Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 18 No. 14 (2024) | OPEN ACCESS

https://doi.org/10.3991/ijim.v18i14.50343

PAPER

A Hybrid-Transformer-Based Cyber-Attack Detection in IoT Networks

Imad Tareq Al-Halboosi[1](✉), Bassant Mohamed Elbagoury[1,2], Salsabil Amin El-Regaily[1], El-Sayed M. El-Horbaty[1]

[1]Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt

[2]Faculty of Computer Science and Computer Engineering, King Salman International University, El-Tor, Egypt

cs.20.19@grad.uotechnology.edu.iq

ABSTRACT

The concept of the Internet of Things (IoT) is significant in today's world and opens up new opportunities for several organizations. IoT solutions are proliferating in fields such as self-driving cars, smart homes, transportation, and healthcare, and new services are constantly being created. Over the previous decade, society has seen a significant expansion in IoT connectivity. In reality, IoT connectivity will expand in a variety of domains over the next few years. Various problems must be overcome to permit effective and secure operations. However, growing connections increase the potential for cyber-attacks since attackers can exploit the broad network of linked devices. Artificial intelligence (AI) detects and prevents cyber assaults by constantly developing and adjusting to new threats and weaknesses. In this study, we offer a novel cyber-detection model for IoT networks based on convolutional neural networks (CNN) transformers. The study aims to enhance the system's ability to identify and detect cyberattacks, new and sophisticated assaults, and its performance. The experimental study findings, using a new cybersecurity CICIoT2023 dataset, show that the CNN-Transformer model can detect IoT hazards with an overall accuracy of 99.49%. In identifying hazardous activity, MLP accuracy is 99.39%, while XGBoost-pipeline accuracy is 99.40%.

KEYWORDS

cyber-security, Internet of Things (IoT), deep learning, transformer, network

1 INTRODUCTION

[1][2] Currently, the Internet of Things (IoT) plays a crucial role in society, offering new capabilities to various industries. IoT projects in transportation and healthcare are becoming increasingly popular, along with industrial systems, self-driving vehicles, smart sensors, mechanical systems, terminals, mechanisms, and innovative applications [1, 2]. However, these systems are vulnerable to a range of cyberattacks and security vulnerabilities. Despite the benefits, several challenges need to be addressed to ensure effective and secure operations. Due to the vast amount of network traffic data in the IoT, the complex data characteristics, and the continuous

Al-Halboosi, I.T., Elbagoury, B.M., El-Regaily, S.A., El-Horbaty, E.-S.M. (2024). A Hybrid-Transformer-Based Cyber-Attack Detection in IoT Networks. *International Journal of Interactive Mobile Technologies (iJIM)*, 18(14), pp. 90–102. https://doi.org/10.3991/ijim.v18i14.50343

Article submitted 2024-03-30. Revision uploaded 2024-06-03. Final acceptance 2024-06-04.

emergence of new types of cyberattacks, intrusion detection methods based on statistical analysis or pattern matching often result in a high false-positive rate and low detection efficiency. Ensuring the safe and reliable operation of the network is a significant challenge [3]. The extensive scale of IoT networks introduces new complexities, such as data security, privacy concerns, and other issues. Ensuring privacy, security, and user satisfaction is crucial for the widespread adoption of IoT technologies. Additionally, IoT systems create new vulnerabilities for potential attacks due to the interconnected nature of the systems. Cybersecurity experts frequently highlight this aspect, emphasizing that IoT expands the attack surface available to hackers. With advancements in artificial intelligence (AI) technology, deep learning has become increasingly popular in developing "intelligent autonomous" IoT security systems due to its robust learning capabilities, flexibility, and portability [4, 5]. Machine learning techniques such as convolutional neural networks (CNN) can be utilized to automatically extract traffic attributes and identify traffic anomalies through categorization [6]. Temporal characteristics involve time series data between each traffic sample in the traffic sequence, and traffic anomalies are identified using time series analysis. Recurrent neural networks (RNN) are commonly used for time series problems, but their linear sequence structure leads to challenges with long-distance dependencies and limited parallel computing capabilities, restricting their use in real-time applications [7]. To overcome the limitations of RNN in sequence analysis tasks, Google introduced the Transformer model, which relies on the attention mechanism to understand the contribution of each input in the sequence to the final outcome through an internal self-attention process. This model considers global information [8]. Transformer has gained popularity in natural language processing [9], object recognition [10], and other fields due to its ability to capture long-distance features effectively and its parallel computing capabilities. To achieve intelligent and efficient identification of cyber-attack behaviors from network traffic data, we propose a cyber-attack detection model, CNN-Transformer, which combines CNN and transformer. The key contributions of this study are summarized as follows:

- We present a novel CNN-Transformer model for cyber-attack detection in IoT networks and compare it with the MLP and XGBoost pipelines.
- In experiments, we utilize the recently published extensive dataset, CICIoT2023, which contains a variety of threats and addresses a gap in the current dataset, to evaluate our model. We address the big data challenge using Spark.
- The CNN Transformer offers significantly better detection performance than other popular detection algorithms.
- We thoroughly evaluate our method with fresh datasets and performance metrics. The different experimental results illustrate our model's durability and efficacy. This research article is organized as follows: We examine related work in Section 2. Section 3 describes the proposed CNN-Transformer paradigm and shows the process of each component. In Section 4, we evaluate public cyber detection benchmark datasets and present the data pretreatment procedure. In Section 5, we present experimental results to confirm the significance of the model. In Section 6, the conclusion and future work are discussed.

2 RELATED WORKS

To identify and prevent cyber assaults on networks, researchers have proposed a variety of network intrusion detection systems. This section presents the current model approaches for detecting transformer-based attacks. The authors of [11] developed

a transformer-based model based on the DDoS assault dataset, CICDDoS2019, combining transformers with a CNN to identify DDoS attacks, with the maximum accuracy attained being 99.82%. Using the CIC-DDoS2019 and CICIDS2017 datasets, [12] proposes an intrusion detection system based on transformers that rebuild feature representations to achieve a balance between dimensionality reduction and feature retention, with a maximum accuracy of 98.58% and 98.45%, respectively. The authors of [13] present a transformer neural network-based intrusion detection system for IoT networks based on the MQTT-IoT-IDS2020 dataset, which achieves 99.9% accuracy. In [14], they present a transformer-based and generative adversarial network (GAN) model for cyber threat-hunting in 6G-enabled IoT networks, using the Edge-IIoT dataset, with an overall accuracy of 95%. In [15], the authors offer a transformer-based intrusion detection system for learning the behaviors and impacts of assaults in a diverse IoT environment. The approach uses a self-attention mechanism to acquire contextual embeddings for input network properties. Experiments using the ToN IoT dataset show an accuracy of 95.78% for multiple classifications and 97.95% for binary classification. The authors of [16] provide a transformer-based intrusion detection approach for analyzing the data features of intrusion behaviors in cloud security. Experimental findings utilizing the CIC-IDS 2018 dataset show that the model is 93% accurate. The authors of [17] offer the transformer-based autoencoder model for anomaly detection in IoT security systems. The model's performance on the DS2OS dataset is tested, with results showing a recall metric of 96.28%. The authors of [18] applied a transformer-based model for malware detection and categorization. The model was tested on the UNSW-NB15 and CIC-IOT23 datasets, with a focus on the payloads of UDP and TCP packets used as inputs. In the multi-classification exercise, UNSW-NB15 achieved an accuracy of 74.24%, whereas the CIC-IOT23 datasets achieved 69.25%. The authors of [19] present a model called BBO-CFAT, which combines the Biogeography-Based Optimization algorithm (BBO) for feature selection with an enhanced Transformer model for conserving context information and saving computational space. Experiments employing the NSL-KDD and CIC-IDS2017 datasets show 97.5% and 99.1% accuracy, respectively (see Table 1).

Table 1. Presents a summary of past investigations

Papers	Years	Dataset	Accuracy (%)	Description
Wang H. [11]	2021	CICDDoS2019	99.82	Developed a transformer-based model based on the DDoS assault dataset.
Wu Z. [12]	2022	CICIDS2017	98.45	Transformers are utilized in intrusion detection systems to create feature representations that balance dimensionality reduction with feature retention.
		CIC-DDoS2019	98.58	
Ullah S. [13]	2023	IDS2020	99.9	A transformer neural network-based intrusion detection system for MQTT-enabled Internet of Things networks.
Ferrag M. [14]	2023	Edge-IIoT	95	A transformer-based and GAN model for cyber threat-hunting in IoT networks.
Wang M. [15]	2023	ToN IoT	95.78	A Transformer-based intrusion detection system for learning the behaviors and impacts of assaults in an IoT environment.
Long Z. [16]	2024	CIC-IDS 2018	93	Transformer-based intrusion detection to analyze the data features of intrusion behaviors in cloud security.
Saghir A. [17]	2023	DS2OS	96.28	Transformer-based autoencoder technique for anomaly detection in IoT security systems.
Stein K. [18]	2024	UNSW-NB15	74.24	A transformer-based model for malware detection and categorization.
		CIC-IOT23	69.25	
Jiang T. [19]	2024	NSL-KDD	97.5	Present a model BBO-CFAT, which combines the BBO for feature selection with an enhanced Transformer model for conserving context information and saving computational space.
		CIC-IDS2017	99.1	

In this study, we utilize CNN-Transformer to detect modern assaults on IoT networks. The approach combines CNN and Transformer methodologies (the hybrid CNN-Transformer algorithm) to identify cyber-attacks in IoT networks using the new CI-CIoT2023 database.

3 PROPOSED MODEL

Transformer is a groundbreaking deep-learning approach that was initially developed to address sequence issues and long-term dependencies. It uses a self-attentive strategy to enable effective parallel processing, thereby enhancing the learning process. The transformer manages both the encoding and decoding processes. While encoding encodes one language, decoding calculates the likelihood of another language based on the past output. In the context of cyber-attack detection, this paper utilized a combination of CNN and transformer to create a hybrid model that surpasses traditional hybrid models. Figure 1 illustrates a structural schematic of the CNN-transformer model.

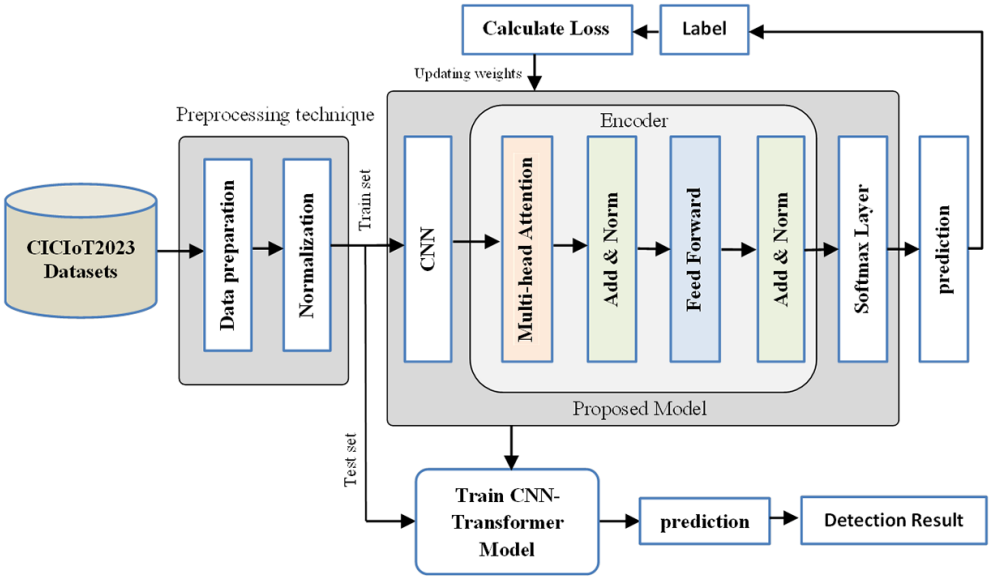

Fig. 1. CNN-transformer architecture

The input to our CNN-Transformer model is a batch of 46 features that are fed into the CNN block. These features represent the input, and the transformer block consists of three blocks. The output of the transformer block serves as the input to the Softmax activation function. The ultimate output consists of eight classes of attacks: DDoS, Recon, DoS, Benign, Web-based, Spoofing, Brute Force, and Mirai.

The suggested CNN processes the input shape, as shown in Figure 2. The CNN layer consists of three batch normalizations, three 1D CNNs with filters (64, 128, 256, and kernel 3), three max pooling layers, flattening, three dense layers (with units 256, 128, and 64), and three dropout layers. Our model utilizes the Selu function, which is defined as:

$$\int(x) = \begin{cases} \lambda\chi, \text{ if } x > 0 \\ \lambda\alpha(e^{x} - 1), \text{ if } x \leq 0 \end{cases} \quad (1)$$

Where λ and α constants with values: $\lambda \approx 1.0505$ and $\alpha \approx 1.6732$.

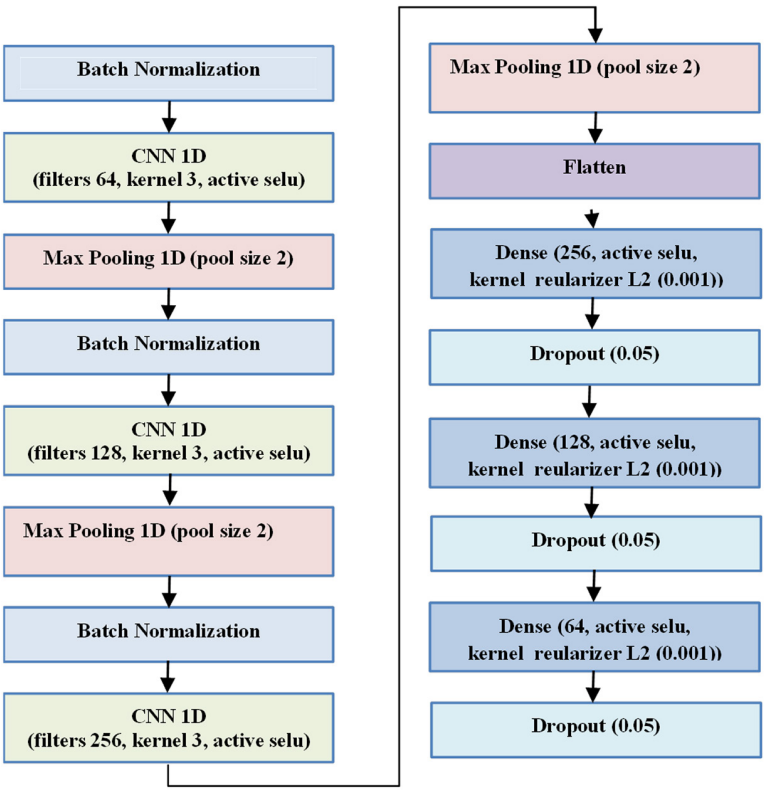

Fig. 2. CNN block

The transformer design includes an encoder block and dense layers (softmax). Figure 3 depicts the encoder block, which consists of a multi-head attention layer, two one-dimensional convolutional layers, two normalization layers, and a feed-forward network. The feed-forward network has a linear layer and SeLU activation function, which processes each embedding vector individually with equal weights. Consequently, each embedding vector undergoes a position-wise feed-forward layer before further transformation. Additionally, multi-head attention is utilized to determine the importance of each head, represented as a one-dimensional vector. A skip connection is then applied to each, involving a simple element-wise addition.

$$x + Sublayer\ (x) \tag{2}$$

The sub-layer might be either multi-head attention or a feed-forward network. Skip connections and transfer prior embeddings to succeeding layers. As a result, the encoder blocks enrich the embedding vectors with extra information obtained via multi-head self-attention computations and feed-forward networks. Each skip connection is followed by layer normalization to mitigate the effect of the covariate shift.

$$LayerNorm(x + Sublayer\ (x)) \tag{3}$$

The key component of the transformer encoder is the multi-head attention layer. This layer allows the model to flexibly learn information from the embedding representations of different features. The multi-head attention layer consists of multiple heads of self-attention, also referred to as "scaled dot-product attention."

$$\text{multi head attention}\ (Q, V, K) = Concat\ (h_1, \ldots, h_i)\ W^o \tag{4}$$

Where *h_i* is computed as

$$h_i = Atten(QW_i^Q, KW_i^K, VW_i^V) \quad (5)$$

In this work, we utilized three heads of multi-head attention layers, eight embedding dimensions, a 0.05 post-attention dropout ratio, and a feed-forward multi-layer factor of [3, 2, 1]. (N-times) represents several blocks (see Figure 3).

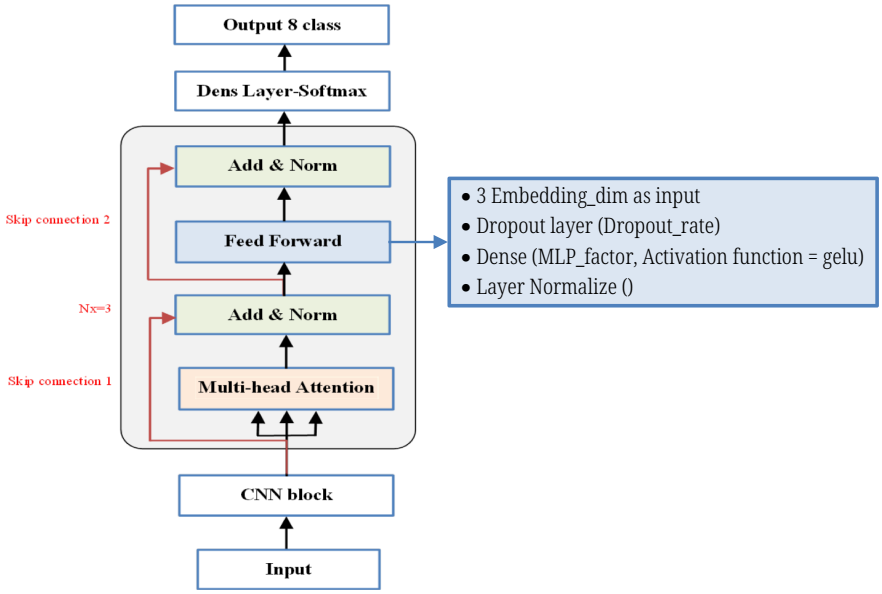

Fig. 3. Encoder transformer

For further experiments on the same database, we utilized multilayer perceptron (MLP) and XGBoost, comparing them with CNN-Transformer. The structure of the MLP model is illustrated in Figure 4.

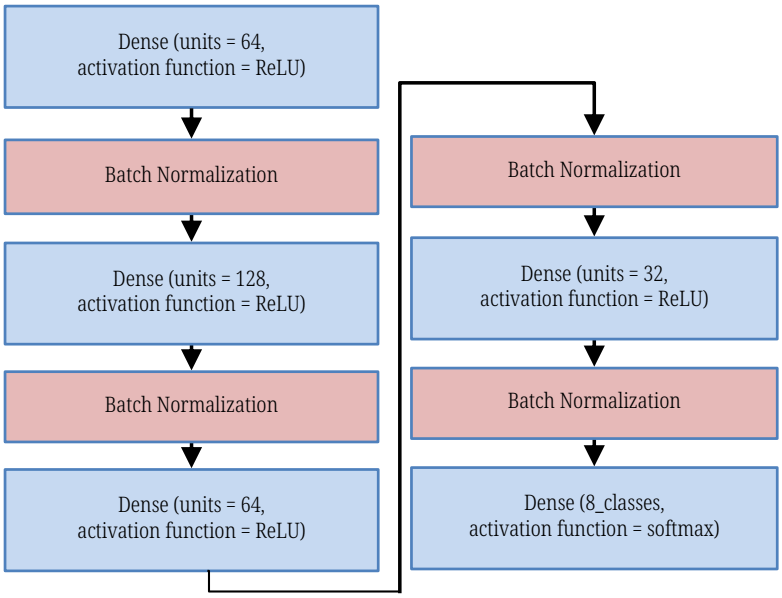

Fig. 4. MLP model

4 EXPERIMENTAL DESIGNS

4.1 CICIoT2023 datasets

The Canadian Institute for Cybersecurity (CIC) has established an innovative and comprehensive IoT threat dataset to promote the development of security analysis applications in real-world IoT deployments [20]. Thirty-three attacks were conducted on an IoT network consisting of 105 devices. These attacks are categorized into seven types (DDoS, Recon, DoS, Web-based, spoofing, brute force, and Mirai), with 46,179,314 typical attack statistics. The training set has a shape of 37,349,263, the test set has a shape of 8,830,051, and the dataset includes 46 features. All attacks are carried out by malicious IoT devices targeting other IoT devices. The attacks directly impacted 67 IoT devices, with an additional 38 Z-Wave and Zigbee devices interconnected to five hubs. Smart home devices, sensors, cameras, and microcontrollers are connected and configured to enable multiple attacks to be executed while capturing the resulting attack traffic. Network activity is monitored using Wireshark and stored in pcap format. Since two data streams are stored, mergecap is used to merge the pcap files for each experiment. The dataset was generated using IoT devices such as audio devices, cameras, hubs, power outlets, home automation systems, lighting, sensors, and NextGen devices. Table 2 provides a summary of the number of attacks and recordings.

Table 2. Types and amounts of records in the CICIoT2023 dataset, as well as the testing and training sets

Type of Event	Data Record	Train Set	Test Set
DDoS	33,984,560	37,349,263	8,830,051
DoS	8,090,738		
Mirai	2,634,124		
Benign	1,098,195		
Spoofing	486,504		
Recon	354,565		
Web	24,829		
BruteForce	13,064		
Total	46,179,314		

4.2 Features preprocessing

When we started working on this paper, the first challenge we encountered was obtaining enough big data to train and test our model. To address this challenge, we divided the data into 169 CSV files and consolidated the 34 classes into eight classes. Each file was processed independently, with 80% allocated for training and 20% for testing. This approach helped mitigate the risk of overfitting and facilitated comprehensive model performance monitoring across the entire dataset. We also ensured that the classes were stratified for both training and testing. Subsequently, we stored the data in separate directories using Spark for training and testing purposes. Additionally, we developed a Spark pipeline to normalize the dataset through MinMax normalization, trained it using the data from the training directory, and

then applied the model to both the training and testing directories before saving the results. The dataset comprised 46 features and classes, including Benign, Brute Force, Mirai, DDoS, Spoofing, DoS, Recon, and Web-based.

In the original dataset paper, it was noted that the protocol type feature is an integer representing the type of protocol. However, we discovered that it is stored as a float.

4.3 Metrics for evaluation

We evaluated our proposed model using various indicators, such as accuracy, recall rate, precision, and F1 score. Additionally, we conducted a systematic benchmark comparison with other relevant approaches. These indicators are commonly used in intrusion detection systems, where true positives (TP) and true negatives (TN) represent accurately predicted values. False positives (FP) and false negatives (FN) indicate incorrectly classified occurrences [21, 22].

Accuracy: The percentage of samples and applications properly categorized in a dataset. A higher accuracy number indicates that the classifier is precise.

$$\text{Accuracy} = (TN + TP)/(FN + TP + TN + FP) \tag{6}$$

Precision refers to the number of accurately detected benign and positive samples and applications in the dataset. A classifier with a higher accuracy value outperforms others and is preferred.

$$\text{Precision} = TP/(TP + FP) \tag{7}$$

F1-Score: The F1 score is calculated by taking the harmonic mean of a classifier's recall and precision.

$$\text{F1-score} = 2.(\text{Recall} * \text{Precision})/(\text{Recall} + \text{Precision}) \tag{8}$$

Recall: This metric calculates the proportion of true positive predictions out of all possible positive predictions.

$$\text{Recall} = TP/TP + FN \tag{9}$$

5 EXPERIMENTAL RESULTS

This experiment evaluated the effectiveness of the CNN-Transformer model in detecting cyberattacks using the new CICIoT2023 datasets. The datasets included 37,349,263 training data, 8,830,051 test data, and 46 features for analysis. Four metrics were employed based on various classes: accuracy, recall, F1 score, and precision. The highest accuracy achieved was 99.47%, with precision at 94.21%, recall at 75.76%, and F1 score at 79.55%, as depicted in Figure 5. The figure shows the proportion of accurately predicted attacks for the eight classes of the confusion matrix displayed in Figure 6. Each metric was computed individually using a learning rate of 0.00005, a batch size of 1024, 50 epochs, and a dropout rate of 0.05. Adam was used as the optimizer for this experiment.

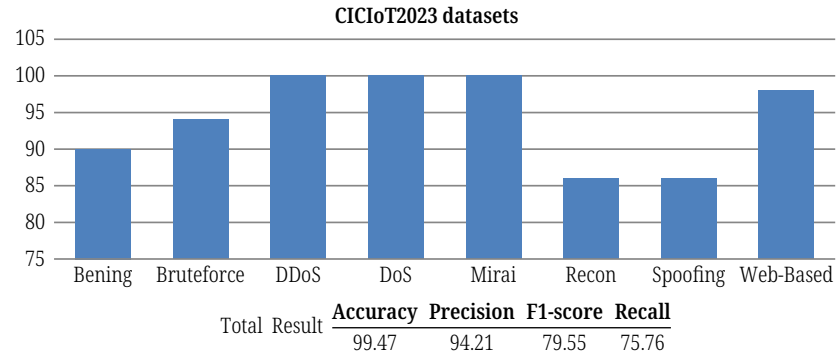

Fig. 5. CICIoT2023 dataset's multi-class classification and CNN-Transformer findings

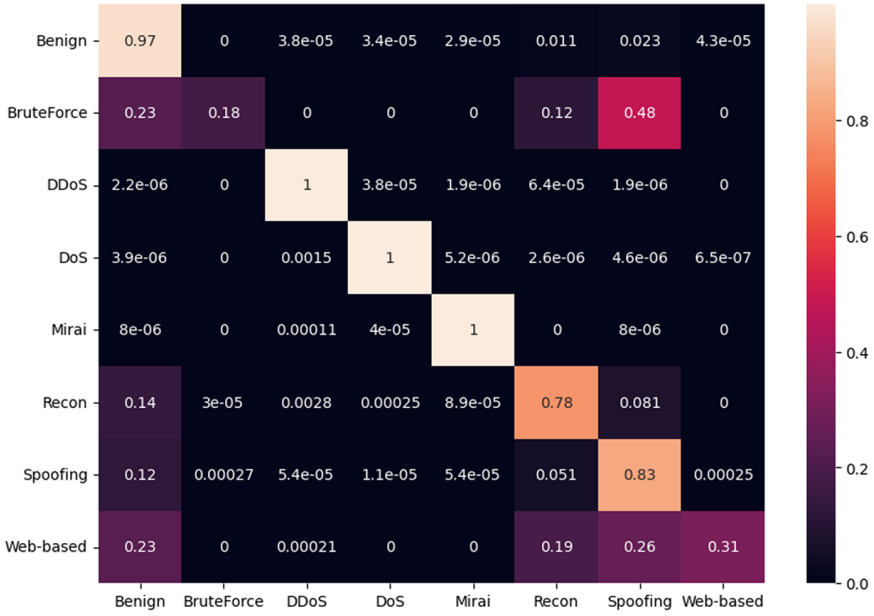

Fig. 6. The confusion matrix displaying the performance of the CNN-Transformer validation for eight classes: Benign, DDoS, BruteForce, DoS, Mirai, Spoofing, Recon, and Web-based (The values in the matrix range from 97 to 31.)

Table 3. Performance of the CICIoT2023 dataset on IoT network data for eight classes

Class Type	Precision	Recall	F1-Score
Benign	0.90	0.97	0.93
BruteForce	0.94	0.18	0.30
DDoS	1.00	1.00	1.00
DoS	1.00	1.00	1.00
Mirai	1.00	1.00	1.00
Recon	0.86	0.78	0.82
Spoofing	0.86	0.83	0.84
Web-based	0.98	0.31	0.48
Accuracy		0.99	
Macro Avg	0.94	0.76	0.80
Weighted Avg	0.99	0.99	0.99

In our tests, we also utilized the multi-layer perceptron (MLP) and XGBoost pipeline algorithms to evaluate the performance of the CICIoT2023 dataset with consistent training and testing values and the original features. The results indicated the CNN-transformer model's accuracy convergence with both MLP and XGBoost, as illustrated in Table 4. With MLP, the accuracy reached 99.39%, precision was 94.66%, recall stood at 75.72%, and the F1 score was 80.11%. The hyperparameters included a learning rate of 0.0002, a batch size of 1024, 100 epochs, and a dropout rate of 0.05. The optimizer used for this experiment was Adam. For the XGBoost pipeline, the dataset of 169 was split into 2 directories initially, allocating 80% (1) of the CSV files to the train directory and 20% (15 CSV files) to the test directory. The data was then mapped from 34 classes to 8 classes, normalized using MinMax normalization, and an XGBoost classifier stage was added to the pipeline. The accuracy was 99.40%, precision was 90.93%, recall was 75.70%, and the F1-score was 79.26%.

Table 4. Comparison of CNN-Transformers model with previous studies based on eight classes

Paper	Model	Accuracy (%)
Neto EC. [20]	Random Forest	99.43
Jony AI. [23]	LSTM	98.75
Wang Z. [24]	DL-BiLSTM	93.13
Gheni HQ. [25]	MLP	97.47
Our model	CNN-Transformers	**99.47**
	MLP	99.39
	XGBoost	99.40

Table 4 compares our results to the original using the same database but a different model, the CNN-Transformers model. The CNN-Transformers model outperformed the other models, achieving an accuracy of 99.47%, precision of 94.21%, f1-score of 79.55%, and recall of 75.76%.

The results demonstrate that, despite the heterogeneity of the IoT data, our technique can extract useful information to enhance classification performance. These findings show that our technique can outperform the latest pure network data from CICIOT2023. In terms of all accuracy criteria, our solution exceeds the most advanced machine learning algorithm for IoT network data.

6 CONCLUSION AND FUTURE WORK

Nowadays, the IoT is becoming increasingly crucial to society. In this setting, developing security solutions is critical to allowing efficient, safe, and reliable IoT operations. This study developed a CNN-Transformer model to identify IoT assaults, with the goal of encouraging the development of security analytics applications in real-world IoT operations and eventually improving the identification of anomalous activities and traffic violations in IoT networks. Furthermore, we tested the performance of CNN-Transformer, MLP, and XGBoost-pipeline on the new CICIoT2023 datasets. The CNN-Transformer we presented achieved an accuracy of 99.49%, MLP had an accuracy of 99.39%, and XGBoost had an accuracy of 99.40%. The experimental findings show that our model outperforms the mainstream conventional and

deep learning intrusion detection algorithms used in other IDSs. These accuracies were achieved using the Adam optimizer. Our future research will focus on accelerating the transformer algorithm for a quick-response intrusion detection system to significantly reduce the harm caused by abnormal events. We also aim to apply our idea to more challenging scenarios, such as edge cloud systems. These decentralized systems present specific challenges that our approach must address.

7 ACKNOWLEDGEMENTS

Mohamed Ahmed (mohamed.ahm.cs@gmail.com) is our chosen engineer and we thank him for his exceptional contributions to the coding of this study.

8 REFERENCES

[1] S. A. Jebur and S. N. Mazkhor, "Development of smart healthcare monitoring system based on IoT," in *AIP Conference Proceedings*, vol. 3079, 2024, no. 1, p. 060001. https://doi.org/10.1063/5.0201958

[2] A. Nauman, Y. A. Qadri, M. Amjad, Y. B. Zikria, M. K. Afzal, and S. W. Kim, "Multimedia Internet of Things: A comprehensive survey," *IEEE Access*, vol. 8, pp. 8202–8250, 2020. https://doi.org/10.1109/ACCESS.2020.2964280

[3] Z. Xiaofeng and H. Xiaohong, "Research on intrusion detection based on improved combination of K-means and multi-level SVM," in *2017 IEEE 17th International Conference on Communication Technology (ICCT)*, Chengdu, China, 2017, pp. 2042–2045. https://doi.org/10.1109/ICCT.2017.8359987

[4] I. Tareq, B. M. Elbagoury, S. A. El-Regaily, and E.-S. M. El-Horbaty, "Deep reinforcement learning approach for cyberattack detection," *Int. J. Online Biomed. Eng.*, vol. 20, no. 5, pp. 15–30, 2024. https://doi.org/10.3991/ijoe.v20i05.48229

[5] H. Rafik, A. Maizate, and A. Ettaoufik, "Data security mechanisms, approaches, and challenges for e-health smart systems," *Int. J. Online Biomed. Eng.*, vol. 19, no. 2, pp. 42–66, 2023. https://doi.org/10.3991/ijoe.v19i02.37069

[6] L. Alzubaidi, A. D. Khamael, H. A. Obeed, A. Saihood, M. A. Fadhel, S. A. Jebur, Y. Chen, A. S. Albahri, J. Santamaría, A. Gupta, and Y. Gu, "MEFF – A model ensemble feature fusion approach for tackling adversarial attacks in medical imaging," *Intelligent Systems with Applications*, vol. 22, p. 200355, 2024. https://doi.org/10.1016/j.iswa.2024.200355

[7] J. Kim *et al.*, "CNN-based network intrusion detection against denial-of-service attacks," *Electronics*, vol. 9, no. 6, p. 916, 2020. https://doi.org/10.3390/electronics9060916

[8] A. Vaswani *et al.*, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998–6008, 2017.

[9] S. A. Jebur, K. A. Hussein, H. K. Hoomod, L. Alzubaidi, and J. Santamaría, "Review on deep learning approaches for anomaly event detection in video surveillance," *Electronics*, vol. 12, no. 1, p. 29, 2022. https://doi.org/10.3390/electronics12010029

[10] S. A. Jebur, A. K. Nawar, L. E. Kadhim, and M. M. Jahefer, "Hiding information in digital images using LSB steganography technique," *International Journal of Interactive Mobile Technologies*, vol. 17, no. 7, pp. 167–178, 2023. https://doi.org/10.3991/ijim.v17i07.38737

[11] H. Wang and W. Li, "DDosTC: A transformer-based network attack detection hybrid mechanism in SDN," *Sensors*, vol. 21, no. 15, p. 5047, 2021. https://doi.org/10.3390/s21155047

[12] Z. Wu, H. Zhang, P. Wang, and Z. Sun, "RTIDS: A robust transformer-based approach for intrusion detection system," *IEEE Access*, vol. 10, pp. 64375–64387, 2022. https://doi.org/10.1109/ACCESS.2022.3182333

[13] S. Ullah *et al.*, "TNN-IDS: Transformer neural network-based intrusion detection system for MQTT-enabled IoT networks," *Computer Networks*, vol. 237, p. 110072, 2023. https://doi.org/10.1016/j.comnet.2023.110072

[14] M. A. Ferrag, M. Debbah, and M. Al-Hawawreh, "Generative AI for cyber threat-hunting in 6G-enabled IoT networks," in *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)*, Bangalore, India, 2023, pp. 16–25. https://doi.org/10.1109/CCGridW59191.2023.00018

[15] M. Wang, N. Yang, and N. Weng, "Securing a smart home with a transformer-based IoT intrusion detection system," *Electronics*, vol. 12, no. 9, p. 2100, 2023. https://doi.org/10.3390/electronics12092100

[16] Z. Long *et al.*, "A transformer-based network intrusion detection approach for cloud security," *Journal of Cloud Computing*, vol. 13, 2024. https://doi.org/10.1186/s13677-023-00574-9

[17] A. Saghir *et al.*, "Explainable transformer-based anomaly detection for Internet of Things security," in *The Seventh International Conference on Safety and Security with IoT (SaSeIoT 2023),* in EAI/Springer Innovations in Communication and Computing, K. P. Tran, S. Li, C. Heuchenne, and T. H. Truong, Eds., Springer, Cham, 2024, pp. 83–109. https://doi.org/10.1007/978-3-031-53028-9_6

[18] K. Stein, A. Mahyari, G. Francia III, and E. El-Sheikh, "A transformer-based framework for payload malware detection and classification," *arXiv preprint arXiv:2403.18223*, 2024. https://doi.org/10.48550/arXiv.2403.18223

[19] T. Jiang, X. Fu, and M. Wang, "BBO-CFAT: Network intrusion detection model based on BBO algorithm and hierarchical transformer," *IEEE Access,* vol. 12, pp. 54191–54201, 2024. https://doi.org/10.1109/ACCESS.2024.3386405

[20] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A real-time dataset and benchmark for large-scale attacks in IoT environment," *Sensors*, vol. 23, no. 13, p. 5941, 2023. https://doi.org/10.3390/s23135941

[21] K. Shaukat, S. Luo, V. Varadharajan, I. A. Hameed, and M. Xu, "A survey on machine learning techniques for cyber security in the last decade," *IEEE Access*, vol. 8, pp. 222310–222354, 2020. https://doi.org/10.1109/ACCESS.2020.3041951

[22] S. A. Jebur, K. A. Hussein, H. K. Hoomod, and L. Alzubaidi, "Novel deep feature fusion framework for multi-scenario violence detection," *Computers*, vol. 12, no. 9, p. 175, 2023. https://doi.org/10.3390/computers12090175

[23] A. I. Jony and A. K. Arnob, "A long short-term memory-based approach for detecting cyber-attacks in IoT using CIC-IoT2023 dataset," *Journal of Edge Computing*, vol. 3, no. 1, pp. 28–42, 2024. https://doi.org/10.55056/jec.648

[24] Z. Wang *et al.*, "A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization," *PeerJ Computer Science,* vol. 9, p. e1569, 2023. https://doi.org/10.7717/peerj-cs.1569

[25] H. Q. Gheni and W. L. Al-Yaseen, "Two-step data clustering for improved intrusion detection system using Ciciot2023 dataset," *SSRN*, 2024. https://dx.doi.org/10.2139/ssrn.4762201

9 AUTHORS

Imad Tareq Al-Halboosi is an employee of the Dewan Al-Waqf Al-Sonny in Iraq. He received master's degree in Computer Science from Middle East University, Jordan, in 2016, and is currently pursuing a Ph.D. degree there. His research interests include Cybersecurity, network computer security, and the Internet of Things, and artificial intelligence applications. You can reach him via email at: cs.20.19@grad.uotechnology.edu.iq.

Bassant Mohamed Elbagoury received Ph.D. and M.Sc. degrees in Computer Science from the Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt, in 2009 and 2005, respectively. She was a Ph.D. Researcher with the NAO Robot team in Humboldt, Germany. She completed her Ph.D. thesis in only two and a half years. Since 2003, she has participated in many international conferences in Paris, Poland, Germany, Jordan, and the USA. Since 2005, she has been a Reviewer in AAAI conferences in the USA. She is currently an Assistant Professor of computer science at the Faculty of Computer Science and Computer Engineering, King Salman International University (KSIU). Her research areas include robotics, artificial intelligence, mobile computing, and cloud computing (E-mail: drbassantcs@gmail.com).

Salsabil Amin El-Regaily is a Lecturer in the Faculty of Computer and Information Sciences at Ain Shams University in Cairo, Egypt. She obtained her PhD in Computer Science in 2019 and currently serves in the Basic Sciences Department within the faculty. Dr. Salsabil previously held positions as a Teaching Assistant and Assistant Lecturer. Her areas of expertise include Structured Programming, Image Processing, and Deep Learning. She is a member of the Ain Shams University Ranking team and serves as the Deputy Director of the Assessment and Evaluation Unit at the Faculty of Computer and Information Sciences. Her professional roles involve being an academic advisor for credit hour programs, a part of the Quality Assurance unit, and a member of the Egypt Government Excellence Award team. Dr. Salsabil has presented numerous research papers in both local and international journals and conferences, focusing on the field of Image Processing, particularly in Medical Imaging. Her Scopus H-Index is 3, with 103 citations to date. In 2021, she was honored by the Association of Arab Universities for having the second-best PhD thesis in the field of Artificial Intelligence. In 2023, she received the Ain Shams Incentive Award in the advanced medical technological science field (E-mail: salsabil_amin@cis.asu.edu.eg).

Professor El-Sayed M. El-Horbaty received his Ph.D. (1985) in Computer Science from London University, UK, his M.Sc. (1978), and B.Sc. (1974) in Mathematics from Ain Shams University, Egypt. He has worked as an academic in Egypt (Ain Shams University), Qatar (Qatar University), and the Emirates (Emirates University, Ajman University, and ADU University). Prof. El-Horbaty's current research areas include Distributed and Parallel Computing, Cloud Computing, Mobile Cloud Computing, Edge Computing, e-health Computing, IoT, and Optimization of Computing Algorithms. His work has been published in many reputed journals, nationally as well as internationally (E-mail: shorbaty@cis.asu.edu.eg).