IOSR Journal of Mobile Computing & Application (IOSR-JMCA)
e- ISSN: 2394-0050, P-ISSN: 2394-0042. Volume 13, Issue 3 (May. – June. 2026), PP 01-09
www.iosrjournals.org

Improve Detecting Anomalies In Iot Networks Using Artificial Intelligence And Deep Recurrent Neural Networks

Zainab Aqeel, Zainb Qasem, Hussein Ali
(Wasit Directorate Of Education /The Ministry Of Education, Iraq)
(Wasit Directorate Of Education /The Ministry Of Education, Iraq)
(Wasit Directorate Of Education /The Ministry Of Education, Iraq)

Abstract:
Intrusion detection systems are used to find the best possible ways to protect IoT networks and the need for defense methods that detect attackers. Intrusion detection systems help the network to resist external attacks. However, it is difficult to use traditional IDS techniques in IoT due to its specific characteristics such as resource-intensive devices, limited specific protocol stacks, and standards. In fact, the new features and mechanisms of IoT cannot be provided by conventional security protocols used in the Internet. In the study deep learning algorithm, classification is done in two stages, in the first stage, preprocessing and classification is done in the recurrent neural network. The MinMaxScaller algorithm is used for normalization and preprocessing, the preprocessed features by this algorithm are given as input to the GRU network and this algorithm finally performs the classification task. According to the results obtained, the use of traditional machine learning methods achieves less accuracy than the use of deep learning models. As a result, the method used in this thesis was useful and the accuracy was 98.97%.

Keyword: *IOT, Feature Selection, Recurrent Neural Networks, GRU*

Date of Submission: 03-05-2026 Date of Acceptance: 13-05-2026

I. Introduction

In the last decade, with the increasing advancement of Internet of Things technology and the increase in communication between devices, the concept of the Internet of Things network has become one of the most important technological initiatives. This network consists of a complex infrastructure of Internet-connected objects that continuously send and receive information to each other. With the increase in the volume of IoT-based information, the importance of anomaly detection in this network has increased and plays a critical role in maintaining the security and proper functioning of IoT-connected systems (Ng et al., 2015).

In this context, deep recurrent models have been considered as one of the most widely used machine learning methods. These models have been very effective in detecting anomalies in temporal data, especially in the Internet of Things network, due to their ability to model temporal patterns and long-term dependencies.

In this research, we investigate and research the field of anomaly detection in the Internet of Things network using deep recurrent models. This research aims to investigate and evaluate new and efficient methods for detecting and predicting anomalies by considering the characteristics of temporal data.

In (2015) (Ahmad & Halim, 2017), Christian Cervantes and his colleagues used an intrusion detection system to detect hole attacks for the Internet of Things environment, which was implemented in the coja simulator. Their proposed system defines four components. The first module is the cluster configuration module, which is responsible for classifying a node. The second module is the routing monitoring module, in which the node monitors the number of transmissions made. The third module is the attacker detection module, which identifies the hole attack node. The fourth module is the attacker isolation module, which isolates the malicious node from the cluster and also generates an alert to notify its neighboring nodes. The simulation result has shown that their method achieves a detection rate of 92%. In (2018) (Aje et al., 2009), in a paper titled “Distributed attack detection scheme using deep learning approach for Internet of Things”, Chilmcourty et al. proposed a distributed deep learning-based detection system for the Internet of Things/fog and investigated the effectiveness of deep models compared to shallow models. The advantages of this research include the high accuracy of the deep model compared to the shallow model, the better performance of the distributed model compared to the centralized model, real-time and low false alarm rate, and its disadvantages include the examination of only four types of attacks, a long training time, and the use of many resources for the training phase. In an article titled "Machine learning ddos detection for consumer internet of things devices" (2018) (Imagej , 2018), Rouhani et al. used 4

steps to detect intrusion: 1-traffic capture, 2-grouping packets by devices and time, 3-feature extraction (stateless-satatefull) and 4-binary classification and algorithms of neural networks, random forests, K-nearest neighbors, support vector machines and decision trees. In this research, the decision tree classification was performed well and an accuracy of 99 was obtained, and the K-nearest neighbors classification also achieved the same accuracy, but the linear SVM classification performed the worst.

II. Research Problem And Objectives

Problem Statement

Smart cities enhance the quality of life of citizens in the city in terms of their energy and water consumption, healthcare requirements, environmental impacts, transportation demand and other important urban services. Smart cities are a complex combination of information and communication technology as well as the Internet of Things, which includes physical (such as sensors and actuators) and non-physical (such as external databases) data sources. Smart physical data sources contribute to the safety and security of citizens' daily lives through intelligent solutions such as human health application programs. These physical devices generate a huge amount of data, which is essential for analytical applications in user services. Cloud computing technologies are used as a centralized ICT architecture to collect a huge amount of data from the entire city scale and create citizen services by developers (Ng et al., 2012).

Data theft (Bala et al., 2022): IoT devices contain a lot of data, much of which is confidential and unique, including credit card details and personal health information, etc. A vulnerable device makes this data vulnerable to theft. Additionally, vulnerable devices can act as a gateway to intrusion into other systems, which is known in hacking as pivoting.

Physical damage (Alcalá Fernández & Alonso Moral, 2015): IoT devices are now widely used in the medical industry, for example, pacemakers, heart monitors, etc. These devices are also vulnerable to security threats. A vulnerable device can be exploited to interfere with a patient's medical care. Although this is very rare, some countries have recently developed malware that is capable of doing this.

In the research methodology and method section, the table of independent and dependent variables of this study and its definitions are given.

Purpose of the study

In this study, it is used to increase the accuracy of anomaly detection in the Internet of Things, so that the advantages of centralized and distributed technology can be used simultaneously to identify abnormal behavior. In this study, the accuracy of anomaly detection in the Internet of Things is investigated and increased with the help of deep learning. For this purpose, the GRU neural network is used to increase the accuracy of anomaly detection. Correlation calculation algorithms such as Pearson and K-Square are used to select features, which improves the accuracy in this type of network. In this study, deep learning and the introduced feature extraction and feature selection techniques will be used to solve the problems of basic references, which will increase the accuracy and speed of anomaly detection.

Objectives of the study

1. Increasing the accuracy of anomaly detection in the Internet of Things network at the edge level of the network and with the help of the GRU neural network.
2. Increasing the accuracy of anomaly detection in Internet of Things networks based on the use of the GRU neural network
3. Using correlation calculation algorithms such as Pearson and Chi-Square to select features and improve the accuracy of anomaly detection in Internet of Things networks.

III. Internet Of Things (IOT)

The Internet of Things (IoT) is a network of physical objects connected to the internet via current communication infrastructure in order to gather, process and exchange data. These are objects with networked digital devices and sensors that can detect information and then communicate it online. These connected devices, known as IoT objects, represent the interaction of sensors with connectivity, people and operational processes to create new applications and enable services (Viola & Jones, 2001)

IoT networks are usually created by connecting devices inside the same limited user environment, which is generally within a 10-meter proximity, and their topology may dynamically alter over time. IoT technology is implemented in various domains like healthcare, national security, business services, and recently to a lesser extent - smart home environments. Nevertheless, there is an exponential growth in the deployment of IoTs, which are more susceptible to security devices. Due to the fact they compose a variety of connected devices, cyberattacks are currently one of the most important concerns against IoT systems since they can target various resources. Consequently, ensuring the security of IoT devices and developing intrusion-resistant IoT networks to safeguard data has become a major research priority.

Security Challenges in IoT

Several challenges hinder the achievement of strong security in IoT environments. Because the concept of interconnected devices is relatively recent, security has not always been prioritized during the early stages of product design. In addition, the rapid growth of the IoT market has encouraged many manufacturers to focus on accelerating product release rather than implementing comprehensive security measures from the outset.

One of the most significant security concerns in IoT systems is the widespread use of default or hard-coded passwords, which can easily lead to unauthorized access and security breaches. Even when passwords are updated, they are often weak and insufficient to resist sophisticated attacks. Another major limitation is that many IoT devices operate under strict resource constraints, lacking the computational power and memory required to support advanced security mechanisms. Consequently, these devices often cannot provide robust protection against cyber threats (Imagej , 2018).

The architecture of the Internet of Things should be designed as an open architecture that relies on open protocols in order to support a wide range of existing network applications. It should also incorporate security, interoperability, and semantic middleware representation to enable global data integration with the Internet. However, there is no universally accepted IoT architecture agreed upon by all researchers, and several models have been proposed in the literature (Bruna & Mallat, 2013).

Some researchers describe IoT as a three-layer architecture, while others advocate a four-layer model. The latter argue that, with the rapid evolution of IoT technologies and applications, the traditional three-layer architecture is no longer sufficient to meet emerging requirements (Bruna & Mallat, 2013 ; Lo et al., 1995).

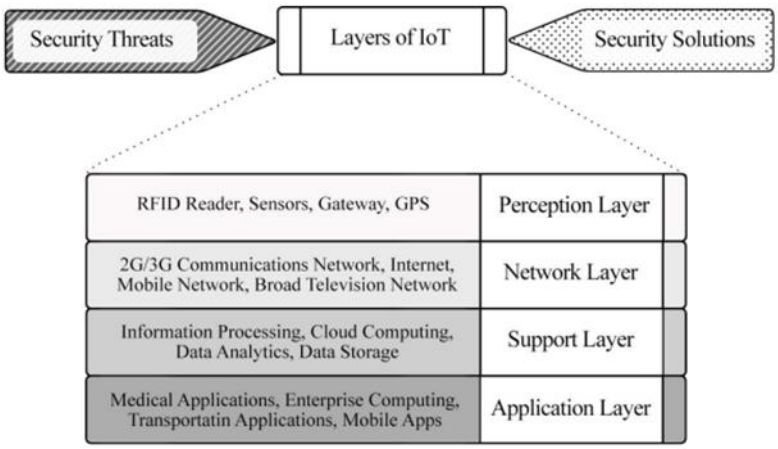

Figure (2-1) Internet of Things architecture (Bruna & Mallat, 2013)

Intrusion Detection System

An intrusion detection system is a device or software that monitors networks or detects and reports any unauthorized activity that disrupts the normal functioning of the system. In other words, intrusion detection is the process of identifying actions that attempt to compromise the integrity and reliability of a resource (Guan et al., 2014).

However, in fact, intrusion detection systems do not directly detect intrusion. In fact, these systems, by examining the ongoing activities, identify suspicious activities and introduce them as intrusions using algorithms or patterns that they contain. It is natural that some of these activities may not actually be intrusions and may simply be an unusual but harmless activity, and the system may have made a mistake in detecting intrusion. An example of an intrusion detection system is shown in Figure (2-2).

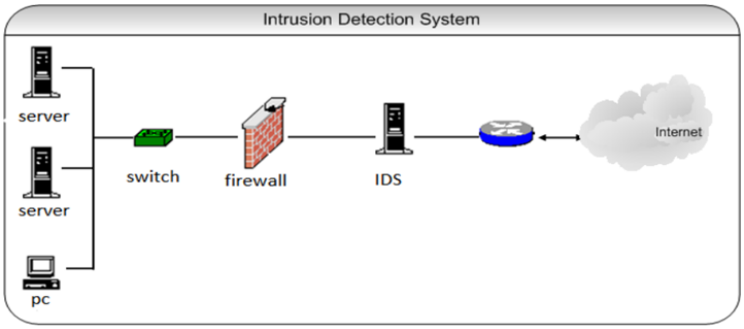

Figure (2-2) An example of an intrusion detection system (Guan et al., 2014)

Artificial Neural Networks

Artificial neural networks are derived from human neural networks. An artificial neural network is an idea for processing information and processes information like the brain. The key element of this idea is the new structure of the information processing system. This system consists of a large number of interconnected processing elements called neurons that work together in complete harmony to solve a problem. An artificial neural network is a computational model that relies on the framework and functionality of neural networks. Neural networks act as the main components of artificial intelligence (Wang et al., 2010). Neural networks consist of three layers under the headings of the input layer, the output layer, and the hidden layer. An artificial neural network is shown in Figure (2-3). As is clear, the network is made up of a number of processing units called neurons in multiple layers. The neurons in each layer are not connected to each other. To solve the problem using a neural network model, appropriate weights must be determined for each layer. The optimal weights for each layer are calculated using training data. Finally, the network is evaluated with test data (Mojumder et al., 2016).

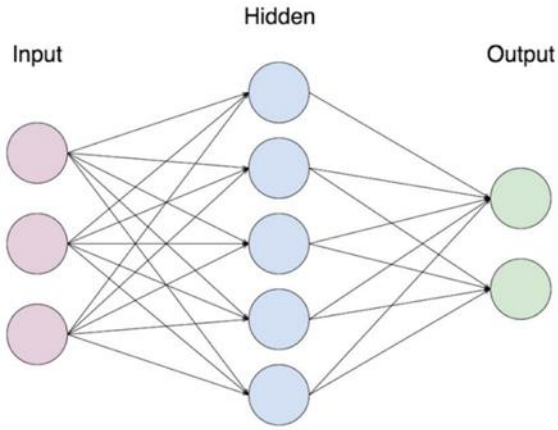

Figure (2-3) General model of artificial neural network (Wang et al., 2010)

GRU Neural Network

The Gated Recurrent Unit (GRU) neural network was introduced in 2014 by Cho et al (Bebis & Amayeh, 2015). as an architecture designed to address limitations found in traditional recurrent neural networks, particularly the vanishing gradient problem, while also reducing the computational complexity associated with the LSTM architecture. GRU is often viewed as a simplified variant of LSTM, as both share a similar conceptual design and frequently achieve comparable performance.

As previously discussed, GRU represents an enhanced form of the recurrent neural network. The key question is what distinguishes this architecture and why it was developed. Its primary objective is to mitigate the vanishing gradient problem in conventional RNNs. GRU accomplishes this through the use of two gating mechanisms: the update gate and the reset gate. These gates are implemented as vectors that control which information should be passed to the output and which should be discarded. A notable advantage of these gates is their ability to learn to preserve information from earlier time steps without unnecessary modification over time. To clarify this concept further, GRU networks are often compared with traditional neural networks.

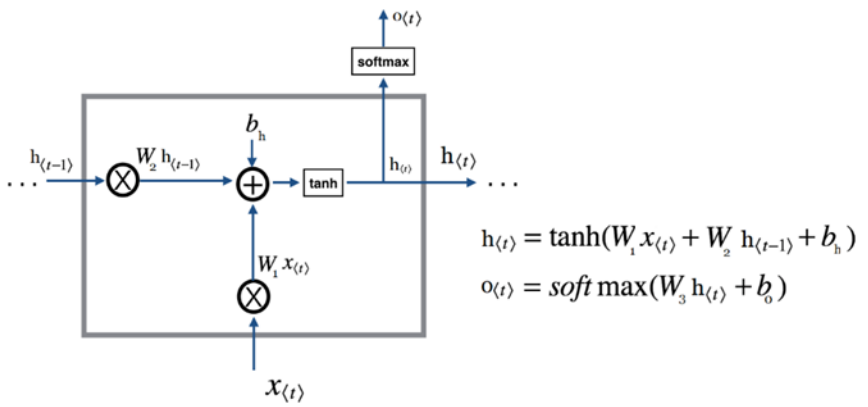

Figure (2-4) A typical RNN cell (Bebis & Amayeh, 2015).

IV. Methodology

Considering the research that has been conducted in the field of intrusion detection in the Internet of Things and the researchers have widely used machine learning and deep learning methods, the proposed method uses deep learning to detect intrusion in the Internet of Things. For this purpose, a hybrid deep learning method, which is actually a combination of two deep learning algorithms, GRU and Pearson, has been used.

In the first stage, preprocessing was performed on the data, which first removed outliers from the data set. Then, in order to use the data as input to the GRU network, it had to be normalized and standardized. The data was divided into three parts: training, validation, and testing. The training data was used to train the model, the validation data was used to test the model's fitness, and finally the test data was used to test the model, and the test set was not included in the training part. According to the reference method, the original data set was divided in such a way that 20% of the total data was considered as the test set and the remaining data, i.e., 80% of the data, was considered as the training set. The training set was again divided in such a way that 20% of the training set was considered as the validation set and the remaining data was considered as the training set. That is, 64% of the total initial data set was considered for model training, 16% was considered for model validation testing, and 20% was considered for testing. Then, the training data was given as input to the Pearson algorithm to select the desired features. Finally, with the help of the GRU deep learning algorithm and using the features extracted in the previous step, the classification operation was performed.

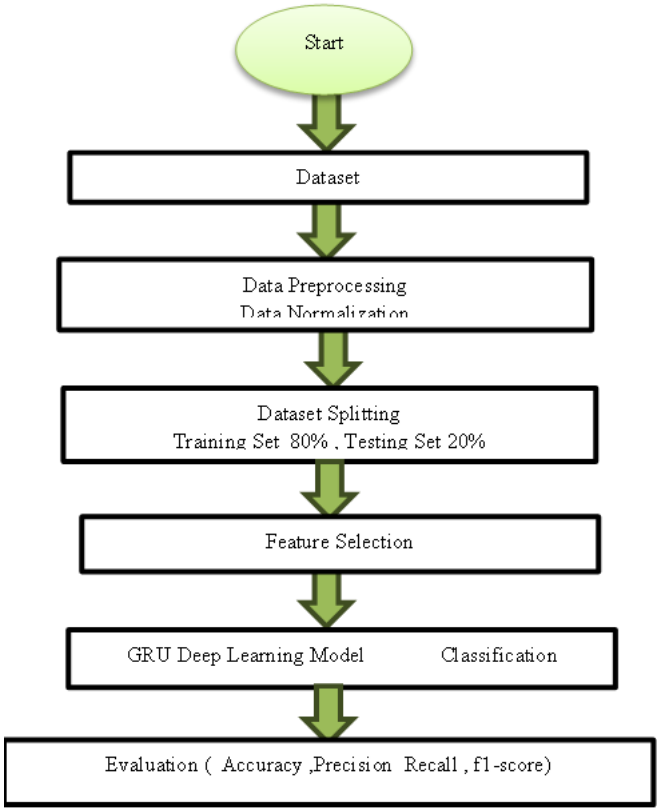

Figure (3-1) Block diagram of the proposed method

Dataset

The BOT-IOT dataset was used to investigate the effectiveness and feasibility of the proposed system. This dataset was created by designing a realistic network environment at the Cyber Range Laboratory of the UNSW Canberra Cyber Center, the latest version of which is from 2019. This dataset contains a mixture of normal and botnet traffic. In this dataset, there is one normal class and four attack classes, which are:

1. Probing attacks: These are malicious activities that collect information about victims by scanning remote systems. The probe can be active or passive. During passive probing, an attacker simply captures all the packets on the network; thus, it works in stealth.
2. DOS and DDOS: These are malicious activities that attempt to disrupt a service. These attacks are carried out by a group of machines called bots, and a remote machine usually targets the server. The goal of such attacks is to disrupt services accessible by legitimate users. These attacks can be divided into bulk and protocol-based attacks based on their attack method.

Data Preprocessing

The BOT-IOT dataset was used to investigate the effectiveness and feasibility of the proposed system. This dataset was created by designing a realistic network environment at the Cyber Range Laboratory of the UNSW Canberra Cyber Center. This dataset contains a combination of normal and botnet traffic. In the preprocessing performed for this dataset, redundant and duplicate records were removed from the dataset. Next, values from the dataset whose values are null were deleted. In the next step, variables that are of object type are known as strings in pandas. In order to convert these samples, the astype() method was used, which selects the appropriate data type. In the dataset used, since the samples are in the form of numbers, it converts them to numeric data. Then, the data was scaled before modeling using normalization. In fact, Sari normalization causes each input variable to be in the range of 10. The dataset was normalized using MinMaxScaller, which converts all values in a column to numbers between zero and one. In this method, the minimum value is subtracted from the current value and then the resulting value is divided by the range of values in the column. In the Scikit-Learn library, MinMaxScaller is used to perform this mathematical operation. Finally, the dataset is converted into three parts: training set, validation set, and test set. In fact, validating the performance of any learning model is an easy way to understand the data coming out of the model and make informed decisions about changes that should be made to the parameters or hyperparameters that affect the model.

GRU neural network training

In this study, the purpose of the proposed model is to detect intrusions in the Internet of Things using pre-processed information from previous stages. For this purpose, deep learning models, namely GRU, have been used. The pre-processed data were divided into three training, validation, and test sets to perform the proposed method. In this proposed model, 688,521 data samples from the BOT-IOT dataset were used, of which 20 data samples, including 137,705, were randomly assigned to the test set, and the remaining data, namely 80 data, including 550,816, were used for training. The training set was again divided into two subsets, namely the validation set and the training set, so that 20% of the training set was considered as the validation set and the rest as the training set. That is, from the total training dataset, the initial 64 were considered for model training and 16 for model validation testing, and 20 were considered for testing. Then, the training and validation sets were given to the model for training and validation. Pearson algorithm is used for feature selection and GRU algorithm is used for anomaly detection model.

V. Results

In this study, the proposed method is implemented using the Tensor Flow framework and the popular Keras library along with the Python programming language. This section consists of three parts: introducing the results obtained from the proposed model, comparing these results with previous work done in this field on different datasets, and introducing the tools used for implementation.

First, the tools used to implement the proposed method are introduced. These tools include Tensor Flow and Keras along with the Python programming language. It is explained how these tools were used to implement and test the proposed method and how they can transform the data into a suitable form for training the model and obtain the results.

Then, the results obtained from the proposed model in this study are presented. These results include various metrics such as accuracy, precision, and other model evaluation metrics that indicate the performance of the model in predicting or detecting different cases.

Evaluation criteria

The most important criteria used to evaluate the results are Accuracy, Precision, Recall and F-score, each of which is defined as follows.

Accuracy: In general, accuracy means how well the model predicts the output. By looking at the accuracy, one can immediately see whether the model is trained correctly or not and how effective it is in general. The accuracy criterion is described in relation (4-1).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad 4\text{-}1$$

Recall: The maximum value of this criterion is one or 100 percent and its minimum value is zero. The Recall criterion is also called the sensitivity criterion and its relationship is expressed according to equation (4-2).

$$Re\,c\,all = \frac{TP}{TP + FP} \qquad 4\text{-}2$$

Pr e cision: The maximum value of this criterion is one or 100 percent and its minimum value is zero. The more cases that the program has predicted incorrectly, which are called False Positives, than the correct predictions or True Positives, the lower the accuracy or Precision value will be, as shown in (4-3).

$$Pr\,e\,cision = \frac{TP}{TP + FP} \qquad 4\text{-}3$$

F1- score: The F1 criterion is a suitable criterion for evaluating the accuracy of the test. This criterion considers Precision and Recall together. The F1 criterion is one in the best case and zero in the worst case, as shown in equation (4-4) for its calculation.

$$F1 - Score = \frac{2 * TP}{2 * TP + FP + FN} \qquad 4\text{-}4$$

Evaluation of the accuracy of the proposed method

After compiling the model, the model is trained using the fit() function and then evaluated using the evaluate () function. After training, an accuracy of 98.99% was obtained.

The proposed GRU model, using appropriate preprocessing and appropriate configuration in the proposed neural network, has been able to achieve an accuracy of 98.97%. Next, the recall precision and fl-score have been calculated for each of the 3 existing classes. The results are shown in Table (3-4).

Table (4-3) Accuracy of each class in the test set

Dataset	f1-score	Precision	Recall
normal	95.70%	96.70 %	94.60%
DOS	99.75%	99.45 %	99.99%
DDOS	99.99%	99.98%	99.96 %
weighted avg	99.97%	99.97%	99.97%

After training the proposed model, the accuracy of the test dataset, which did not participate in the training process and had different data from validation and training, was obtained using the evaluate method, which was equal to 99.97%.

The accuracy obtained for each of the classes in the test dataset is as follows:

For the normal traffic class, out of 94 samples, the normal of the proposed model was able to recognize 89 samples, and the calculated precision, recall, and fl-score values were equal to 95.70%, 96.70%, and 94.60%, respectively. The results for the normal class are shown in Figure (4-1).

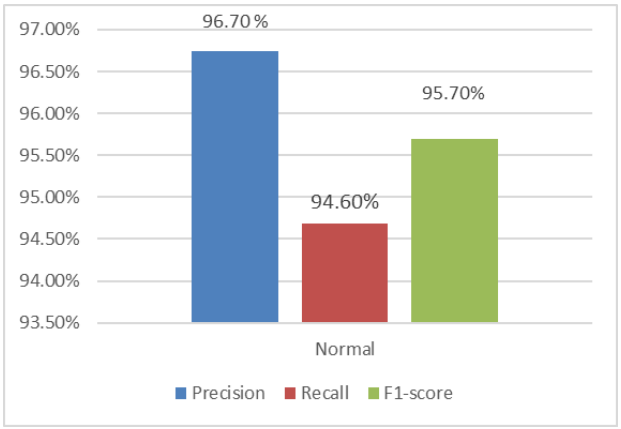

Figure (4-1) Normal class results chart

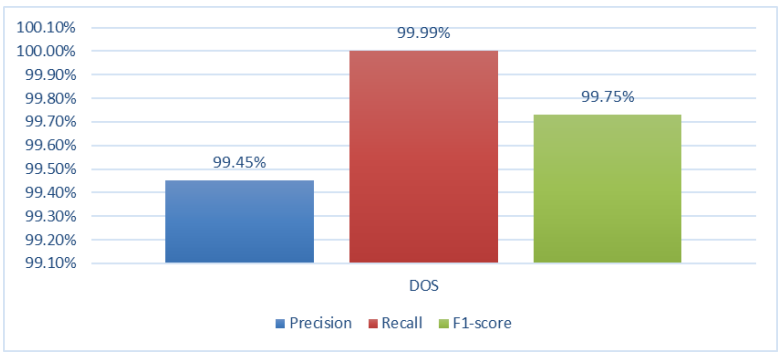

Figure (4-2) Results graph for the DOS class

For the DDOS class, out of 115,461 samples, the proposed model was able to recognize 115,438 samples and the calculated precision, recall and fl-score values were equal to: 99.99%, 99.45% and 99.97% respectively. The results graph for the DDOS class is shown in Figure (4-3).

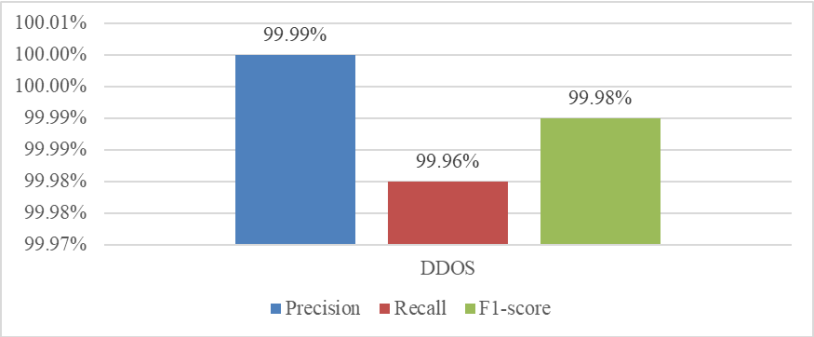

Figure (4-3) Results graph for the DDOS class

Evaluation of the efficiency of the proposed method compared of other method

A comparison of the proposed method and previous works is given, according to which for the BOT-IOT dataset, the proposed method given in the first row has achieved the highest recognition accuracy, and the recognition accuracy for the GRU algorithm is shown in the second row. The recognition accuracy in the cited sources that have used machine learning and deep learning algorithms is in the lower ranks. As shown in the table, using the combined deep learning method proposed in this thesis achieves better results than using a machine learning method or a deep learning method alone.

Method	Detection Rate
GRU (preposed Method)	99.97%
SVM (Cáceres-Matos et al., 2023)	99.11%
NN (Cáceres-Matos et al., 2023)	99.40%
KNN (Cáceres-Matos et al., 2023)	99.73%
RNN (Cabrerizo et al., 2015)	99.93%
LSTM (Cabrerizo et al., 2015)	99%
KNN (Cabrerizo et al., 2015)	97%
SVM (Cabrerizo et al., 2015)	78%
NB (Cabrerizo et al., 2015)	99.97%

VI. Conclusion

In the study, the goal is to use a method for detecting intrusions in the Internet of Things that can detect intrusions and network traffic with high accuracy and low error. This method should have acceptable performance. Considering that the new features and mechanisms of the Internet of Things cannot be provided by conventional security protocols used in the Internet, deep learning methods have been used. In the study deep learning algorithm, classification is done in two stages, in the first stage, preprocessing and classification is done in the recurrent neural network. The MinMaxScaller algorithm is used for normalization and preprocessing, the preprocessed features by this algorithm are given as input to the GRU network, and this algorithm finally performs the classification task. According to the results obtained, the use of traditional machine learning methods achieves less accuracy than the use of deep learning models. As a result, the method used in this thesis was useful and an accuracy of 98.97% was obtained.

References

[1]. Ng, C. B., Tay, Y. H., & Goi, B. M. (2015). A Review Of Facial Gender Recognition. Pattern Analysis And Applications, 18(4), 739-755.
[2]. Ahmad, H., & Halim, H. (2017). Determining Sample Size For Research Activities: The Case Of Organizational Research. Selangor Business Review, 20-34.
[3]. Aje, O. I., Odusami, K. T., & Ogunsemi, D. R. (2009). The Impact Of Contractors' Management Capability On Cost And Time Performance Of Construction Projects In Nigeria. Journal Of Financial Management Of Property And Construction, 14(2), 171-187.
[4]. Imagej. (2018). Available: Https://Imagej.Nih.Gov
[5]. Ng, C. B., Tay, Y. H., & Goi, B. M. (2012). Vision-Based Human Gender Recognition: A Survey. Arxiv Preprint Arxiv:1204.1611.
[6]. Bala, N., Gupta, R., & Kumar, A. (2022). Multimodal Biometric System Based On Fusion Techniques: A Review. Information Security Journal: A Global Perspective, 31(3), 289-337.
[7]. Alcalá Fernández, J., & Alonso Moral, J. M. (2015). A Survey Of Fuzzy Systems Software: Taxonomy, Current Research Trends, And Prospects.
[8]. Viola, P., & Jones, M. (2001, December). Rapid Object Detection Using A Boosted Cascade Of Simple Features. In Proceedings Of The 2001 IEEE Computer Society Conference On Computer Vision And Pattern Recognition. CVPR 2001 (Vol. 1, Pp. I-I). Ieee.

[9]. Bruna, J., & Mallat, S. (2013). Invariant Scattering Convolution Networks. IEEE Transactions On Pattern Analysis And Machine Intelligence, 35(8), 1872-1886.
[10]. Lo, S. C. B., Chan, H. P., Lin, J. S., Li, H., Freedman, M. T., & Mun, S. K. (1995). Artificial Convolution Neural Network For Medical Image Pattern Recognition. Neural Networks, 8(7-8), 1201-1214.
[11]. Guan, Y., Wei, X., & Li, C. T. (2014). On The Generalization Power Of Face And Gait In Gender Recognition. International Journal Of Digital Crime And Forensics (IJDCF), 6(1), 1-8.
[12]. Wang, J. G., Li, J., Lee, C. Y., & Yau, W. Y. (2010, December). Dense Sift And Gabor Descriptors-Based Face Representation With Applications To Gender Recognition. In 2010 11th International Conference On Control Automation Robotics & Vision (Pp. 1860-1864). IEEE.
[13]. Mojumder, U., Sarker, T. T., Monika, G. M., & Ratul, N. A. (2016). Vehicle Model Identification Using Neural Network Approaches (Doctoral Dissertation, BRAC University).
[14]. Bebis, G., & Amayeh, G. (2015). Hand-Based Biometric Analysis (No. Patent Application Number: US-Patent-Appl-SN-11/820,474).
[15]. Cáceres-Matos, R., Gil-García, E., Vázquez-Santiago, S., & Cabrera-León, A. (2023). Factors That Influence The Impact Of Chronic Non-Cancer Pain On Daily Life: A Partial Least Squares Modelling Approach. International Journal Of Nursing Studies, 138, 104383.
[16]. Cabrerizo, F. J., Morente-Molinera, J. A., Pérez, I. J., López-Gijón, J., & Herrera-Viedma, E. (2015). A Decision Support System To Develop A Quality Management In Academic Digital Libraries. Information Sciences, 323, 48-58.