Grenze International Journal of Engineering and Technology, June Issue

A Hybrid Deep Learning Model for Attack Detection and Classification in IoT Network

Ajay A V[1], Pramod H B[2] and Gowtham M[3]
[1]Department of Computer Science and Engineering, JSS Science and Technology University, Mysuru, Karnataka, India
Email: ajush1331@gmail.com
[2]Department of Information Science and Engineering, Adichunchanagiri Institute of Technology, Chikkamagaluru, Karnataka, India
Email: hbpramod@gmail.com
[3]Department of Information Science and Engineering, JSS Science and Technology University, Mysuru, Karnataka, India.
Email: gouthamgouda@gmail.com

***Abstract*— The advent of Internet of Things (IoT) devices has introduced a wide array of security challenges. We find that which security breaches faced in IoT systems' and what class of them they fall into is a priority in securing private information. In this paper we propose a hybrid deep learning model for the detection and classification of security threats in the Internet of Things. We combine the power of Random Forest with CNN Features, Decision Tree and Extreme machine learning to produce a model which is very good at analysing and classifying security threats. It takes into consideration features in the raw IoT data, and a few of the features in our model considers time and trend aspects in the data. For the training and evaluation of the model we put together a large set of our own simulated security attack scenarios in IoT networks. The results show that the model performs very well on the tasks proposed, thus showing givenness of the system for real time security supervision. Our model which utilises deep learning technologies also answers the deficiencies there are in the security of the Internet of Things. With the hybrid model proposed security personnel can quickly detect and reply to damaging intrusions which could affect the Internet of Things and important information. Future work could analyze the model performance, scalability and the diverse ecosystems of the Internet of Things.**

***Index Terms*— IoT, IoT attacks, Denial-of-Service, Random Forest, Convolution Neural Network, Intrusion Detection System (IDS), Deep Learning.**

I. INTRODUCTION

The Internet of Things (IoT) is a network of devices such as physical gadgets, automobiles, home appliances and other objects. These devices contains sensors, actuators, network connectivity and software, so that they collect the data from the devices and exchange it with other devices, applications and people. The data collected by devices can be analysed and used to automate the operations, boost efficiency and to develop new services and applications. The industries like transportation, agriculture, manufacturing healthcare and smart homes can be transformed by IoT[5]. For instance the IoT devices in healthcare industry can be used to monitor patients, monitor their medical data remotely and alerts staff if there are any potential problems. Similarly the sensors in transport industry can be used to monitor traffic patterns, optimize the routes so that increases the safety.

Grenze ID: 01.GIJET.12.2.102_21

In agriculture IoT sensors measures soil moisture levels, crop growth and meteorological conditions to optimize irrigation and increase the crop yields. As IoT technology advances, it become more predominant in our daily lives, transforming the interaction with physical world. Deep Learning(DL) is way of machine learning for pattern recognition, classification and perdition using artificial neural networks. Here the networks are learn from large amount of data and then model functions as human brain. DL algorithms process and modify input data using numerous layers of artificial neurons[11].

Mathematical operations are applied on the data in each layers and output will be sent to next layer. The arrangement of layers will be in a hierarchical structure, each layer learns increasingly complex feature from the data. DL can learn and extract the features from raw data automatically. DL algorithms can generalize to novel data and learn from large datasets, makes them suitable for applications involving high dimensional and complex data. DL is a growing area and numerous research are place to develop new algorithms and architectures that solves more challenging problems. A hybrid DL model can build by combining different DL methods into a single model, such as convolutional layers for image processing, recurrent layers for sequence modeling and attention layers for context based learning[12]. The model can be trained with different methods such as back propagation, reinforcement learning and unsupervised learning. These models are being used in speech recognition, computer vision and robotics. With the evolution in the DL, models are going to play vital role in solving complex problems.

A. Importance of Security in IoT

Since the devices in IoT are extremely interconnected, the security is a prime concern because they are vulnerable to various cyber attacks. As the IoT network has more than one interconnected devices and system, each of them can be a potential pathway for an attacker. Hence, it is crucial to make each part of the network secure and the communication between them. Intrusions in IoT networks may lead to negative effects such as unauthorized access to information, interruption of services, and even physical injury to individuals and environment[6].

IoT devices capture data that is more sensitive in nature, and unauthorized access to such data may result in privacy breaches, identity theft, and economic losses. Intrusions in industrial IoT systems may lead to physical damage to machinery, production downtime, and even endanger lives. In order to avoid such harmful effects, security practices like authentication, access control, encryption, and intrusion detection are crucial in IoT systems. IoT devices must be designed secure by default and must be provided with regular updates and patches to meet vulnerability fixes. Network traffic must be monitored and analyzed for detecting any abnormal activity and intrusions, and necessary measures should be taken in case of any security incidents. With growing deployment of IoT solutions across different fields, it is important to make the devices and networks secure in order to avoid any unauthorized access or malicious attacks that can lead to serious impacts. Some of the most common network intrusion attacks are enumerated below:

- U2R: A User to Root (U2R) attack in the context of IoT is an illegal attempt to gain administrative privileges to an IoT device or system.
- Probe: It's a type of cyber-attack that targets IoT devices in which probes are used to probe IoT devices for vulnerabilities.
- R2L: Remote to Local(R2L) is a type of cyber-attack used to target IoT devices to achieve unauthorised access to a network or device.
- DOS: It's a type of cyber-attack where an army of hacked IoT devices is used to flood a specific website or network, rendering it inaccessible to genuine users.
- Replay: Replay attacks are an attack in IoT systems where an attacker captures and stores a message that was transmitted between two devices.
- Blackhole: An attack whereby a hijacked device is used to form a "blackhole" within the network, dropping or intercepting traffic before it reaches its destination.
- Sinkhole: A sinkhole IoT attack is a sort of cyber-attack that targets internet-connected IoT devices, which may have lax security measures.
- Spoofing: Spoofing in IoT is a sort of cyber-attack where an attacker manipulates their identity so that they can have unauthorized access to an IoT device, network or application.

B. Problem Statement

The amount of data causes a reduction in detection accuracy and an increase in false reports of network intrusion. This is one of the most important issues when the system is under attack by unknown intruders. The main aim must be to increase detection accuracy and decrease false alarm incidence.

II. Related Work

Connected to the Internet, IoT devices are vulnerable to security breaches as a result of their vast networks and interconnectedness[1]. Studies show that about 70% of IoT devices are vulnerable to serious security threats[2]. The increased connectivity and use of IoT technology has resulted in an alarming rise in security threats. The amount of cybercrime in the IoT environment is growing at a greater rate than security measures can be put in place to counteract such threats. The generic word "attack" concerning IoT information security refers to various forms of vulnerabilities, security threats, and cybercrimes which threaten the integrity of the system. These "attacks" can be denial of service (DoS), where the IoT system is bombarded with such a huge volume of traffic that it can no longer remain responsive or efficient, breaches of sensitive information, or unauthorized access to sensitive information. Furthermore, IoT devices hold and collect sensitive data, thereby exposing them to unauthorized personnel who may exploit this sensitive information for inappropriate deliberate gain[3]. The analysis of these attacks through the medium of machine learning algorithms could provide valuable insight into countermeasure development and greater security for these devices. Machine learning algorithms could be used to highlight patterns and aberrant behavioral characteristics in network traffic flow, which could give insight into the different types of cyber attacks on IoT devices[8]. Further machine learning algorithms could be formatted or trained, in order to classify and categorize different types of attacks, including covert eavesdropping, traffic analysis, camouflage attacks, and spoofing. By analyzing the behavioral characteristics of information flow in the form of network traffic, it may be possible to highlight the aberrant traffic, which would be the factor indicating a possible attack of sorts. The algorithms can also be used to continuously assess and improve the security measures of IoT systems making it possible to detect and respond in real time emerging threats. It has also been suggested that neural networks can be used to build improved anomaly based IoT intrusion detection systems[4]. The machine learning algorithms such as neural networks have shown to provide promising results for developing enhanced anomaly based IoT intrusion detection systems with accurate precision. Nevertheless, it should be noted that the machine learning algorithms can improve IoT security but are not foolproof. They are vulnerable to attacks and require continuous monitoring and updating to stay ahead of evolving threats[9]. In addition, the rapid rise of numerous attacks and threats in IoT systems has rendered previous security mechanisms inadequate. Therefore, it is crucial to continuously enhance and modernize IoT security measures. We can achieve these security measures by implementing the model as hybrid. We could classify different types of IoT attacks such as DOS, DDOS, control hijacking, firmware attacks etc[10]. For the classification and detection of all these attacks the algorithms that were mostly used were CNN[7] and LSTM. Machine learning algorithms like KNN, SVM, DT, NB, random forest (RF), ANN, and LR were used but after research we can conclude that hybrid deep learning models are more accurate and precise.

Building on these foundations, the research community has increasingly focused on deep learning architectures steadily to try to cope with the complexity of network traffic in IoT. Long Short-Term Memory (LSTM) and Convolutional Neural Networks (CNN) shown particular promise for classifying the wide array of attacks seen in IoT systems. However, a significant research gap persists. While these models are often used as end-to-end classifiers, there has been less exploration into a hybrid approach where a deep learning model serves purely as an automated feature extractor. In this context, we tackle an existing problem by addressing the need for a new practical methodology. The core of this strategy will place an emphasis on using a CNN for the learning and extraction of high-level feature representations from raw data. These representations are then introduced into a robust ensemble classifier, namely the Random Forest, for the final classification. This system is designed to capitalize on the great feature learning techniques of Deep Learning with the proven classification power of its traditional counterparts, thus providing a new improved and efficient method of Intrusion Detection.

III. Methodology

Our study will then apply this methodology to the NSL-KDD dataset, an established standard in contribution to the evaluation of IDS. The initial step in the methodology is data pre-processing where libraries like NumPy and Pandas will be used to clean up the data and to specifically deal with irrelevant features. Following this, we shall apply the K-Means clustering to perform initial grouping of the data and the determination of centroids thus segregating the anomalous from the normal data traffic.

Following this pre-processing, the data is passed to our core contribution: the hybrid deep learning model. The detailed architecture of this model, which combines CNN-based feature extraction with Decision Tree and Random Forest classifiers, is illustrated in Figure 1.

A. Decision Tree

Decision trees are a kind of supervised machine learning in which the data is regularly separated according to a certain parameter. The tree is built with two elements. The decisions and results are completed in the leaves. The data is separated at the decision nodes Regression and classification problems can be resolved using decision tree approaches, in contrast to conventional supervised learning algorithms [15].

B. Random Forest

Random forests are often used in data science contests as well as real-world problems. They are typically accurate, rarely require parameter adjusting, and don't call for feature scaling or category feature encoding[13]. Compared to more complex models like neural networks, they may also be simpler to comprehend. A random forest is composed of a number of random decision trees. There are two types of randomness that are found in the trees. First a random sample of the original data is used to make the tree. The second type of randomization means that a subset of the features is randomly picked to produce the best split at each node of the tree[14].

C. Proposed Hybrid Model Architecture

The core of our proposed method lies in its hybrid architecture, which is designed to leverage the distinct strengths of CNNs for automated feature extraction and ensemble tree-based models for robust classification. The architecture is a sequential multi-stage architecture in which the output of each stage becomes the refined input to the next. A conceptual diagram of this architecture is in Figure 1.

Stage 1: Data Transformation and Reshaping: The NSL-KDD data set, like most network traffic data, is inherently tabular. To reshape this data to be compatible with a CNN learnt on spatial hierarchies (like images), we must first pass through a data transformation phase. Each network connection which is represented as a 1D feature vector is reshaped to yield an input as a 2D matrix. This reshaping allows the Convolutional Neural Network to treat the features as if they belong to a small image and recognize local structures and inter-relationships between adjacent features of the vector.

Stage 2: Feature Extraction via CNN: The reshaped data matrix then enters the lightweight CNN. It is important to note that the CNN part of our model is not applied to end-to-end classification. It is rather used as a powerful automated feature extractor. The architecture of the network itself consists of layers of convolution, followed by layers of ReLU, and finally layers of max-pooling. The purpose of these layers is to learn the features of the input data in an hierarchical and abstracted way which it would have previously been impossible or very difficult to do manually.

Stage 3: Feature Fusion and Classification: The fusion occurs at the interface of the CNN and the tree-based classifiers. Having passed through the convolution and pooling layers, the output feature maps are then flattened into a single, one-dimensional feature vector. This vector is a representation of the high-level abstract features learned by the CNN.

This flattened feature vector is then passed as the input to both the RF and DT classifiers. Here, the term "Random Forest with CNN Features" indicates the RF model trained on this advanced set of features generated by the CNN, rather than that of the raw input data. Both the RF and DT models then operate independently in performing the final classification task, predicting whether the network traffic relates to normal behaviour or a specific category of attack. This systematic process ensures that the classifiers base their decisions on a rich, distilled group of features, which significantly enhances detection accuracy and lowers the false positive rate.

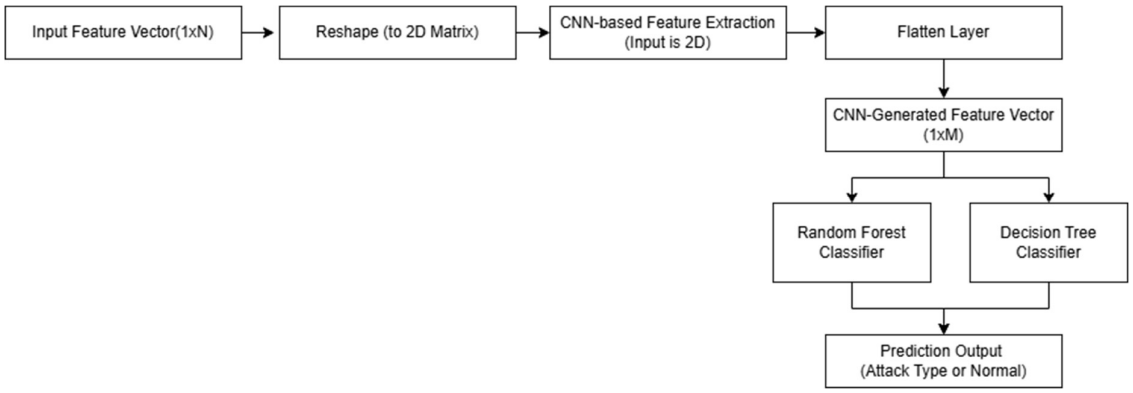

Figure 1. Detailed Architecture of the Proposed Hybrid Model

IV. RESULT ANALYSIS

In order to check the feasibility of our model three different types of analyses were performed. We first evaluated the performance of our hybrid model against its fundamental parts, then studied the classification accuracy of it like some attack types, and finally compared it against other existing modeled state-of-the-art methods.

A. Internal Model Performance

The performance of the our hybrid model was compared against a number of different baseline algorithms, including a pure Decision Tree (DT), SVM, a single CNN, a hybrid CNN-RNN. This is shown in Table I, where it is noted that our proposed hybrid model which has features extracted with a CNN and then classified with ensemble classifiers gives the best performance of any of the important metrics.

The model gives an accuracy of 96.92% and True Positive Rate (TPR) of 96.16%, while decreasing the False Positive Rate (FPR) to 8.59%. This low FPR ensures that it is simple to reduce the number of false alarms in a real security environment. The high performance attained tends to indicate that the hybrid architecture is in order.

TABLE I. PERFORMANCE EVALUATION OF INTERNAL MODELS

Model	TPR (%)	FPR (%)	Accuracy (%)
Decision Tree (DT)	90.33	12.63	92.47
SVM	93.47	11.46	94.75
CNN	92.74	10.23	93.05
CNN-RNN	95.94	9.12	91.45
Proposed Hybrid Model	96.16	8.59	96.92

B. Attack-Specific Classification Accuracy

When all the overall performance had been analyzed, attention was given to the actual performance of the model in classifying the eight attack types inherent in the data set studied. Table II gives the classification accuracy for all types including R2L, U2R, Probe, DoS, Replay, Blackhole, Sinkhole and Spoofing. The analysis has yielded results that show all attack types are identified with a high degree of accuracy, demonstrating the robustness of the system and its ability to distinguish between dissimilar forms of invasive behavior in network traffic.

TABLE II. CLASSIFICATION ACCURACY OF DETECTED ATTACK TYPES

Attack Type	Accuracy (%)
U2R	92.86
Probe	93.53
R2L	90.59
DoS	94.65
Replay	92.86
Blackhole	93.53
Sinkhole	90.59
Spoofing	94.65

C. Comparative Evaluation with Benchmark Models

In order to effectively evaluate the performance of the proposed hybrid model a thorough quantitative comparison was made with a selection of state-of-the-art deep learning models that are commonly used in IDSs. These included not only Gated Recurrent Unit (GRU) and LSTM models, both are well known for their overall good performance in the coping of sequential, high-dimensional data such as those representing network traffic behaviour, but also a combined CNN and LSTM architecture and an Autoencoder-based anomaly detection model. The aforementioned models were all trained and tested on the same NSL-KDD data in order to ensure that a fair and direct comparison was made.

The performance measures of the comparative evaluation are given in Table III. The results show that recurrent models (such as LSTM and GRU) do, indeed, give a good result; however, the proposed hybrid model gives a markedly superior performance on all the main performance measures. To expand on this LSTM and GRU gave respectively 95.17% and 95.48% on the accuracy performance measure, whereas the CNN-LSTM model of model G yields an accuracy of 96.92%. In addition the FPR of model G of 8.59% is the lowest of all the evaluated systems. This reduction of the FPR of a model is highly important for real world deployability of an IDS as this implies a low number of false alarms allowing security personnel to zero in on real threats. The

enhanced results support the notion that the hybrid model has enhanced performance over all others evaluated due to the combination of CNN performed dimensionality reduction capability with Random Forest and Decision Tree classifiers leading to enhanced discriminatory power.

TABLE III. COMPARATIVE PERFORMANCE ANALYSIS WITH STATE-OF-THE-ART IDS MODELS ON THE NSL-KDD DATASET

Model	Accuracy (%)	TPR (%)	FPR (%)
LSTM[16]	93.75	93.10	11.02
GRU[16]	94.12	93.88	10.78
Autoencoder IDS [17]	94.81	94.20	10.15
CNN-LSTM [18]	95.68	95.05	9.42
Proposed Hybrid Model	96.92	96.16	8.59

V. DISCUSSION

While we believe our proposed hybrid model is very effective in identifying and classifying IoT attacks, it does have its limitations and drawbacks. Most importantly, the limitation is with the dataset itself. While the NSL-KDD dataset is widely accepted as a standard test, it is outdated, which means that some of the complexity and unique traffic patterns of modern large-scale IoT networks may be lost. Also, like many security datasets, it can suffer from diversity in the distribution of attack types, meaning that the proposed model might perform better on given classes of attacks that are more prevalent.

From a scalability perspective, the computational load on the proposed model is also a consideration. The procedure of reshaping input vectors into 2D matrices and running them through a CNN to extract the most relevant features is much more computationally intensive than applying a simple classifier to the raw data stream.

Adoption of such a model on resource-poor edge devices or gateways throughout an IoT ecosystem will therefore probably be excessively difficult from a latency and processing perspective. The provision of real-time detection systems across thousands of devices would therefore possibly require a large computational infrastructure

Finally, the K-Means clustering procedure is, while effective for pregrouping of data, sensitive to the specific initial selection of centroids and chosen number of clusters. An inadequate or inappropriate initialization would have a knock-on effect on the performance of subsequent classification by the deep learning model.

V. CONCLUSION AND FUTURE WORK

In this study, we proposed and validated a hybrid deep learning model for the detection and classification of security incidents in IoT environments. By usnig a CNN-based feature extractor with the discrimination ability of Random Forest and Decision Tree classifiers, our model achieved very high performance in terms of accuracy (96.92%) and a very low false positive rate (8.59%).

The results of the experimentation highlighted its effectiveness in detecting different types of attacks, such as DoS, Probe, and R2L, and showed an improved performance against many single deep learning models available.

However, despite these rather encouraging results, we see the limitations discussed and several clear directions for future work.

- First, to allow the model to be generalizable, it must be validated on more recent and larger performance datasets that are specific to IoT, such as ToN-IoT or Bot-IoT, which better represent traffic and attacks currently presented in smart environments.
- Second, it is very important that the question of scalability is addressed. Future works will consider techniques such as pruning and quantization to improve the concept and generate a light version of the model for application on edge devices which have limited computational resources.
- Finally, we will consider also an online learning framework, since it will allow to define dynamic modes that allow the model to learn and adapt to new attacks that have not been studied so far ("zero-day" attacks), without the need for the complete retraining of the model, allowing it to be more dynamic and more robust with regard to the solutions of security on the evolving IoT scenarios.

The results of the present research show the effectiveness of the hybrid approach to the context of IoT, in that it can strengthen its security much better, at the same time that we offer lines of improvement for the results obtained from the point of view of theoretical performance and practice applied to real situations.

REFERENCES

[1] D. Rani, N. S. Gill, and P. Gulia, "Classification of security issues and cyber attacks in layered internet of things," Journal of Theoretical and Applied Information Technology, vol. 100, no. 13, pp. 4895-4913, 2022.

[2] I. M. Anand, and S. Anand, "What Has IoT Got to Do with HR and People: A Case of Delloitte," Turkish Journal of Computer and Mathematics Education, vol. 12, no. 5, pp. 797-803, 2021.

[3] Y. G. Kim, B. Mendoza, O. Kwon, and J. Yoon, "Task-specific feature selection and detection algorithms for iot-based networks," Journal of computer and communications, vol. 10, no. 10, pp. 59-73, 2022.

[4] M. F. M. Sam, A. F. M. F. Ismail, K. A. Bakar, A. Ahamat, and M. I. Qureshi, "The effectiveness of IoT based wearable devices and potential cybersecurity risks: A systematic literature review from the last decade," International journal of online and biomedical engineering, vol. 18, no. 9, pp. 56-73, 2022.

[5] J. Roldán, J. Boubeta-Puig, J. L. Martínez, and G. Ortiz, "Integrating complex event processing and machine learning: An intelligent architecture for detecting IoT security attacks," Expert Systems with Applications, vol. 149, p. 113251, 2020.

[6] D. Li, Z. Cai, L. Deng, X. Yao, and H. H. Wang, "Information security model of block chain based on intrusion sensing in the IoT environment," Cluster computing, vol. 22, pp. 451-468, 2019.

[7] M. V. de Assis, L. F. Carvalho, J. J. Rodrigues, J. Lloret, and M. L. Proença Jr, "Near real-time security system applied to SDN environments in IoT networks using convolutional neural network," Computers & Electrical Engineering, vol. 86, p. 106738, 2020.

[8] I. Ullah, and Q. H. Mahmoud, "Design and development of a deep learning-based model for anomaly detection in IoT networks," IEEE Access, vol. 9, pp. 103906-103926, 2021.

[9] M. Anwer, S. M. Khan, and M. U. Farooq, "Attack detection in IoT using machine learning," Engineering, Technology & Applied Science Research, vol. 11, no. 3, pp. 7273-7278, 2021.

[10] U. Islam, A. Muhammad, R. Mansoor, M. S. Hossain, I. Ahmad, E. T. Eldin, and M. Shafiq, "Detection of distributed denial of service (DDoS) attacks in IOT based monitoring system of banking sector using machine learning models," Sustainability, vol. 14, no. 14, p. 8374, 2022.

[11] A. K. Sahu, S. Sharma, M. Tanveer, and R. Raja, "Internet of Things attack detection using hybrid Deep Learning Model," Computer Communications, vol. 176, pp. 146-154, 2021.

[12] A. Sagu, N. S. Gill, and P. Gulia, "Hybrid deep neural network model for detection of security attacks in IoT enabled environment," International Journal of Advanced Computer Science and Applications, vol. 13, no. 1, 2022.

[13] S. B. Atitallah, M. Driss, and I. Almomani, "A novel detection and multi-classification approach for IoT-malware using random forest voting of fine-tuning convolutional neural networks," Sensors, vol. 22, no. 11, p. 4302, 2022.

[14] M. B. Farukee, M. Z. Shabit, M. R. Haque, and A. S. Sattar, "DDoS attack detection in IoT networks using deep learning models combined with random forest as feature selector," in International Conference on Advances in Cyber Security, Singapore: Springer, 2020, pp. 118-134.

[15] J. Su, S. He, and Y. Wu, "Features selection and prediction for IoT attacks," High-Confidence Computing, vol. 2, no. 2, p. 100047, 2022.

[16] M. Roopak, G. Y. Tian, and J. Chambers, "An intrusion detection system against DDoS attacks in IoT networks," in 2020 10th annual computing and communication workshop and conference (CCWC), 2020, pp. 0562-0567.

[17] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, A. Shabtai, D. Breitenbacher, and Y. Elovici, "N-baiot—network-based detection of iot botnet attacks using deep autoencoders," IEEE Pervasive Computing, vol. 17, no. 3, pp. 12-22, 2018.

[18] M. A. F. Al-Husainy, and F. T. Al-Azawi, "A novel CNN-LSTM framework for attack detection in internet of things," Journal of Network and Computer Applications, vol. 165, p. 102697, 2020.