

Research Methodology in Computer Science

Student Name: Ahmed Hasan Alhamdi

Student ID: 262001048

Course: Research Methodology in Computer Science

Instructor: Dr. Mohammad Anbar

Academic Year: 2026

AI-Based Detection of Botnet Attacks in Computer Networks

Table of Contents

Abstract

Chapter 1: Introduction

1.1 Introduction.....

1.2 Background.....

1.2.1 Botnet Attacks.....

1.2.2 Evolution of Botnet Detection.....

1.2.3 Artificial Intelligence in Botnet Detection.....

1.3 Problem Statement.....

1.4 Research Objectives.....

1.5 Research Scope.....

Chapter 2: Literature Review.....

2.1 Introduction.....

2.2 Bot-IoT Dataset and Network Forensic Analysis.....

2.3 Statistical Learning-Based Botnet Detection.....

2.4 Deep Learning Using Raw Network Traffic.....

2.5 Reliable Machine Learning for IIoT Botnet Detection.....

2.6 Detection of Unknown Botnets.....

2.7 Lightweight Hybrid Deep Learning.....

2.8 Hybrid Machine Learning Detection.....

2.9 Traditional Detection in IoT Environments.....

2.10 Research Gap.....

2.11 N-Balot: Network-Based Detection Using Deep Autoencoders.....	
2.12 Kitsune: Online Network Intrusion Detection.....	
2.13 CTU-13 and Empirical Botnet Detection Comparison.....	
2.14 IoT-23 Dataset and Real Malware Traffic.....	
2.15 Systematic Review of Machine Learning for IoT Botnet Detection.....	
2.16 Explainable Artificial Intelligence for Botnet Detection.....	
2.17 Recent Review of IoT Botnet and DDoS Detection.....	
2.18 Comparative Analysis of Previous Studies.....	
2.19 Literature Review Synthesis.....	
2.20 Identified Research Gap.....	

Chapter 3: Research Methodology

3.1 Research Design.....	
3.2 Proposed AI Botnet Detection Framework.....	
3.3 Dataset Selection.....	
3.4 Data Preparation.....	
3.5 Class Imbalance Handling.....	
3.6 Feature Selection.....	
3.7 Training and Testing Strategy.....	
3.8 Artificial Intelligence Models.....	
3.8.1 Random Forest.....	
3.8.2 Gradient Boosting / XGBoost.....	
3.8.3 Artificial Neural Network.....	
3.8.4 Deep Learning Model.....	

3.9 Model Evaluation.....	
3.10 Confusion Matrix.....	
3.11 Model Selection Criteria.....	
3.12 Proposed Experimental Tools.....	
3.13 Ethical Considerations.....	
3.14 Research Limitations.....	
3.15 Expected Results.....	
3.16 Expected Contribution.....	
3.17 Future Research.....	

Gantt Chart.....	
-------------------------	--

References.....	
------------------------	--

Conclusion.....	
------------------------	--

Abstract

Botnet attacks represent a significant cybersecurity threat because they enable attackers to remotely control large numbers of compromised computers, Internet of Things (IoT) devices, and other network-connected systems. These compromised devices can be coordinated to perform malicious activities such as Distributed Denial-of-Service (DDoS) attacks, information theft, spam distribution, network scanning, and unauthorized access. The increasing number and diversity of connected devices have made botnet detection more difficult, particularly when attackers modify their communication patterns or use previously unseen botnet variants.

Traditional botnet detection techniques frequently depend on predefined signatures, static rules, or manually selected indicators. Although these techniques can effectively identify previously known attacks, their ability to detect new and evolving botnets is limited. Research has therefore increasingly investigated Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) approaches capable of learning behavioral patterns from network traffic. Recent work includes lightweight CNN-LSTM architectures for resource-constrained IoT environments, zero-shot approaches for detecting previously unseen botnet families, and hybrid machine-learning models that combine several learning algorithms.

This research proposes an AI-based framework for detecting botnet attacks by analyzing network traffic characteristics and classifying traffic as legitimate or malicious. The study will use a quantitative experimental methodology and publicly available cybersecurity datasets, with the Bot-IoT dataset considered as a primary candidate. Bot-IoT was constructed in a realistic cyber-range environment containing both normal and botnet traffic and includes attacks such as DDoS, DoS, reconnaissance, keylogging, and data exfiltration.

Several AI classification algorithms will be evaluated using performance measures including accuracy, precision, recall, F1-score, false-positive rate, and confusion matrices. The expected contribution is a systematic comparison of AI techniques and the development of an effective detection framework capable of achieving high botnet detection performance while maintaining acceptable computational requirements.

Keywords: Artificial Intelligence, Botnet Detection, Cybersecurity, Machine Learning, Deep Learning, Network Intrusion Detection, Internet of Things.

Chapter 1: Introduction

1.1 Introduction

Modern computer networks have become increasingly interconnected as organizations and individuals rely on cloud services, mobile devices, smart systems, and Internet of Things technologies. Although this connectivity has created significant economic and technological benefits, it has also increased the number of systems that can potentially be compromised by cyber attackers.

One particularly serious cybersecurity threat is the botnet. A botnet is formed when multiple compromised devices are remotely controlled and coordinated by an attacker. Once compromised, these devices can collectively participate in malicious operations such as Distributed Denial-of-Service attacks, network scanning, data theft, spam campaigns, malware distribution, and other unauthorized activities. IoT devices are particularly attractive targets because large numbers of connected devices may have limited computing resources and may operate with security mechanisms that differ from those found on conventional computers. Research evaluating traditional botnet-detection systems has shown that techniques developed for conventional botnets may face difficulties when applied directly to IoT-based botnets.

The increasing sophistication of botnet behavior creates an important challenge for network security. Traditional intrusion-detection approaches often depend on attack signatures, fixed rules, blacklists, or predetermined thresholds. Such approaches remain useful for identifying known malicious behavior but may become less effective when the attacker changes communication patterns or when a previously unseen botnet appears.

Artificial Intelligence offers an alternative approach by enabling security systems to learn patterns from network traffic rather than relying exclusively on manually defined signatures. Machine-learning algorithms can identify statistical relationships between network-flow characteristics, while deep-learning architectures can discover more complex patterns from high-dimensional or sequential traffic data.

For example, research has proposed deep-learning-based network traffic analyzers that automatically derive useful characteristics from raw network data. Other work has explored zero-shot learning to identify botnet samples belonging to both previously seen and unseen families. These developments suggest that AI can potentially improve the adaptability of botnet-detection systems.

This research therefore investigates how artificial intelligence can be applied to detect botnet attacks in computer networks. The proposed work will examine existing AI-based approaches, evaluate selected machine-learning and deep-learning models, and propose a detection framework based on network traffic analysis.

1.2 Background

1.2.1 Botnet Attacks

Botnets consist of multiple compromised network-connected devices operating under the control of an attacker. A compromised device may appear to operate normally while simultaneously communicating with malicious infrastructure or participating in coordinated attacks.

Botnets can be used for several types of malicious activities, including:

- Distributed Denial-of-Service attacks.
- Unauthorized network scanning.
- Credential and information theft.
- Malware distribution.
- Spam campaigns.
- Keylogging.
- Data exfiltration.
- Coordinated attacks against network infrastructure.

The Bot-IoT research dataset illustrates the diversity of possible botnet-related activities. Its experimental environment combines legitimate and malicious traffic and contains categories including DDoS, DoS, operating-system and service scanning, keylogging, and data exfiltration. The full dataset contains more than 72 million records, while a smaller subset was also produced to make experimentation more manageable.

1.2.2 Evolution of Botnet Detection

Early botnet detection approaches relied heavily on signatures and manually defined behavioral indicators. These techniques search network communications for characteristics previously associated with malicious software.

Although signature-based detection can provide efficient recognition of known threats, modern botnets can change network behavior, use alternative communication mechanisms, or produce traffic that resembles legitimate activity. IoT botnets also introduce additional challenges because IoT protocols and network architectures may differ from those used by conventional computing platforms. Woodiss-Field et al. experimentally examined traditional tools including BotMiner, BotProbe, and BotHunter in the context of IoT botnets, illustrating the importance of evaluating whether established botnet-detection assumptions remain valid in newer environments.

As a result, researchers increasingly investigate behavioral and data-driven detection techniques.

1.2.3 Artificial Intelligence in Botnet Detection

Machine learning allows a detection system to construct classification models based on examples of legitimate and malicious network traffic. Algorithms such as Random Forest, Support Vector Machine, Artificial Neural Networks, gradient-boosting methods, and other classifiers can be trained using network-flow characteristics.

Deep learning extends this capability by using multiple computational layers to automatically learn complex representations from data. Researchers have proposed architectures combining Convolutional Neural Networks and Long Short-Term Memory networks to capture both local and sequential characteristics of botnet behavior.

Alshehri et al. proposed SkipGateNet, a lightweight CNN-LSTM architecture intended to identify Mirai and Bashlite botnets while considering the resource limitations of IoT environments.

Ali et al. proposed a hybrid machine-learning system combining ANN, CNN, LSTM, and RNN components. Their work emphasizes the potential benefit of combining different learning architectures rather than relying on a single classifier.

Another important direction concerns unknown attacks. Luu et al. investigated a zero-shot deep-learning approach designed to detect botnet samples from both previously observed and unseen botnet families on unknown devices.

These studies demonstrate the growing importance of AI in botnet detection but also indicate unresolved challenges concerning generalization, computational efficiency, dataset dependence, and detection of novel attacks.

1.3 Problem Statement

Botnet attacks continue to pose a significant challenge to computer and IoT networks because compromised devices can generate coordinated malicious traffic while remaining difficult to distinguish from legitimate network activity.

Conventional botnet detection systems commonly depend on known signatures, manually configured rules, or predefined behavioral patterns. These techniques can perform effectively when the malicious pattern has previously been identified; however, their effectiveness may decrease when botnets modify their behavior, operate on new device types, or belong to previously unseen attack families.

Recent research demonstrates promising results from machine-learning and deep-learning approaches. However, different studies use different datasets, network characteristics, feature-selection techniques, and AI models, which makes it difficult to determine which approach provides the most appropriate balance among detection accuracy, false-positive rate, generalization, and computational requirements.

The ability to identify previously unseen botnets also remains an important research problem. Zero-shot research has explicitly addressed detection of unseen botnet families, demonstrating that generalization beyond attacks present in the training data is a distinct requirement rather than simply achieving high classification performance on known samples.

Therefore, the central research problem addressed in this study is:

How can Artificial Intelligence be used to develop an accurate and efficient botnet attack detection framework that distinguishes legitimate network traffic from malicious botnet activity while reducing false-positive rates and maintaining the ability to respond to evolving attack behavior?

1.4 Research Objectives

The main objective of this research is to investigate and develop an AI-based approach for detecting botnet attacks in computer networks.

The specific objectives are:

Objective 1

To analyze the characteristics and security impact of botnet attacks in modern computer and IoT networks.

This objective will examine botnet architecture, attack behavior, network communication characteristics, and common malicious activities associated with compromised devices.

Objective 2

To critically review existing traditional, machine-learning, and deep-learning approaches for botnet detection.

The study will compare existing methods and identify their major advantages, limitations, and research gaps.

Objective 3

To develop an AI-based botnet detection framework using network traffic features.

The framework will include data preprocessing, feature analysis, AI model training, classification, and performance evaluation.

Objective 4

To compare selected machine-learning and deep-learning algorithms for botnet traffic classification. Candidate algorithms may include:

- Random Forest.
- XGBoost or another gradient-boosting classifier.
- Artificial Neural Network.
- Deep Neural Network.
- CNN/LSTM-based models where computationally feasible.

Objective 5

To evaluate the proposed framework using standard cybersecurity classification metrics.

Performance will be examined using:

- Accuracy.
- Precision.
- Recall.
- F1-score.
- False Positive Rate.
- Confusion Matrix.
- ROC-AUC where appropriate.

Objective 6

To identify a suitable balance between detection performance and computational efficiency.

This objective is particularly important for environments containing IoT or edge devices with limited computing resources.

1.5 Research Scope

This research focuses on AI-based detection of botnet attacks through the analysis of network traffic. The study will primarily investigate network-based detection rather than host-based malware analysis. It will therefore concentrate on traffic characteristics that can be observed from network flows or packet-derived features.

A publicly available labelled cybersecurity dataset will be used for experimentation. The Bot-IoT dataset is considered the primary candidate because it contains a combination of legitimate and botnet traffic collected in a realistic cyber-range environment. Its attack categories include DDoS, DoS, scanning, keylogging, and data exfiltration.

Where appropriate, datasets such as N-BaIoT may be discussed or used for additional evaluation. Research using N-BaIoT has employed it to validate machine-learning approaches for IIoT botnet detection, and the dataset contains a large feature set for network-based analysis.

The research will concentrate on supervised AI classification, although anomaly detection and zero-shot methods will be examined in the literature review because of their relevance to unknown attacks.

The study does not aim to develop offensive botnet software, execute attacks against real external networks, or compromise real devices. All experimentation will use authorized research datasets and controlled analysis environments.

Chapter 2: Literature Review

2.1 Introduction

Artificial-intelligence-based botnet detection has become an active area of cybersecurity research because conventional approaches frequently struggle with the scale, diversity, and changing behavior of modern network attacks. Existing work can broadly be grouped into dataset development, classical machine learning, statistical anomaly detection, deep learning, lightweight detection, hybrid learning, and approaches designed for unknown attacks.

2.2 Bot-IoT Dataset and Network Forensic Analysis

Koroniotis et al. developed the Bot-IoT dataset to provide researchers with realistic legitimate and botnet traffic for cybersecurity experimentation. The environment was constructed in the UNSW Canberra Cyber Range Lab and generated both benign and malicious traffic. The dataset includes DDoS, DoS, reconnaissance, keylogging, and data-exfiltration activities.

The importance of this work extends beyond providing data. AI-based cybersecurity models strongly depend on the quality and representativeness of their training data. A labelled dataset such as Bot-IoT enables researchers to train and compare different classification techniques under reproducible conditions.

A limitation, however, is that performance obtained using a controlled research dataset may not necessarily represent performance in every operational network. Therefore, evaluation should consider model generalization rather than accuracy alone.

2.3 Statistical Learning-Based Botnet Detection

Ashraf et al. proposed IoTBoT-IDS, a statistical-learning-based anomaly-detection framework intended to detect botnets in smart-city networks. The approach demonstrates that botnet detection does not necessarily require conventional signature matching and can instead be based on statistical patterns derived from network behavior.

This type of detection is valuable because anomaly-oriented methods can potentially identify suspicious behavior that has not been explicitly encoded in a signature database. Nevertheless, anomaly-detection methods must carefully control false-positive rates because unusual legitimate behavior can also appear statistically abnormal.

2.4 Deep Learning Using Raw Network Traffic

Shahhosseini et al. proposed a deep-learning-based network traffic analyzer for botnet detection that automatically extracts useful characteristics from raw traffic data.

Automatic feature learning is an important advantage of deep neural networks because manually designing features requires domain expertise and may overlook complex relationships.

On the other hand, deep-learning systems can require greater computational resources and larger training datasets than conventional machine-learning classifiers.

2.5 Reliable Machine Learning for IIoT Botnet Detection

Taher et al. investigated a machine-learning approach for Industrial Internet of Things botnet detection and used the N-BaIoT dataset for validation. The dataset used in their evaluation contains 115 features.

Their research highlights the role of feature analysis and model reliability in intrusion detection. Rather than assuming that more input features automatically produce better detection, effective cybersecurity models must determine which characteristics provide meaningful discriminatory information.

2.6 Detection of Unknown Botnets

Luu et al. proposed a zero-shot deep-learning approach intended to detect botnet samples from both seen and unseen families and on unknown IoT devices.

This study addresses one of the major weaknesses of traditional supervised classification. A conventional model learns from labelled examples belonging to known categories, but operational cybersecurity systems must frequently deal with attack variants that were absent from training data.

The research therefore demonstrates the importance of evaluating AI systems not only according to their ability to reproduce training labels but also according to their ability to generalize to new threats.

2.7 Lightweight Hybrid Deep Learning

Alshehri et al. proposed SkipGateNet, which combines CNN and LSTM components with learnable skip connections for botnet attack detection. The architecture specifically targets efficient detection of Mirai and Bashlite attacks in resource-constrained IoT environments.

The work is particularly relevant because highly accurate cybersecurity models may have limited practical value when their computational requirements exceed the capacity of the systems on which they must operate.

2.8 Hybrid Machine Learning Detection

Ali et al. introduced a hybrid machine-learning approach that combines ANN, CNN, LSTM, and RNN models. The work was published in IEEE Access in 2024 and demonstrates an ensemble-oriented strategy for improving botnet attack identification.

Hybrid approaches can potentially exploit the strengths of different learning architectures. However, their increased model complexity may also increase processing time and memory requirements. Therefore, performance improvement must be considered together with implementation cost.

2.9 Traditional Detection in IoT Environments

Woodiss-Field et al. investigated traditional botnet-detection systems including BotMiner, BotProbe, and BotHunter and examined how effectively methods originally developed for conventional botnets could identify IoT-based bots.

This work is important because it demonstrates that botnet detection techniques cannot automatically be assumed to transfer successfully from traditional computing environments to IoT ecosystems. Network protocols, device behavior, traffic patterns, and resource constraints may differ substantially.

2.10 Research Gap

The reviewed literature demonstrates considerable progress in AI-based botnet detection. Machine-learning models provide efficient classification, deep-learning techniques can automatically learn complex representations, hybrid models combine multiple architectures, and zero-shot methods attempt to detect previously unseen threats.

However, several important research issues remain.

First, many studies are optimized and evaluated on specific datasets. High performance on a single dataset does not necessarily demonstrate effective generalization to other network environments.

Second, sophisticated deep-learning and hybrid models can impose significant computational requirements. This issue is particularly important for IoT and edge environments.

Third, detection of unknown or evolving botnets remains more difficult than classification of known attacks. The existence of dedicated zero-shot botnet research illustrates that this challenge remains distinct from ordinary supervised classification.

Fourth, cybersecurity evaluation must consider more than overall accuracy. A highly imbalanced dataset can produce misleading accuracy values, making precision, recall, F1-score, false-positive rate, and confusion matrices essential for evaluating practical detection performance.

Based on these gaps, this research proposes a comparative AI-based botnet detection framework that evaluates multiple learning algorithms using common preprocessing and evaluation procedures. This will enable a clearer assessment of the trade-off between classification effectiveness and computational complexity.

Chapter 3: Research Methodology

3.1 Research Methodology

This research will use a quantitative experimental methodology.

The experimental approach is appropriate because the research aims to train AI models using labelled network traffic and quantitatively compare their ability to distinguish legitimate traffic from botnet activity.

The methodology will consist of the following phases:

Phase 1: Literature Review

Relevant journal articles, conference papers, cybersecurity datasets, and technical publications will be reviewed.

The literature review will focus on:

- Botnet architecture and behavior.
- Traditional botnet detection.
- Machine-learning-based detection.
- Deep-learning-based detection.
- Hybrid models.
- Lightweight intrusion detection.
- Detection of unknown botnets.
- Feature-selection approaches.
- Evaluation metrics.

Phase 2: Dataset Selection

The Bot-IoT dataset will be considered as the primary experimental dataset.

It contains legitimate and simulated botnet traffic produced in the UNSW Canberra Cyber Range Lab and provides PCAP, Argus, and CSV representations.

A reduced dataset subset may be used when necessary to maintain practical computational requirements.

Phase 3: Data Preprocessing

The dataset will undergo preprocessing before model training.

The preprocessing procedure may include:

1. Removing duplicate records.
2. Identifying missing values.
3. Removing irrelevant identifiers.
4. Converting categorical variables into numerical representations.
5. Normalizing or standardizing numerical features when required.
6. Examining class imbalance.
7. Dividing the data into training, validation, and testing subsets.

A stratified split will be preferred to preserve the proportion of malicious and legitimate traffic.

Phase 4: Feature Selection

Feature-selection techniques will be investigated to reduce unnecessary input variables and computational complexity.

Potential approaches include:

- Correlation analysis.
- Mutual information.
- Recursive Feature Elimination.
- Feature importance derived from tree-based classifiers.

Reducing redundant features may improve training efficiency and facilitate deployment on resource-constrained systems.

Phase 5: Model Development

Several AI models will be trained and compared.

The proposed baseline models are:

Random Forest

Random Forest will provide a strong classical machine-learning baseline and can naturally model nonlinear relationships between network characteristics.

XGBoost / Gradient Boosting

A gradient-boosting model will be evaluated because boosting methods are frequently effective when analyzing structured tabular cybersecurity data.

Artificial Neural Network

A feed-forward neural network will be implemented to evaluate the effectiveness of nonlinear neural classification.

Deep Learning Model

Depending on available computational resources, a deeper neural network or CNN/LSTM-based approach will be evaluated to determine whether representation learning provides a significant advantage.

Phase 6: Model Evaluation

The trained models will be evaluated using the independent test set.

The main evaluation metrics will include:

Accuracy

Accuracy measures the proportion of correctly classified samples.

Precision

Precision measures the proportion of predicted attacks that are actually malicious.

Recall

Recall measures the proportion of actual attacks that are successfully detected.

F1-Score

F1-score provides a harmonic balance between precision and recall.

False Positive Rate

False-positive rate measures how frequently legitimate traffic is incorrectly classified as malicious.

Confusion Matrix

The confusion matrix will provide detailed information about true positives, true negatives, false positives, and false negatives.

Where appropriate, ROC-AUC will also be used.

Phase 7: Comparative Analysis

The models will be compared according to:

- Classification performance.
- False-positive rate.
- Training time.
- Prediction time.
- Number of selected features.
- Computational complexity.

The final framework will be selected based on the overall balance among these factors rather than accuracy alone.

3.2 Expected Results

The proposed research is expected to demonstrate that AI-based techniques can effectively distinguish legitimate traffic from botnet-related network activity.

The expected outcomes include:

1. Identification of network characteristics that provide the greatest contribution to botnet detection.
2. Development of a structured AI-based botnet detection workflow covering data preprocessing, feature selection, training, classification, and evaluation.
3. Quantitative comparison of classical machine-learning and neural-network models.
4. Identification of the algorithm providing the best balance between detection rate and computational efficiency.

5. Reduction of false-positive classifications compared with poorly optimized baseline approaches.
6. Improved understanding of the limitations of AI-based botnet detection, particularly regarding unknown attacks and dataset generalization.
7. Practical recommendations for future implementation of AI-assisted network intrusion detection systems.

The research does not assume in advance that a specific model will achieve the best results. Final conclusions will be based on experimental evaluation.

Preliminary References

- [1]N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779-796, 2019.
- [2]J. Ashraf, M. Keshk, N. Moustafa, M. Abdel-Basset, H. Khurshid, A. D. Bakhshi, and R. R. Mostafa, "IoTBoT-IDS: A novel statistical learning-enabled botnet detection framework for protecting networks of smart cities," *Sustainable Cities and Society*, vol. 72, 103041, 2021.
- [3]M. Shahhosseini, H. Mashayekhi, and M. Rezvani, "A deep learning approach for botnet detection using raw network traffic data," *Journal of Network and Systems Management*, 2022.
- [4]F. Taher, M. Abdel-Salam, M. Elhoseny, and I. M. El-Hasnony, "Reliable machine learning model for IIoT botnet detection," *IEEE Access*, vol. 11, pp. 49319-49336, 2023, doi: 10.1109/ACCESS.2023.3253432.
- [5]C. D. Luu, V. Q. Nguyen, and T. S. Pham, "A zero-shot deep learning approach for unknown IoT botnet attack detection," in *2023 RIVF International Conference on Computing and Communication Technologies*, 2023, doi: 10.1109/RIVF60135.2023.10471793.
- [6]M. S. Alshehri, J. Ahmad, S. Almakdi, M. Al Qathrady, Y. Y. Ghadi, and W. J. Buchanan, "SkipGateNet: A lightweight CNN-LSTM hybrid model with learnable skip connections for efficient botnet attack detection in IoT," *IEEE Access*, vol. 12, pp. 35521-35538, 2024, doi: 10.1109/ACCESS.2024.3371992.
- [7]M. Ali, M. Shahroz, M. F. Mushtaq, S. Alfarhood, M. Safran, and I. Ashraf, "Hybrid machine learning model for efficient botnet attack detection in IoT environment," *IEEE Access*, vol. 12, pp. 40682-40699, 2024, doi: 10.1109/ACCESS.2024.3376400.
- [8]A. Woodiss-Field et al., "Examination of traditional botnet detection on IoT-based networks," *Sensors*, vol. 24, no. 3, 1027, 2024.

Chapter 2: Literature Review - Continued

2.11 N-BaloT: Network-Based Detection Using Deep Autoencoders

Meidan et al. introduced an important network-based approach for detecting IoT botnet attacks using deep autoencoders. Their work focused on identifying anomalous traffic generated by compromised IoT devices rather than relying exclusively on known malicious signatures.

The researchers experimentally infected nine commercial IoT devices with two major IoT botnets, Mirai and BASHLITE, and collected network behavior from both legitimate and compromised states. Deep autoencoders were then trained to learn the normal behavioral patterns of each device. Deviations from the learned patterns were considered potential malicious activities.

One of the main strengths of this approach is that it uses anomaly detection rather than conventional attack-signature matching. This potentially enables the system to identify previously unknown malicious behavior. The study also demonstrates that device-specific traffic characteristics can provide useful indicators of compromise. (arXiv)

However, a limitation is that separate behavioral models may be required for different categories of devices. In large and heterogeneous IoT networks, maintaining individual detection profiles may increase operational complexity.

This study is highly relevant to the proposed research because it demonstrates the potential of neural-network-based anomaly detection and provides a benchmark dataset that can be used to evaluate AI-based botnet detection systems.

2.12 Kitsune: Online Network Intrusion Detection

Mirsky et al. proposed Kitsune, an online network intrusion detection system based on an ensemble of autoencoders known as KitNET.

Unlike many supervised learning approaches that require large quantities of labelled attack data, Kitsune is designed to learn normal network behavior in an unsupervised manner. It processes network traffic continuously and uses several small autoencoders rather than a single large neural network.

The architecture was designed with computational efficiency in mind. The authors demonstrated that Kitsune could operate even on limited hardware such as a Raspberry Pi while detecting a range of malicious network behaviors. (arXiv)

This represents an important advantage for IoT and edge environments because network gateways may not have sufficient processing capacity to execute large deep-learning architectures.

However, as with most anomaly-detection systems, determining the threshold separating unusual legitimate activity from malicious behavior remains a challenge. Changes in normal network behavior can potentially increase false-positive rates.

For the present research, Kitsune highlights two important design objectives:

1. The AI detector should achieve strong classification performance.
2. The computational requirements of the detector should remain practical.

2.13 CTU-13 and Empirical Botnet Detection Comparison

García et al. developed and evaluated botnet detection techniques using what became known as the CTU-13 dataset.

The dataset consists of thirteen different botnet scenarios and includes three principal categories of network traffic:

- Botnet traffic.
- Normal traffic.
- Background traffic.

Their research compared different botnet detection methods under common experimental conditions and emphasized the importance of using realistic, labelled datasets when evaluating security systems (ساینس دایرکت).

A significant contribution of the study is its demonstration that no single detection method necessarily performs best during every stage of botnet activity. Some detection techniques may identify communication or command-and-control behavior effectively, while others may perform better during attack execution.

The study therefore supports the argument that botnet detection should not be evaluated using accuracy alone. Detection performance should also consider the type of attack behavior being observed and the cost associated with false classifications.

One limitation of CTU-13 for contemporary IoT research is its age. The dataset was generated before many recent IoT botnets emerged. Nevertheless, it remains academically valuable because it provides realistic malicious, normal, and background network traffic.

2.14 IoT-23 Dataset and Real Malware Traffic

Garcia, Parmisano, and Erquiaga introduced the IoT-23 dataset, which provides labelled network traffic generated by both malicious and legitimate IoT devices.

The dataset contains 23 scenarios: 20 malware captures and three benign IoT traffic captures. Malware traffic was generated from infected IoT devices, while benign scenarios were obtained from real consumer IoT equipment. (Stratosphere Laboratory)

IoT-23 is particularly valuable because it includes realistic malicious behavior rather than relying entirely on synthetic traffic generation.

The availability of detailed flow labels enables researchers to study specific malicious activities such as command-and-control communication and attack execution. As a result, the dataset provides another potential source for validating the generalizability of a model trained on Bot-IoT.

However, IoT-23 is relatively large. The full dataset contains hundreds of millions of network flows, meaning that preprocessing and model training may require substantial storage and computing resources. (Stratosphere Laboratory)

For this research, IoT-23 may therefore be considered as an optional secondary dataset or a reduced sample may be used for additional validation.

2.15 Systematic Review of Machine Learning for IoT Botnet Detection

Nazir et al. conducted a systematic review investigating the use of machine learning and deep learning for IoT botnet detection.

The researchers initially examined 1,567 records and ultimately included 25 primary studies published between 2018 and 2023. Their analysis considered machine-learning algorithms, deep-learning approaches, benchmark datasets, evaluation measures, and preprocessing techniques .
(ساینس دایرکت)

The review concluded that AI techniques show significant potential for detecting IoT botnet attacks and can outperform conventional signature-based approaches in many experimental environments.

An important finding, however, was that reported performance varies according to:

- The dataset used.
- The selected features.
- Data-preprocessing procedures.
- The learning algorithm.
- Evaluation metrics.

This finding supports the methodology of the present research because the proposed work will compare multiple models using the same dataset, preprocessing procedure, and evaluation metrics.

Such controlled comparison is necessary to determine whether observed performance improvements result from the AI model itself rather than differences between experimental settings.

2.16 Explainable Artificial Intelligence for Botnet Detection

Saied and Guirguis investigated the use of Explainable Artificial Intelligence (XAI) for IoT botnet detection.

The researchers argued that achieving high detection accuracy alone may not be sufficient for practical cybersecurity systems because network administrators must also understand why an AI model classifies traffic as malicious.

Their proposed approach integrates several explainability techniques, including:

- SHAP.
- LIME.
- Rule extraction and model distillation.

The study applied explainability techniques to a Random Forest-based botnet detector and demonstrated how individual network characteristics could be interpreted to understand the model's decisions. (Nature)

This direction is particularly important because many advanced AI techniques behave as black-box models.

In a security operations environment, an alert such as:

"Botnet Detected: 98% probability"

provides less actionable information than an explanation such as:

"The traffic was classified as malicious because of unusually high connection frequency, repeated destination behavior, abnormal packet statistics, and scanning characteristics".

Explainability can therefore improve trust in the detection system and support analysts when investigating alerts.

A limitation is that additional explainability mechanisms introduce extra processing and implementation complexity.

Nevertheless, XAI represents an important future extension for the proposed framework.

2.17 Recent Review of IoT Botnet and DDoS Detection

A systematic literature review published in Sensors in 2024 examined IoT botnet DDoS attacks and the detection techniques proposed to counter them.

The review emphasizes that IoT devices remain attractive botnet targets partly because resource constraints can limit the implementation of conventional security mechanisms. It categorizes existing botnet detection approaches and identifies unresolved research problems associated with evolving attacks. (MDPI)

The study reinforces an important observation for the current research: an effective IoT security solution must balance three competing objectives:

Detection performance + adaptability + computational efficiency

A model with extremely high accuracy but excessive computational demands may be unsuitable for practical IoT deployment.

2.18 Comparative Analysis of Previous Studies

The reviewed literature indicates that AI-based botnet detection can be approached using several different learning strategies.

Study	Technique	Dataset/Environment	Main Strength	Main Limitation
García et al.	Behavioral botnet detection	CTU-13	Realistic comparative evaluation	Older botnet scenarios
Meidan et al.	Deep Autoencoders	N-BaloT	Unsupervised anomaly detection	Device-dependent models
Mirsky et al.	Ensemble Autoencoders	Kitsune	Lightweight online detection	Anomaly threshold sensitivity
Koroniotis et al.	ML/DL analysis	Bot-IoT	Large realistic IoT dataset	Dataset imbalance
Shahhosseini et al.	Deep Learning	Network traffic	Automatic feature learning	Higher computational cost
Taher et al.	Machine Learning	N-BaloT	Extensive feature analysis	Dataset-specific validation
Luu et al.	Zero-Shot Deep Learning	IoT botnet traffic	Detection of unseen botnets	Increased model complexity
Alshehri et al.	CNN-LSTM	IoT traffic	Lightweight architecture	More complex than classical ML
Ali et al.	ANN-CNN-LSTM-RNN Ensemble	UNSW-NB15	Strong ensemble performance	Computational overhead
Nazir et al.	Systematic Review	25 studies	Broad comparison of ML/DL	No new experimental detector

Saied & Guirguis	Random Forest + XAI	N-BalIoT	Explainable predictions	Explainability overhead
Recent Sensors Review	Multiple methods	Multiple studies	Identifies current research gaps	Review rather than experimental model

2.19 Literature Review Synthesis

The literature reveals several consistent patterns.

First, machine learning and deep learning have demonstrated considerable potential for recognizing botnet-related network behavior.

Second, the quality and characteristics of the dataset strongly influence reported performance.

Third, deep-learning algorithms can automatically discover complex network patterns but generally require more computational resources.

Fourth, lightweight AI models are becoming increasingly important because many IoT and edge systems operate with limited processing capacity.

Fifth, detecting previously unseen attacks remains substantially more difficult than detecting attacks represented in the training dataset.

Sixth, a highly accurate AI detector is not necessarily a practical detector. False-positive rates, prediction speed, resource consumption, and interpretability must also be considered.

The present study therefore proposes a comparative framework that evaluates multiple AI techniques under consistent experimental conditions.

2.20 Identified Research Gap

Based on the reviewed literature, four major research gaps are identified.

Gap 1: Dataset Dependency

Many proposed botnet detection systems report excellent performance on one benchmark dataset but provide limited evidence of generalization to other network environments.

Gap 2: Accuracy Versus Efficiency

Deep-learning and hybrid models can provide strong classification results, but their computational complexity can limit deployment on IoT gateways and edge systems.

Gap 3: Detection of Unknown Attacks

Most supervised-learning systems are optimized for botnet traffic observed during training. Previously unseen botnet families remain challenging.

Gap 4: Model Interpretability

Highly complex models may classify attacks accurately without providing understandable explanations for their decisions.

Consequently, this study focuses on identifying an AI model that offers a practical balance among:

Detection\ Accuracy]

False\ Positive\ Rate]

Computational\ Efficiency]

Generalization]

rather than optimizing only a single metric.

Chapter 3: Research Methodology

3.1 Research Design

This research adopts a quantitative experimental research design.

The experimental method is selected because the primary aim is to evaluate whether selected artificial-intelligence algorithms can accurately differentiate between legitimate network traffic and botnet traffic.

The independent variable is the AI classification algorithm, while the primary dependent variables are:

- Accuracy.
- Precision.
- Recall.
- F1-score.
- False-positive rate.
- ROC-AUC.
- Training time.
- Prediction time.

3.2 Proposed AI Botnet Detection Framework

The proposed framework consists of seven primary stages:

Network Dataset



Data Cleaning



Data Preprocessing



Feature Selection



AI Model Training



Traffic Classification



Performance Evaluation

The framework receives labelled network traffic records as input and outputs a classification indicating whether traffic is legitimate or associated with malicious botnet activity.

3.3 Dataset Selection

The primary dataset proposed for this research is Bot-IoT.

Bot-IoT was developed by researchers at UNSW Canberra using a cyber-range environment combining legitimate and malicious network traffic. The complete dataset is large, and its authors also provide a 5% subset containing approximately three million records to facilitate experimentation (ساینس دایرکت).

The dataset includes malicious activities such as:

- DDoS.
- DoS.
- Reconnaissance.
- Service scanning.
- Keylogging.
- Data exfiltration.

The reduced dataset is preferred for the initial experiment because it provides sufficient observations while reducing computational requirements.

3.4 Data Preparation

Prior to training the AI models, the dataset will undergo a structured preprocessing process.

Step 1: Data Inspection

The dataset will be inspected to identify:

- Missing values.
- Duplicate records.
- Data types.
- Class distribution.
- Categorical variables.
- Numerical variables.

Step 2: Data Cleaning

Duplicate observations and records containing unusable information will be removed where appropriate.

Step 3: Feature Removal

Identifiers that are unlikely to provide meaningful behavioral information may be removed.

Examples may include:

- Raw IP identifiers.
- Unique flow IDs.
- Timestamp identifiers where they create data leakage.

Care will be taken not to remove behavioral information required for attack classification.

Step 4: Categorical Encoding

Categorical network characteristics will be transformed into numerical values using methods such as:

- One-hot encoding.
- Label encoding.

Step 5: Numerical Scaling

Algorithms that are sensitive to the magnitude of numerical variables will use normalization or standardization.

For example:

$$z = \frac{x - \mu}{\sigma}$$

[

where:

- x represents the original feature value.
- μ represents the feature mean.
- σ represents its standard deviation.

3.5 Class Imbalance Handling

Cybersecurity datasets frequently contain unequal distributions of normal and malicious traffic.

An imbalanced dataset can produce misleading accuracy.

For example, if:

99 %of traffic = malicious

= 1% legitimate

a model predicting every sample as malicious could obtain 99% accuracy despite having no useful discrimination capability.

Therefore, class distribution will be analyzed before training.

Potential approaches include:

- Random undersampling.
- Random oversampling.
- Class weighting.
- SMOTE where appropriate.

Performance will then be evaluated using precision, recall, and F1-score in addition to accuracy.

3.6 Feature Selection

Feature selection will be performed to identify the network characteristics that contribute most strongly to botnet classification.

Possible techniques include:

Correlation Analysis

Highly correlated redundant variables may be identified and reduced.

Mutual Information

Mutual information will estimate how strongly each feature contributes to predicting the target label.

Random Forest Feature Importance

A Random Forest model can rank variables according to their contribution to classification decisions.

Recursive Feature Elimination

Recursive Feature Elimination may be applied to gradually remove low-value variables.

Feature reduction can potentially decrease:

- Training time.
- Memory usage.
- Prediction latency.
- Overfitting.

3.7 Training and Testing Strategy

The dataset will be divided into independent subsets.

A potential division is:

- 70% Training Data
- 10% Validation Data
- 10% Testing Data

Alternatively:

- 80% Training Data

- 20% Testing Data

with cross-validation conducted on the training set.

A stratified sampling strategy will be used where possible so that each subset retains representative class proportions.

The test dataset will remain isolated during model training.

3.8 Artificial Intelligence Models

3.8.1 Random Forest

Random Forest combines multiple decision trees to produce a final classification.

Advantages include:

- Strong performance on tabular data.
- Ability to model nonlinear relationships.
- Feature-importance estimates.
- Limited preprocessing requirements.
- Relatively high interpretability compared with deep neural networks.

Random Forest will therefore serve as an important baseline.

3.8.2 Gradient Boosting / XGBoost

Gradient boosting develops multiple weak decision models sequentially, with each model attempting to correct errors made by previous models.

XGBoost is commonly effective for structured datasets and may provide strong performance with relatively efficient prediction.

The algorithm will be compared against Random Forest to determine whether boosting produces a significant improvement.

3.8.3 Artificial Neural Network

A feed-forward Artificial Neural Network will be developed.

A possible architecture is:

Input Layer

↓

Dense Layer



ReLU



Dropout



Dense Layer



Output Layer

For binary classification, the final layer may use the sigmoid activation function:

]

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

[

The resulting probability represents the estimated likelihood that the traffic is malicious.

3.8.4 Deep Learning Model

If computational resources permit, a deeper neural architecture will also be evaluated.

Potential models include:

- Deep Neural Network.
- 1D CNN.
- CNN-LSTM.

The purpose is not simply to use the most complicated architecture but to determine whether the additional complexity produces a meaningful improvement over simpler models.

3.9 Model Evaluation

Accuracy

]

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

[

Precision

]

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

[

High precision means that when the model generates a botnet alert, the prediction is usually correct.

Recall

]

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

[

High recall is important in cybersecurity because false negatives represent malicious traffic that the detector failed to identify.

F1-Score

]

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

[

F1-score balances precision and recall.

False Positive Rate

]

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

[

A low false-positive rate is operationally important because excessive false alerts can overwhelm cybersecurity analysts.

3.10 Confusion Matrix

Model classification will also be represented through a confusion matrix.

	Predicted Normal	Predicted Botnet
Actual Normal	True Negative	False Positive

Actual Botnet	False Negative	True Positive
----------------------	----------------	---------------

3.11 Model Selection Criteria

The final AI model will not automatically be selected according to the highest accuracy.

Instead, the following factors will be considered together:

]

Final\ Score= $f(A, P, R, F1, FPR, T, C)$

[

where:

- (A) = Accuracy.
- (P) = Precision.
- (R) = Recall.
- (F1) = F1-score.
- (FPR) = False-positive rate.
- (T) = Prediction time.
- (C) = Computational complexity.

This produces a more practical evaluation of cybersecurity performance.

3.12 Proposed Experimental Tools

The experiment may be implemented using:

- Python.
- Pandas.
- NumPy.
- Scikit-learn.
- Matplotlib.
- TensorFlow or PyTorch.
- Jupyter Notebook.

Pandas and NumPy will support data preparation.

Scikit-learn will be used for classical machine-learning models and evaluation metrics.

TensorFlow or PyTorch may be used for neural-network implementation.

3.13 Ethical Considerations

Cybersecurity research involving botnets requires particular ethical care because the subject relates to malicious software and unauthorized network activity.

This research will not create an operational botnet or compromise real external systems.

Only legitimate academic datasets and controlled computing environments will be used.

The research will follow the following principles:

1. No unauthorized access to computer systems.
2. No execution of malicious attacks against external networks.
3. No collection of private user network traffic.
4. Use only authorized or publicly available research data.
5. Report findings for defensive cybersecurity purposes.
6. Protect any sensitive information discovered during experimentation.

The purpose of the project is exclusively to improve defensive botnet detection.

3.14 Research Limitations

Several limitations should be recognized.

Dataset Dependence

Results obtained using Bot-IoT may not fully represent every real-world network.

Simulated Environment

Some Bot-IoT traffic was generated in a controlled cyber-range environment rather than a production enterprise network.

Unknown Attacks

A supervised classifier may perform poorly on botnet families that differ substantially from those included in the training dataset.

Computational Resources

Deep-learning architectures may require greater processing power and training time.

Concept Drift

Normal and malicious network behaviors change over time.

A model trained using historical data may therefore require periodic retraining.

Class Imbalance

Large differences between attack categories may affect model training and performance evaluation.

3.15 Expected Results

It is expected that machine-learning and deep-learning techniques will demonstrate a strong ability to distinguish botnet traffic from legitimate network activity.

However, the research does not assume that the most complex model will automatically perform best.

The anticipated findings are:

1. Effective Botnet Classification

At least one selected AI model is expected to produce high precision, recall, and F1-score.

2. Identification of Important Features

The research is expected to identify a reduced group of network characteristics that strongly influence botnet detection.

3. Reduced Computational Requirements

Feature selection is expected to decrease model training and prediction requirements without causing a major reduction in detection performance.

4. Comparison of AI Techniques

The experiment will provide a controlled comparison between classical machine-learning algorithms and neural-network techniques.

5. Detection Framework

A structured framework will be produced consisting of:

Traffic Collection → Preprocessing → Feature Selection → AI Classification → Alert

6. Practical Recommendations

The study will recommend an AI algorithm suitable for future implementation in an intrusion detection environment.

3.16 Expected Contribution

The principal contribution of this research will be a systematic AI-based framework for botnet attack detection.

The proposed study contributes by:

- Comparing multiple algorithms under identical conditions.
- Evaluating more than accuracy alone.
- Analyzing feature importance.
- Considering resource efficiency.
- Examining the limitations of supervised botnet detection.
- Providing recommendations for practical AI-assisted cybersecurity deployment.

3.17 Future Research

Future research may extend the proposed framework in several directions.

Cross-Dataset Evaluation

A model trained on Bot-IoT could be evaluated using N-BaIoT or IoT-23 to investigate generalization.

Zero-Day Botnet Detection

Unsupervised, semi-supervised, and zero-shot techniques could be evaluated against previously unseen attacks.

Explainable AI

SHAP and LIME could be integrated to explain why individual traffic flows are classified as malicious.

Federated Learning

Future systems may train collaboratively across multiple IoT networks without transferring raw sensitive traffic to a central server.

Edge Deployment

Lightweight models could be deployed on:

- Routers.
- IoT gateways.

- Edge servers.

Continuous Learning

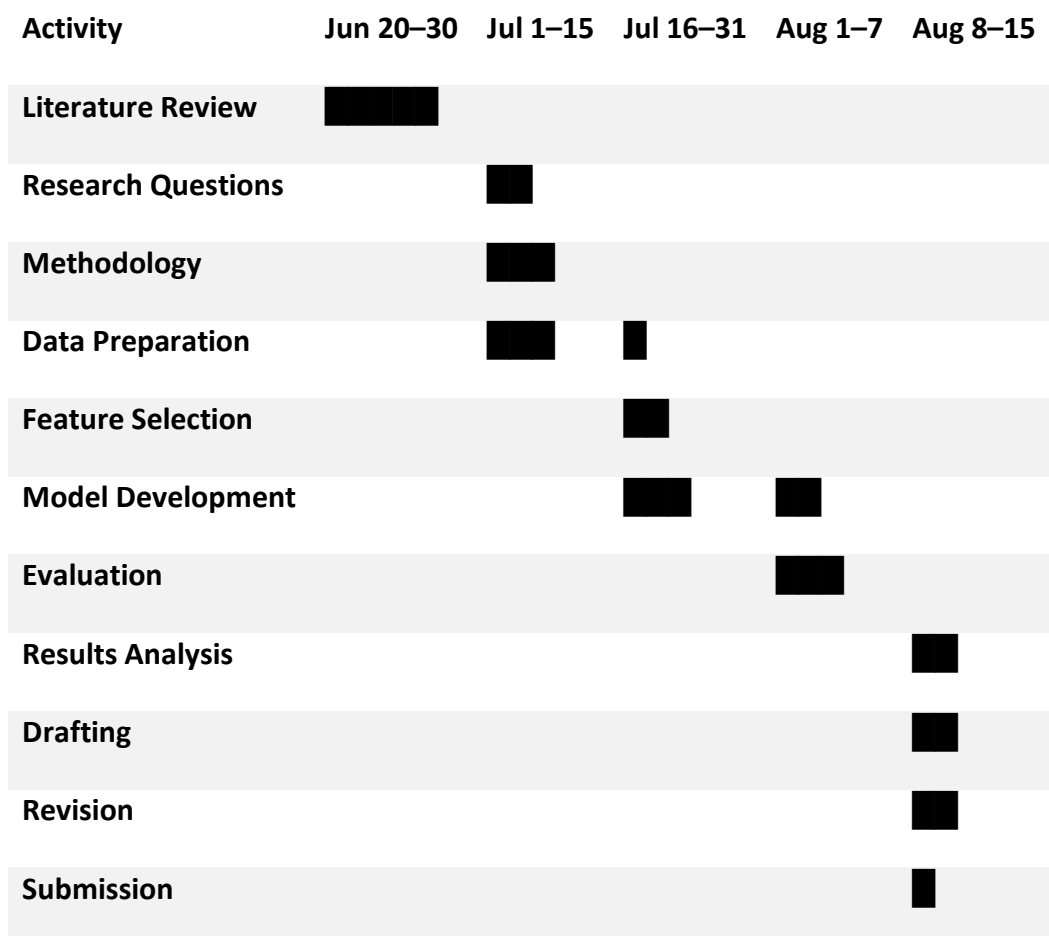
Models could be updated periodically as new network traffic and botnet variants become available.

Gantt Chart

The research work can be organized according to the following schedule:

Research Activity	Start Date	End Date	Duration
Literature Review	20 Jul 2026	30 Jun 2026	11 days
Formulate Research Problem and Questions	1 Jul 2026	5 Jul 2026	5 days
Develop Research Methodology	6 Jul 2026	12 Jul 2026	7 days
Dataset Collection and Preparation	13 Jul 2026	22 Jul 2026	10 days
Feature Selection	23 Jul 2026	27 Jul 2026	5 days
AI Model Development	28 Jul 2026	4 Aug 2026	8 days
Performance Evaluation	5 Aug 2026	8 Aug 2026	4 days
Results Analysis	9 Aug 2026	10 Aug 2026	2 days
Research Drafting	11 Aug 2026	12 Aug 2026	2 days
Review and Revision	13 Aug 2026	14 Aug 2026	2 days
Final Submission	15 Aug 2026	15 Aug 2026	1 day

Simplified Gantt Representation



References

- [1] N. Koroniotis, N. Moustafa, E. Sitnikova, and B. P. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779-796, 2019, doi: 10.1016/j.future.2019.05.041.
- [2] Y. Meidan, M. Bohadana, Y. Mathov, Y. Mirsky, D. Breitenbacher, A. Shabtai, and Y. Elovici, "N-BaloT: Network-based detection of IoT botnet attacks using deep autoencoders," *IEEE Pervasive Computing*, vol. 17, no. 3, pp. 12-22, 2018.
- [3] Y. Mirsky, T. Doitshman, Y. Elovici, and A. Shabtai, "Kitsune: An ensemble of autoencoders for online network intrusion detection," in *Proceedings of the Network and Distributed System Security Symposium (NDSS)*, 2018.
- [4] S. García, M. Grill, J. Stiborek, and A. Zunino, "An empirical comparison of botnet detection methods," *Computers & Security*, vol. 45, pp. 100-123, 2014, doi: 10.1016/j.cose.2014.05.011.
- [5] S. Garcia, A. Parmisano, and M. J. Erquiaga, "IoT-23: A labeled dataset with malicious and benign IoT network traffic," *Stratosphere Laboratory*, 2020, doi: 10.5281/zenodo.4743746.
- [6] J. Ashraf, M. Keshk, N. Moustafa, M. Abdel-Basset, H. Khurshid, A. D. Bakhshi, and R. R. Mostafa, "IoTBoT-IDS: A novel statistical learning-enabled botnet detection framework for protecting networks of smart cities," *Sustainable Cities and Society*, vol. 72, Art. no. 103041, 2021.
- [7] F. Taher et al., "Reliable machine learning model for IIoT botnet detection," *IEEE Access*, vol. 11, pp. 49319-49336, 2023, doi: 10.1109/ACCESS.2023.3253432.
- [8] C. D. Luu, V. Q. Nguyen, and T. S. Pham, "A zero-shot deep learning approach for unknown IoT botnet attack detection," in *2023 RIVF International Conference on Computing and Communication Technologies*, 2023, doi: 10.1109/RIVF60135.2023.10471793.
- [9] M. S. Alshehri, J. Ahmad, S. Almakdi, M. Al Qathrady, Y. Y. Ghadi, and W. J. Buchanan, "SkipGateNet: A lightweight CNN-LSTM hybrid model with learnable skip connections for efficient botnet attack detection in IoT," *IEEE Access*, vol. 12, pp. 35521-35538, 2024, doi: 10.1109/ACCESS.2024.3371992.
- [10] M. Ali, M. Shahroz, M. F. Mushtaq, S. Alfarhood, M. Safran, and I. Ashraf, "Hybrid machine learning model for efficient botnet attack detection in IoT environment," *IEEE Access*, vol. 12, pp. 40682-40699, 2024, doi: 10.1109/ACCESS.2024.3376400.
- [11] A. Nazir et al., "Advancing IoT security: A systematic review of machine learning approaches for the detection of IoT botnets," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 10, Art. no. 101820, 2023, doi: 10.1016/j.jksuci.2023.101820.

[12]M. Saied and S. Guirguis, "Explainable artificial intelligence for botnet detection in Internet of Things," Scientific Reports, vol. 15, Art. no. 7632, 2025, doi: 10.1038/s41598-025-90420-6.

" [13]Systematic literature review of IoT botnet DDoS attacks and evaluation of detection techniques," Sensors, vol. 24, no. 11, Art. no. 3571, 2024, doi: 10.3390/s24113571.

Conclusion

Botnet attacks remain a significant cybersecurity concern, particularly as Internet of Things devices continue to expand across consumer, industrial, healthcare, transportation, and smart-city environments. Traditional signature-based security solutions provide useful protection against known threats but may be less effective against rapidly changing and previously unseen botnet behavior.

Artificial intelligence provides an alternative approach by learning statistical and behavioral characteristics directly from network traffic. Machine-learning models such as Random Forest and gradient boosting can offer strong classification performance with relatively manageable computational requirements, while neural and deep-learning approaches can identify more complicated traffic relationships.

The literature reviewed in this research demonstrates that no single technique provides a complete solution. Model effectiveness depends on dataset quality, traffic characteristics, preprocessing procedures, computational resources, and the characteristics of the botnet being detected.

For this reason, the proposed study develops a comparative AI-based detection framework rather than assuming that one algorithm is universally superior. Bot-IoT will serve as the primary experimental dataset, and selected machine-learning and neural-network classifiers will be trained under consistent preprocessing and testing conditions.

Accuracy, precision, recall, F1-score, false-positive rate, computational requirements, and prediction speed will be considered when evaluating the models.

The expected outcome is the identification of an AI technique capable of delivering an effective balance between high botnet detection performance and practical computational efficiency.

Ultimately, the research demonstrates how artificial intelligence can strengthen network intrusion detection while also identifying critical future research requirements, including cross-dataset generalization, explainable AI, lightweight edge deployment, continuous learning, and detection of previously unseen botnet families.