PAGES: 174-187
9/9/23
JOURNAL OF SCIENTIFIC RESEARCH AND TECHNOLOGY(JSRT) VOLUME-1 ISSUE-6 SEPTEMBER
Registered under MSME Government of India ISSN: 2583-8660

Detection of Mirai Botnet Attacks on IoT devices Using Deep Learning

Antariksh Sharma[1], Prof. Vibhakar Mansotra[2] Kuljeet Singh[3]

[1]Student, Department of Computer Science & IT, University of Jammu, Jammu, India
antarikshsharma.cs@gmail.com

[2]Professor, Department of Computer Science & IT, University of Jammu, Jammu. India

[3]Scholar, Department of Computer Science & IT, University of Jammu, Jammu. India

ABSTRACT

The Internet of Things is a staple in the workplace, especially in the fields of office automation (OA) and operational technology (OT). As a consequence, businesses may set up a wide variety of IoT and IIoT gadgets across their operations. A setup like this makes areas with no prior cyber security concerns more vulnerable to assault. IoT devices in these shared spaces may have an effect on mission-critical systems like intranet and database servers due to the data collection and monitoring capabilities of the IoT systems. Thus, even hazards involving seemingly innocent IoT equipment like smart toilets and smart coffee makers may have a major effect, depending on the environment in which they are deployed. In light of this, it is important to consider the potential security issues that might lead to successful attacks on IoT systems and devices as part of any implementation of the IoT. These subtypes of the Mirai assault are known as ACK, SYN, Plain UDP, UDP flood, and Scan. These are the most important results from this study: The Dataset provides solid outcomes across the board and for each of the assaults it covers. Measures such as accuracy, precision, recall, and F1-score were employed to evaluate the datasets' reliability. Our findings on the N-BaIoT dataset show that the CNN model outperforms LSTM and GRU in terms of accuracy, precision recall, and f1 score. Improvements are possible by further refining these three methods. The development of a reliable approach to identify botnet assaults will need much future research and technological development. Because there are so few publicly accessible figures on IoT network traffic, it has been presumed that the vast majority of the traffic is loT network activity. Additionally, additional complications may appear while dealing with streaming data. Streaming-based learning needs further empirical investigation as a potential solution to this issue.

Keywords: Mirai Botnet, IoT, Deep learning, CNN, LSTM, GRU.

I. INTRODUCTION

1.1 OVERVIEW

The Internet of Things was once conceived as a way to link physical devices together through the web. Smart homes, smart workplaces, the smart grid, smart healthcare, smart agriculture, smart transportation, smart cities, etc., are just some of the many areas where the Internet of Things has had a transformative effect. According to a digital assessment conducted by McKinsey, the economic potential of IoT is large and increasing. By 2030, it might generate a global value of between $5.5 and $12.6 trillion [1]. The Internet of Things (IoT) is a core technology of the 21st century, connecting the real world with digital systems to boost efficiency, save costs, and free up more time for humans. A recent research found that the frequency of cyber assaults is increasing along with the number of susceptible IoT devices. It is possible to launch a DDoS or botnet assault. The most common kind of cyber attack witnessed recently is an assault, and their frequency and quantity have both increased over the last decade. DDoS assaults, the most common and prevalent kind, prevent authorised users from accessing resources. Despite the beneficial transformation, IoT's primary worry is security. Network attacks like denial of service (DoS) and spoofing [2] may also affect devices connected to the internet. Web apps and other software used by IoT devices might be exploited due to security flaws. Internet of Things (IoT) device deployment in mission-critical settings raises the stakes for cybercrime. If the IoT devices and apps are not adequately protected, a cyberattack on these vital facilities might have catastrophic consequences. The machines in a botnet are all infected with malware, and they're all linked together to be managed by one or more command and control servers.

Intruders send spam emails, commit click fraud, bring down websites with distributed denial of service (DDoS) assaults, etc. via the Botnet. While botnets have been around for some time, their size, complexity, and risk have

all increased because to the widespread use of unsecured IoT devices. Cybercriminal gangs may hijack internet-connected IoT devices and launch widespread assaults using them. Malware installed on IoT devices gives cybercriminals control over them, allowing them to use their computing power to launch distributed denial of service (DDoS) attacks against larger targets, send spam, steal information, and even conduct covert surveillance using IoT devices equipped with a camera or microphone. Attacks have also been carried out using a massive Botnet comprised of hundreds of thousands or even millions of IoT devices. Network security may be compromised by a large number of IoT devices that go undetected. And if the security system can't find the device, it can't quickly find the dangers to the device, too. These devices and their network connections are typically hidden from view by network security systems. One of the most malleable tools at our disposal right now is the internet of things (IoT). The IoT is adaptable and expandable because of the general availability of the internet, the growing speed and capacity of network connections, and the large range of objects that may be connected. Food production, manufacturing, finance, healthcare, and energy are just some of the industries that have been revolutionised by the Internet of Things. In particular, this is true of its offshoot, the IIoT (industrial IoT) [4]. Smart homes, buildings, and even whole cities have emerged as a direct consequence of this phenomenon. The possible implications of the Internet of Things (IoT) need to be recognised as its prevalence grows.

II. LITERATURE REVIEW

Husain et al. (2020) To improve detection of botnet attacks across different datasets, the authors of "Towards a Universal Features Set for IoT Botnet Attacks Detection" propose a universal feature set. When testing the trained machine learning models across three distinct botnet attack datasets, the suggested features set shows outstanding performance for identifying the assaults.

Noor et al. (2020) In order to efficiently and effectively detect attacks on IoT devices, the authors of the paper "Detecting Botnet Attacks in IoT Environments An Optimized Machine Learning Approach" propose an optimised ML-based framework based on a combination of the Bayesian optimization Gaussian Process (BO-GP) algorithm and the decision tree (DT) classification model. The Bot-IoT-2018dataset is used to test the effectiveness of the proposed framework. The experimental findings demonstrate the efficiency and resilience of the proposed optimised framework in detecting botnet assaults in IoT settings, with high detection accuracy, precision, recall, and F-score.

Singh et al. (2021) A 1D-CNN based model is presented in this study for the classification and analysis of network attacks. Using a specialised kind of convolutional neural networks called 1D-CNN, we can detect assaults by first differentiating between "normal" traffic and "attack data" (CNN). To do this, we use metrics like recall, accuracy, and F1-score to assess how well a model performs on the CICIDS2017 dataset, which contains 14 distinct attack types over 8 files. Individual sub-datasets and merged datasets were used to construct unique 1D-CNN based DL models. The model is further evaluated by contrasting its results with those obtained from an artificial neural network (ANN) simulation. The bulk of the class labels obtained outstanding scores in each of the assessment metrics, showing that the suggested model has performed better and showed significant capacity in identifying network assaults.

Alkahtani et al. (2021) This article presents a CNN-LSTM model for detecting botnet attacks in IoT settings. To identify botnet assaults, such as BASHLITE and Mirai, on nine commercial IoT devices, researchers in this study recommended a convolutional neural network and long short-term memory (CNN-LSTM) technique for hybrid deep learning. Extensive empirical study was conducted using an actual N-BaIoT dataset taken from a genuine system, and it included both benign and harmful patterns. While the proposed system achieved good accuracy (88.53%) in identifying botnet attacks from thermostat devices, experimental results revealed the superiority of the CNN-LSTM model with accuracies of 90.88 percent and 88.61 percent, respectively, in detecting botnet attacks from doorbells (Danminin and Ennio brands). With regard to accuracy metrics, the suggested system detected botnet assaults from security cameras with accuracies of 87.19%, 89.23%, 87.76%, and 89.64%, respectively. Overall, the CNN-LSTM model achieved state-of-the-art accuracy in identifying botnet assaults from a wide range of IoT gadgets.

Zewairi et al. (2022) Using supervised and shallow Deep learning classifiers, the authors of "Discovering unknown Botnet assaults on IoT devices" report on their findings. Investigate how well supervised shallow and deep learning classifiers can detect undiscovered botnet assaults on Internet of Things gadgets. Using a popular dataset, researchers examined the efficacy of shallow and deep supervised learning classifiers (i.e., the Aposemat IoT-23 dataset). Over the course of 1000 tests, we looked at the binary and multiclass classification

issue with respect to 12 unknown attack types and 38 unknown attack subtypes. Overall, the findings demonstrated a weighted average classification error rate that was rather high (61.46-86.40 percent), highlighting the need for new methods of detection of unknown threats.

Alyousfi et al. (2022) This research, titled "Preserving Location Privacy in the IoT against Advanced Attacks using Deep Learning," analyses the use of location-based services (LBS) to launch assaults on smart devices and discusses many methods for identifying these attacks. Since using an LBS necessitates transmitting the user's actual position in order to complete tasks, doing so leaves users vulnerable to privacy attacks. Some examples of assaults in LBS are Map Matching Attacks (MMAs) and Semantic Location Attacks (SLAs).

Idriss et al. (2021) This work implements and tests an intrusion detection system using a specialised Bot-IoT dataset to protect Internet of Things devices from common Botnet assaults. Our Bot IDS achieves encouraging results with 99.94% in validation accuracy, 0.58 % in validation loss, and a prediction execution time of less than 0.34 ms when compared to other deep learning approaches as simple RNN, LSTM, and GRU.

Sadaf et al. (2020) Using a deep learning methodology with an Autoencoder (AE) and an Isolation Forest (IF) for the fog environment, the authors of the study "Intrusion Detection Based on Autoencoder and Isolation Forest in Fog Computing" present a method (Auto-IF) for intrusion detection. Since distinguishing attack packets from regular ones in real time is of most significance to fog devices, our method focuses only on binary classification of the incoming traffic. Using the standard NSL-KDD dataset, we verify the efficacy of the suggested technique. When compared to numerous other state-of-the-art incursions detection techniques, our approach obtains a high accuracy rate of 95.4%.

Ivanova et al. (2020) This research article analyses network traffic using feedforward neural networks to identify IoT-based DDoS attacks. A model was proposed that may be used to defend against key logging, data exfiltration, OS fingerprinting, and service scans, as well as DoS and DDoS assaults including TCP, UDP, and HTTP flood. Such network traffic is easily distinguished from typical network flows. All neurons in the network's single hidden layer are activated by the hyperbolic tangent, and Adam optimization is used as the network's solver. The number of secret neurons might be adjusted to meet varying needs for precision and throughput. Extensive testing on the Bot IoT dataset demonstrates that models built with 8 or 10 characteristics work well.

Xie et al. (2022) When it comes to anomaly detection, the authors of the publication "IoT data analytics using deep learning" mix an LSTM-NN and an N-B model with a Gaussian distribution. The LSTM-Gauss-NBayes method produces respectable outcomes on three real-world datasets.

III. OBJECTIVES

The objectives of proposed work are as follows:

- Research and Analysis of Botnet Attacks on Internet of Things Devices.
- Examine and analyse the different Deep Learning methods currently used for Attack Detection.
- To deploy/construct a Machine Learning model for monitoring IoT networks for signs of intrusion.
- To educate and verify the Deep Learning model.
- Evaluation of the model is to be carried out on the test dataset.
- Investigate how well Deep Learning works in identifying malicious activity on networks.

IV. MATERIALS AND METHODS

4.1 Data Description:

4.1.1 Botnet Traffic in Datasets

All the data collected from a botnet and used to create a dataset is in the form of a comma-separated values (csv) file. All of the attributes are shown in the columns, and each row contains information on a single packet. Therefore, each row stands in for a different data packet, and each column represents a different set of characteristics for that data set. This is crucial because. Deep learning algorithms can read and process.csv files with relative ease.

4.1.2 DATASET

Since these datasets are freely accessible to the public, we choose to utilise them in our research. During the course of our literature review, we looked at many potential datasets. We created the N-BaIoT dataset to study

the spread of Mirai virus across a variety of Internet of Things (IoT) gadgets. Each unit included both raw traffic data, sometimes known as benign data, and infected traffic data.
A Dataset of N-BaIoT Devices When it comes to identifying IoT botnets, this dataset is a top contender. Many similar works have utilised this dataset for Mirai and Bashlite Malware detection, as we have shown. The University of California, Irvine, is the publisher of this dataset. This data collection includes 9 consumer-facing IoT devices that have been compromised by 2 distinct botnets. However, Mirai was the sole target for this particular effort. In this data collection, 9 different gadgets were utilised.
1. Danmini Doorbell
2. Ecobee Thermostat
3. Ennio Doorbell
4. Philips B120N10 Baby Monitor
5. Samsung SNH 1011 N Webcam
6. Provision PT 838 Security Camera
7. Provision PT 737E Security Camera
8. SimpleHome XCS7 1002 WHT Security Camera
9. SimpleHome XCS7 1003 WHT Security Camera

IoT devices, a command and control server, and a scanning and loading server were installed in a secure laboratory setting to begin the data gathering procedure for this dataset. After the network was turned on, raw traffic data was collected instantly. Port mirroring on the switch through which all traffic passes allowed it to be sniffed, and wireshark was used to capture the data. In this data collection, Mirai assaults such as Scan, Ack flooding, Syn flooding, UDP flooding, and UDP flooding with fewer choices were carried out. Figure 1 presents the fundamental structure of the dataset under consideration.

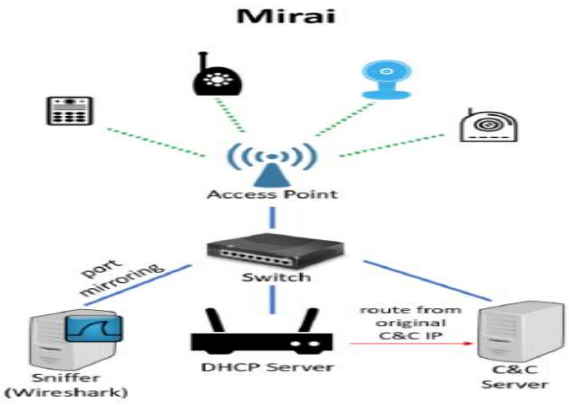

Figure 1 – Lab setup for detecting IoT botnet attacks [21]

Four out of the five Mirai virus variants launch DDoS assaults. Table 1 displays information about the available datasets.

Table 1 Type of Attack

Attacks	Description
Scan	Automatic scanning for vulnerable devices
ACK	Ack flooding
SYN	Syn flooding
UDP	UDP flooding
UDPplain	UDP flooding with fewer options, optimized for higher packets per second

Because the data size is too large for a system with a modest configuration, we selected just four of the nine smart devices available. Some of the absent attacks appear in the case of other gadgets. Table 2 displays the features of the chosen device and the attacks that were considered. In this case, N is the total number of occurrences.

Table 2 Description about data

Device Description	BENIGN (N)	SCAN (N)	ACK (N)	SYN (N)	UDP (N)	UDPplain (N)
PhilipsB120N10 (Baby Monitor)	175240	103621	91123	118128	217034	80808
ProvisionPT737E (SecurityCamera)	62154	96781	60554	65746	156248	56681
Damini (Doorbell)	49548	107685	102195	122573	237665	81982
Ecobee (Thermostat)	13113	43192	113285	116807	151481	87368

4.2 Deep Learning: Deep learning is a subfield of machine learning and AI that attempts to simulate the way that people learn. In the field of data science, which also covers fields like statistics and predictive modelling, deep learning plays a crucial role. Simply said, deep learning is a method for computerising predictive analytics. In contrast to the linear structure of classical ML algorithms, the deep learning algorithm stacks complexity and abstraction level upon level. Deep learning is a subfield of machine learning characterised by the use of multilayered neural networks. These neural networks "learn" from extensive datasets in an effort to mimic human brain activity, albeit they are still far from brain supremacy. A single-layer neural network can still produce approximations, but more complex networks with hidden layers may improve accuracy. Many AI apps and services rely on deep learning to boost automation by handling analytical and physical activities that previously required human participation. Digital assistants, voice-enabled TV remotes, and credit card fraud detection are just some of the common goods and services that rely on deep learning technology (such as self-driving cars). Below, we'll go through the three most common Deep Learning methods that have been put into practise:

- Convolutional Neural Network (CNN)
- Long Short Term Memory (LSTM)
- Gated Recurrent Unit (GRU)

V. RESEARCH METHODOLOGY AND PROPOSED MODELS

This study's suggested technique is an exact replication of the traditional Deep learning pipeline. The whole data set was then assessed using the suggested model. Figure 2 depicts a modelling process flowchart.

FLOW CHART: This flow chart shows the execution of models.

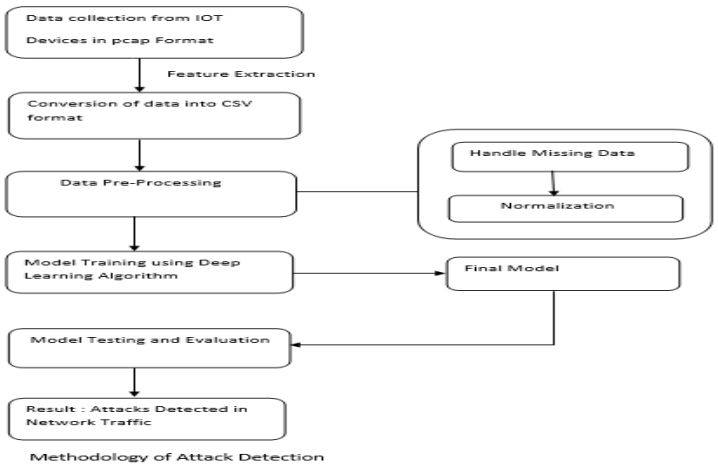

Figure 2 Flow Chart

5.5 Experimental Setup and Metrics

The hardware and software configuration is

Table 3 Model Configuration

Hardware setup	Software setup
i5 Processor	Google Collaboratory
RAM- 8GB	Spyder
Hard Disk- 1TB	Libraries

Table 4 Libraries used in models

Library	version	Explanation
NumPy	1.20.3	NumPy is an open-source numerical and popular Python library. It can be used to perform a variety of mathematical operations on arrays and matrices.
Pandas	1.3.4	Pandas are a data science and analysis Python library that allows developers to build intuitive and seamless high-level data structures.
TensorFlow	2.9.1	TensorFlow is a free and open-source Python library that specializes in differentiable programming. The library offers a collection of tools and resources that help make building DL and ML models
Keras	2.9.0	It also offers a fully functioning model for creating neural networks as it integrates with objectives, layers, optimizers, and activation functions.
Sci-kit Learn	0.24.2	It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistence interface in Python

VI. EXPERIMENTAL RESULTS

Here, we go through the findings of our tests with three distinct Deep learning approaches on the N-BaIoT dataset collected from a wide variety of devices. The results may be shown using the following metrics: F1 Score, Accuracy, Precision, Recall, and N (Test Samples).

6.1 N-BaIoT Dataset Result: The tables show the results of N-BaIoT dataset with different models.

Preformance assessment results for each attack type across all four devices are shown in Tables 5–8.

Table 5 shows the the performance evaluation of various attacks using LSTM in Device 1 Damini (Doorbell)

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	9928	96.80	95.99	96.79	95.89
Scan	21568	95.23	94.30	95.12	94.42
ACK	20485	95.62	94.32	95.64	95.00
SYN	24597	94.02	93.21	94.00	93.29
UDP	47324	93.11	92.95	93.08	92.98
UDP plain	16428	94.98	93.00	94.99	93.59

Table 6 shows the performance evaluation of various attacks using CNN model in Device 1.

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	9928	98.70	97.89	98.69	98.79
Scan	21568	98.13	97.20	98.22	98.12
ACK	20485	97.52	97.22	97.54	97.39
SYN	24597	97.52	97.30	97.54	97.43
UDP	47324	98.11	98.95	98.13	96.98
UDP plain	16428	98.79	96.63	98.80	97.99

Table 7 shows the performance evaluation of various attacks using GRU model in Device 1

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	9928	96.81	95.86	96.70	96.11
Scan	21568	95.20	94.32	95.11	94.21
ACK	20485	95.62	94.42	95.66	95.17
SYN	24597	95.02	94.21	95.00	94.26
UDP	47324	93.11	92.96	93.10	92.92
UDP plain	16428	95.98	93.11	95.99	94.45

Table 8 displays the performance evaluation of various attacks using LSTM model in Device 2 Provision PT-737E (Security Camera)

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	12470	94.96	94.48	94.98	94.61
Scan	19410	95.00	94.61	94.99	94.52
ACK	12101	94.99	94.95	94.98	94.97
SYN	13083	96.99	95.37	96.97	96.23
UDP	31211	97.00	96.12	97.99	96.66
UDP plain	11358	96.99	95.47	96.98	95.88

Table 9 displays the performance evaluation of various attacks using CNN model in Device 2

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	12470	98.70	97.89	98.69	98.79
Scan	19410	97.23	96.31	97.20	97.02
ACK	12101	97.52	97.22	97.54	97.39
SYN	13083	98.22	97.30	98.24	98.13
UDP	31211	98.11	98.95	98.13	96.98
UDP plain	11358	97.99	96.63	97.98	97.49

Table 10 displays the performance evaluation of various attacks using GRU model in Device 2

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	12470	96.58	95.98	96.55	96.29
Scan	19410	93.20	92.30	93.24	92.42
ACK	12101	95.62	94.32	95.64	95.00
SYN	13083	95.02	94.21	95.00	94.29
UDP	31211	93.22	92.95	93.28	92.78
UDP plain	11358	95.87	94.00	95.90	94.59

Table 11 displays the performance evaluation of various attacks using LSTM model in Device 3 Ecobee (Thermostat)

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	2634	94.86	94.48	94.90	94.63
Scan	8594	95.00	94.61	94.99	94.52
ACK	22595	93.99	92.65	93.98	93.27
SYN	23317	96.99	95.37	96.97	96.23
UDP	30539	95.00	94.12	95.99	94.66
UDP plain	17371	94.99	93.47	94.98	93.88

Table 12 displays the performance evaluation of various attacks using CNN model in Device 3

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	2634	98.50	97.89	98.59	97.99
Scan	8594	98.13	97.20	98.22	97.62
ACK	22595	97.52	97.22	97.54	97.39

SYN	23317	98.22	97.30	98.24	97.83
UDP	30539	98.11	98.95	98.13	96.96
UDP plain	17371	98.79	96.63	98.80	97.87

Table 13 displays the performance evaluation of various attacks using GRU model in Device 3

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	2634	96.66	96.20	96.58	96.30
Scan	8594	93.00	92.42	93.12	92.63
ACK	22595	95.62	94.30	95.64	95.00
SYN	23317	95.33	94.21	95.00	94.29
UDP	30539	93.20	92.95	93.22	92.78
UDP plain	17371	95.84	94.10	95.89	94.56

Table 14 displays the performance evaluation of various attacks using LSTM model in Device 4 Philips B120N10 (Baby Monitor)

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	35067	93.86	93.48	93.90	93.63
Scan	20884	95.00	94.61	94.99	94.52
ACK	18319	94.45	93.65	94.52	93.88
SYN	23557	96.90	95.37	96.97	96.23
UDP	43263	95.08	94.10	95.18	94.66
UDP plain	16101	94.99	93.47	94.98	93.88

Table 15 displays the performance evaluation of various attacks using CNN model in Device 4

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	35067	98.20	97.88	98.28	97.99
Scan	20884	98.13	97.20	98.22	97.62
ACK	18319	98.52	98.22	98.54	98.39
SYN	23557	98.20	97.40	98.12	97.73

UDP	43263	98.11	98.95	98.13	96.96
UDP plain	16101	98.79	96.63	98.80	97.87

Table 16 displays the performance evaluation of various attacks using GRU model in Device

Major Attacks	N	Accuracy	Precision	Recall	F1_score
Benign	35067	97.48	96.88	97.45	97.11
Scan	20884	95.20	94.30	95.24	94.42
ACK	18319	95.62	94.32	95.64	95.00
SYN	23557	95.02	94.21	95.00	94.29
UDP	43263	93.22	92.95	93.28	92.78
UDP plain	16101	95.84	94.43	95.88	94.99

Table17 displays the comparative analysis of attacks using CNN, LSTM and GRU models. The metrics used to show the attacks in these models is Precision, Recall, and F1 Score in the N-BaIoT data with a single device Danmini (Doorbell), for botnet infections using Mirai attacks.

DEVICE 1 (Doorbell)	Precision			Recall			F1_score		
	CNN	LSTM	GRU	CNN	LSTM	GRU	CNN	LSTM	GRU
Benign	97.89	95.99	95.86	98.69	96.79	96.70	98.79	95.89	96.11
Scan	97.20	94.30	94.32	98.22	95.12	95.11	98.12	94.42	94.21
ACK	97.22	94.32	94.42	97.54	95.64	95.66	97.39	95.00	95.17
SYN	97.30	93.21	94.21	97.54	94.00	95.00	97.43	93.29	94.26
UDP	98.95	92.95	92.96	98.13	93.08	93.10	96.98	92.98	92.92
UDP plain	96.63	93.00	93.11	98.80	94.99	95.99	97.99	93.59	94.45

Table 18 displays the comparative analysis of attacks using CNN, LSTM and GRU models. The metrics used to show the attacks in these models is Precision, Recall, and F1 Score in the N-BaIoT data file 5 with a single device Provision PT-737E (Security Camera) for botnet infections using Mirai attacks. In comparison of these models the metrics of all models have slightly changes.

Device 2 (Security Camera)	Precision			Recall			F1_score		
	CNN	LSTM	GRU	CNN	LSTM	GRU	CNN	LSTM	GRU
Benign	97.89	94.48	95.98	98.69	94.98	96.55	98.79	94.61	96.29
Scan	96.31	94.61	92.30	97.20	94.99	93.24	97.02	94.52	92.42
ACK	97.22	94.95	94.32	97.54	94.98	95.64	97.39	94.97	95.00
SYN	97.30	95.37	94.21	98.24	96.97	95.00	98.13	96.23	94.29
UDP	98.95	96.12	92.95	98.13	97.99	93.28	96.98	96.66	92.78
UDP plain	96.63	95.47	94.00	97.98	96.98	95.90	97.49	95.88	94.59

Table 19 displays the comparative analysis of attacks using CNN, LSTM and GRU models. The metrics used to show the attacks in these models is Precision, Recall, and F1 Score in the N-BaIoT data file 5 with a single device Ecobee (Thermostat) for botnet infections using Mirai attacks. In comparison of these models the metrics of all models have slightly changes.

Device 3 (Thermostat)	Precision			Recall			F1_score		
	CNN	LSTM	GRU	CNN	LSTM	GRU	CNN	LSTM	GRU
Benign	97.89	94.48	96.20	98.59	94.90	96.58	97.99	94.63	96.30
Scan	97.20	94.61	92.42	98.22	94.99	93.12	97.62	94.52	92.63
ACK	97.22	92.65	94.30	97.54	93.98	95.64	97.39	93.27	95.00
SYN	97.30	95.37	94.21	98.24	96.97	95.00	97.83	96.23	94.29
UDP	98.95	94.12	92.95	98.13	95.99	93.22	96.96	94.66	92.78
UDP plain	96.63	93.47	94.10	98.80	94.98	95.89	97.87	93.88	94.56

Table 20 displays the comparative analysis of attacks using CNN, LSTM and GRU models. The metrics used to show the attacks in these models is Precision, Recall, and F1 Score in the N-BaIoT data file 5 with a single device Philips B120N10 (Baby Monitor) for botnet infections using Mirai attacks. In comparison of these models the metrics of all models have slightly changes.

Device 4	Precision			Recall			F1_score		
	CNN	LSTM	GRU	CNN	LSTM	GRU	CNN	LSTM	GRU
Benign	97.88	93.48	96.88	98.28	93.90	97.45	97.99	93.63	97.11
Scan	97.20	94.61	94.30	98.22	94.99	95.24	97.62	94.52	94.42

ACK	98.22	93.65	94.32	98.54	94.52	95.64	98.39	93.88	95.00
SYN	97.40	95.37	94.21	98.12	96.97	95.00	97.73	96.23	94.29
UDP	98.95	94.10	92.95	98.13	95.18	93.28	96.96	94.66	92.78
UDP plain	96.63	93.47	94.43	98.80	94.98	95.88	97.87	93.88	94.99

Table 21 Evaluation matrix for all four different smart devices The metrics used to show the attacks in these three deep learning models .

Device	**Precision**			**Recall**			**F1_score**		
	CNN	**LSTM**	**GRU**	**CNN**	**LSTM**	**GRU**	**CNN**	**LSTM**	**GRU**
Device 1 (Doorbell)	97.53	93.96	94.15	98.15	94.94	95.26	97.78	94.20	94.52
Device 2 (Security Camera)	97.38	95.17	93.96	97.96	96.15	94.94	97.63	95.48	94.23
Device 3 (Thermostat)	97.53	94.12	94.03	98.25	95.30	94.91	97.61	94.53	94.26
Device 4 (Baby Monitor)	97.71	94.11	94.51	98.34	95.09	95.41	97.76	94.46	94.76

6.2 Model summary This section discusses model summary of various devices using three different deep learning Techniques.

```
cnn_model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 1, 60)             7020

 conv1d_1 (Conv1D)           (None, 1, 40)             2440

 dropout (Dropout)           (None, 1, 40)             0

 flatten (Flatten)           (None, 40)                0

 dense (Dense)               (None, 20)                820

 dense_1 (Dense)             (None, 6)                 126

=================================================================
Total params: 10,406
Trainable params: 10,406
Non-trainable params: 0
_________________________________________________________________
```

Figure 3 CNN model for Provision PT737E (Security Camera)

```
Lstm_model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 1, 92)             76912

 lstm_1 (LSTM)               (None, 1, 72)             47520

 lstm_2 (LSTM)               (None, 60)                31920

 dense (Dense)               (None, 6)                 366

=================================================================
Total params: 156,718
Trainable params: 156,718
Non-trainable params: 0
_________________________________________________________________
```

Figure 4 LSTM Model for ECOBEE (Thermostat)

```
GRU_model.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 gru (GRU)                   (None, 1, 90)             56160

 gru_1 (GRU)                 (None, 1, 80)             41280

 gru_2 (GRU)                 (None, 70)                31920

 dense (Dense)               (None, 6)                 426

=================================================================
Total params: 129,786
Trainable params: 129,786
Non-trainable params: 0
_________________________________________________________________
```

Figure 5 GRU Model for Damini (DOORBELL)

6.3 Observation: N-BaIoT includes data from nine different connected home IoT gadgets. Every single assault on Doorbell, Baby Monitor, Thermostat, and Security camera devices was uncovered by our model. Due to the sheer volume of data, we could only settle on four of the nine devices we considered. What if some of the missing assaults occurred on other devices? In addition, Bashlite and mirai assaults are separated into their own category in each dive. Both webcams and doorbells fell victim to the Mirai malware, although the latter was unable to compromise the latter. We use three distinct models—CNN, LSTM, and GRU—and they all provide varying outputs from the identical input data. The majority of the attacks in this dataset are DoS and DDoS assaults. More assaults with appropriate labelling are needed in the dataset so that accurate findings may be found.

6.4 Future Scope

The primary goal of this research was to conduct a comparative analysis of existing deep learning algorithms using a freshly released dataset. We want to continue developing this project in the future by amassing our own dataset to use in addressing the limitations of existing ones. Although the research shows that considerable effort has been put towards identifying botnet assaults early on, we believe there is still much room for improvement in this area. Additionally, several different deep learning techniques may be tested to boost botnet detection effectiveness..

VII. CONCLUSION

As the number of internet-connected devices that may be compromised continues to rise, botnet assaults have become a major concern for network safety. For the detection of botnet attacks, several machine learning-based algorithms have been published so far; however, only a few Deep Learning approaches have been used to detect Mirai botnet attacks on IoT devices. To mitigate the dangers of DDoS assaults on IoT devices, we developed a system based on a deep learning algorithm to better detection of Mirai botnet attacks. If DDoS assaults can be

detected in their early phases, network administrators may move more swiftly to cut off Internet access to the vast majority of IoT devices, therefore improving security and slowing the spread of botnets. In this study, we compared the performance of three different models of convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and genetic recurrent neural networks (GRUs) for detecting network attacks on IoT smart devices. To mitigate the dangers of DDoS assaults on Internet of Things (IoT) gadgets, we built a solution that makes use of deep learning algorithms. If DDoS assaults can be detected in their early phases, network administrators may move more quickly to cut off Internet access to the vast majority of IoT devices, therefore improving security and slowing the spread of botnets. In this study, we used the N-BaIoT dataset constructed from nine commercial IoT devices attacked by the BASHLITE and Mirai botnets: the Damini, Ennio, Ecobee, Phillips B120N/10, Provision PT-737E, Provision PT-838, Simple Home XCS7-1002-WHT, Simple Home XCS7-1003-WHT, and Samsung SNH1011N. The Mirai assault may be broken down into the following types: ACK, SYN, Plain UDP, UDP flood, and Scan. The study's most important conclusions are as follows: The Dataset provides both comprehensive and granular insights on the many threats it describes. The reliability of the dataset was evaluated using a number of different metrics. We found that the CNN model outperformed LSTM and GRU on the N-BaIoT dataset in terms of accuracy, precision recall, and f1 score. Better outcomes may be attained by further refinement of these three methods. A reliable approach to identify botnet assaults will need much research and technology in the future. Since there are so little publicly accessible facts on IoT network traffic, it has been presumed that the massive network traffic is loT network traffic. Moreover, other complications may arise while dealing with streaming data. Streaming-centric empirical research is required to better understand this issue.

REFERENCES:

1. Hussain, F., Abbas, S. G., Fayyaz, U. U., Shah, G. A., Toqeer, A., & Ali, A. (2020). Towards a universalfeatures set for IoT botnet attacks detection. In 2020 IEEE 23rd International Multitopic Conference (INMIC) (pp. 1-6). IEEE.
2. Injadat, M., Moubayed, A., &Shami, A. (2020). Detecting botnet attacks in IoT environments: An optimized machine learning approach. In 2020 32nd International Conference on Microelectronics (ICM) (pp. 1-4). IEEE
3. M. Al-Zewairi, S.A. lmajaliand M. Ayyash. Unknown security attack detection using shallow and deep ANN classifiers. Electronics, 9(12), 2022
4. Alkahtani, H., &Aldhyani, T. H. (2021). Botnet Attack Detection by Using CNN-LSTM Model for Internet of Things Applications. Security and Communication Networks, 2021.
5. Alazzam, H., Alsmady, A., &Shorman, A. A. (2019, December). Supervised detection of IoT botnet attacks. In Proceedings of the second international conference on data science, E-Learning and information systems (pp. 1-6).
6. Ahmed, Z., Danish, S. M., Qureshi, H. K., & Lestas, M. (2019, September). Protecting iots from mirai botnet attacks using blockchains. In 2019 IEEE 24th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD) (pp. 1-6). IEEE.
7. Sharmila, B. S., &Nagapadma, R. (2021). Multi Core DN N based IDS for Botnet Attacks using KPCA Reduction Techniques.
8. Abbas, S. G., Zahid, S., Hussain, F., Shah, G. A., &Husnain, M. (2020, December). A Threat Modelling Approach to Analyze and Mitigate Botnet Attacks in Smart Home Use Case. In 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE) (pp. 122-129). IEEE.
9. Alqahtani, M., Mathkour, H., & Ben Ismail, M. M. (2020). IoT botnet attack detection based on optimized extreme gradient boosting and feature selection. Sensors, 20(21), 6336.
10. Oliveira, S., Linhares, C., Travençolo, B., &Miani, R. (2020). Investigation of amplification-based DDoS attacks on IoT devices. INFOCOMP Journal of Computer Science, 19(1).
11. K. Singh, A. Mahajan and V. Mansotra.(2021) 1D-CNN based Model for Classification and Analysis of Network Attacks. International Journal of Advanced Computer Science and Applications, 12(11), pp. 604-613, 2021.
12. H. Alkahtani and T.H. Aldhyani.(2021) Botnet attack detection by using CNN-LSTM model for Internet of Things applications. Security and Communication Networks, pp.1-23, 2021.
13. M. Al-Zewairi, S.A. lmajaliand M. Ayyash.(2022) Unknown security attack detection using shallow and deep ANN classifiers. Electronics, 9(12), 2022.
14. A.S. Alyousef, K. Srinivasan, M.S. Alrahhal, M.A. lshammari and M. Al-Akhras.(2022) Preserving Location Privacy in the IoT against Advanced Attacks using Deep Learning. International Journal of Advanced Computer Science and Applications, 13(1), 2022.