D.: An intrusion detection system model based on bidirectional LSTM. In: 2021 3rd International Conference on Cybernetics and Intelligent System (ICORIS), pp. 1–6 (2021). https://doi.org/10.1109/ICORIS52787.2021.9649612
2. Asad, M., Asim, M., Javed, T., Beg, M.O., Mujtaba, H., Abbas, S.: DeepDetect: detection of distributed denial of service attacks using deep learning. Comput. J. **63**(7), 983–994 (2020)
3. Azizjon, M., Jumabek, A., Kim, W.: 1D CNN based network intrusion detection with normalization on imbalanced data. In: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pp. 218–224 (2020). https://doi.org/10.1109/ICAIIC48513.2020.9064976
4. Canadian Institute for Cybersecurity: A Realistic Cyber Defense Dataset (CSE-CIC-IDS2018). https://registry.opendata.aws/cse-cic-ids2018/. Accessed 02 June 2022
5. Cyrus, C.: IoT cyberattacks escalate in 2021, according to Kaspersky (2021). https://www.iotworldtoday.com/2021/09/17/iot-cyberattacks-escalate-in-2021-according-to-kaspersky/

6. Dib, M., Torabi, S., Bou-Harb, E., Assi, C.: A multi-dimensional deep learning framework for IoT malware classification and family attribution. IEEE Trans. Netw. Serv. Manage. **18**(2), 1165–1177 (2021)
7. Du, X., et al.: SoK: exploring the state of the art and the future potential of artificial intelligence in digital forensic investigation. In: Proceedings of the 15th International Conference on Availability, Reliability and Security. ARES 2020, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3407023.3407068
8. Emeç, M., Özcanhan, M.H.: A hybrid deep learning approach for intrusion detection in IoT networks. Adv. Electr. Comput. Eng. **22**(1), 3–12 (2022)
9. Farhan, R.I., Maolood, A.T., Hassan, N.F.: Optimized deep learning with binary PSO for intrusion detection on CSE-CIC-IDS2018 dataset. J. Al-Qadisiyah Comput. Sci. Math. **12**(3), 16 (2020)
10. Friday, K., Bou-Harb, E., Crichigno, J., Scanlon, M., Beebe, N.: On Offloading Network Forensic Analytics to Programmable Data Plane Switches. World Scientific Publishing, Singapore (2021)
11. Harell, A., Makonin, S., Bajić, I.V.: Wavenilm: a causal neural network for power disaggregation from the complex power signal. In: ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8335–8339 (2019). https://doi.org/10.1109/ICASSP.2019.8682543
12. Khan, M.A.: HCRNNIDS: hybrid convolutional recurrent neural network-based network intrusion detection system. Processes **9**(5), 834 (2021)
13. Kim, A., Park, M., Lee, D.H.: AI-IDS: application of deep learning to real-time web intrusion detection. IEEE Access **8**, 70245–70261 (2020). https://doi.org/10.1109/ACCESS.2020.2986882
14. Kim, J., Shin, Y., Choi, E.: An intrusion detection model based on a convolutional neural network. J. Multimed. Inf. Syst. **6**(4), 165–172 (2019)
15. Li, J., Qu, Y., Chao, F., Shum, H.P., Ho, E.S., Yang, L.: Machine learning algorithms for network intrusion detection. AI in Cybersecur., 151–179 (2019)
16. Lin, P., Ye, K., Xu, C.-Z.: Dynamic network anomaly detection system by using deep learning techniques. In: Da Silva, D., Wang, Q., Zhang, L.-J. (eds.) CLOUD 2019. LNCS, vol. 11513, pp. 161–176. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-23502-4_12
17. Liu, H., Lang, B.: Machine learning and deep learning methods for intrusion detection systems: a survey. Appl. Sci. **9**(20), 4396 (2019). https://doi.org/10.3390/app9204396
18. Ma, C., Du, X., Cao, L.: Analysis of multi-types of flow features based on hybrid neural network for improving network anomaly detection. IEEE Access **7**, 148363–148380 (2019). https://doi.org/10.1109/ACCESS.2019.2946708
19. Ma, H., Chen, C., Zhu, Q., Yuan, H., Chen, L., Shu, M.: An ECG signal classification method based on dilated causal convolution. Comput. Math. Methods Med. (2021)
20. van den Oord, A., et al.: WaveNet: a generative model for raw audio. In: Arxiv (2016). https://arxiv.org/abs/1609.03499
21. Pektaş, A., Acarman, T.: A deep learning method to detect network intrusion through flow-based features. Int. J. Netw. Manage. **29**(3), e2050 (2019) https://doi.org/10.1002/nem.2050. https://onlinelibrary.wiley.com/doi/abs/10.1002/nem.2050
22. Rizvi, S.M., Syed, T., Qureshi, J.: Real-time forecasting of petrol retail using dilated causal CNNs. J. Ambient Intell. Humanized Comput. **13**(2), 989–1000 (2022)

23. Sharafaldin, I., Lashkari, A.H., Ghorbani, A.A.: Toward generating a new intrusion detection dataset and intrusion traffic characterization. In: Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP 2018), vol. 1, pp. 108–116 (2018). https://doi.org/10.5220/0006639801080116
24. Sun, P., et al.: DL-IDS: extracting features using CNN-LSTM hybrid network for intrusion detection system. Secur. Commun. Netw. (2020)
25. Varanasi, V.R., Razia, S.: CNN implementation for IDS. In: 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), pp. 970–975 (2021). https://doi.org/10.1109/ICAC3N53548.2021.9725426
26. Wei, P., Li, Y., Zhang, Z., Hu, T., Li, Z., Liu, D.: An optimization method for intrusion detection classification model based on deep belief network. IEEE Access **7**, 87593–87605 (2019). https://doi.org/10.1109/ACCESS.2019.2925828
27. van de Wiel, E., Scanlon, M., Le-Khac, N.A.: Enabling non-expert analysis of large volumes of intercepted network traffic. In: Peterson, G., Shenoi, S. (eds.) Advances in Digital Forensics XIV, pp. 183–197. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-99277-8_11
28. Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., Gao, M., Hou, H., Wang, C.: Machine learning and deep learning methods for cybersecurity. IEEE Access **6**, 35365–35381 (2018)