IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Journal Européen des Systèmes Automatisés
Vol. 58, No. 6, June, 2025, pp. 1189-1196
Journal homepage: http://iieta.org/journals/jesa

A Survey of Artificial Intelligence Techniques in Intrusion Detection for the Internet of Things

Mourad Hana[1*], Ibtissame Aouraghe[2], Oussama El Haouari[3], Ghizlane Khaissidi[1], Mostafa Mrabti[1]

[1] Laboratoire d'Ingénierie, Systèmes et Applications, ENSA, USMBA, Fez 30000, Morocco
[2] National School of Applied Sciences, ENSA, University Sultan Moulay Slimane, Beni Mellal 23000, Morocco
[3] LASET, Laboratory of Applied Sciences and Emerging Technologies, ENSA, USMBA, Fez 30000, Morocco

Corresponding Author Email: mourad.hana@usmba.ac.ma

https://doi.org/10.18280/jesa.580609

Received: 5 May 2025
Revised: 10 June 2025
Accepted: 19 June 2025
Available online: 30 June 2025

Keywords:
security, artificial intelligence, Machine Learning, deep learning, Internet of Things

ABSTRACT

The Internet of Things (IoT) is revolutionizing living standards. As IoT-based solutions proliferate rapidly, ensuring their security has become an increasingly pressing concern for both manufacturers and consumers. There's a growing trend towards leveraging artificial intelligence (AI) techniques to bolster IoT security. This survey paper is dedicated to examining recent advancements in applying AI to intrusion detection within the IoT realm. The selected articles are categorized based on the used AI algorithms. This research offers a thorough examination of recent breakthroughs in AI aimed at enhancing IoT security. It synthesizes and organizes recent relevant research findings, providing a comprehensive discussion on research challenges, ongoing issues, and areas requiring future investigation.

1. INTRODUCTION

As digital systems become more connected, the Internet of Things (IoT) is playing an increasingly important role in everyday life. From smart homes and wearable devices to industrial automation, IoT technologies are now embedded in how we live and work. But this rapid growth comes with new security challenges. Many IoT devices were designed with functionality in mind, not security, which makes them vulnerable to cyberattacks. Traditional tools that rely on predefined rules or known threat signatures are often too limited to handle the speed, scale, and complexity of modern attacks [1].

To tackle these challenges, researchers are increasingly turning to artificial intelligence (AI), particularly its subfields Machine Learning (ML) and deep learning (DL). What makes these methods powerful is that they can learn from traffic patterns and identify unusual behavior even when the attack is entirely new without needing prior knowledge of specific threats [2-4]. Several recent studies have shown that learning-based systems can reduce false positives and detect more subtle or novel intrusions compared to older, rule-based methods [5, 6].

In this paper, we explore how Machine Learning and deep learning are being applied to intrusion detection in IoT networks. We review recent progress, examine practical challenges still facing researchers and practitioners, and outline some of the directions where the field seems to be heading. Our goal is to give both researchers and security professionals a clearer picture of the current landscape and of what needs to happen next to improve the security of increasingly connected environments.

2. MATERIALS AND METHODS

To better understand how artificial intelligence is being used for intrusion detection in IoT systems, we carried out a targeted review of recent scientific research. We followed a structured approach to make sure the studies we included were both relevant and of high quality.

Our search was conducted using several major academic databases: Web of Science, Scopus, IEEE Xplore, ScienceDirect, and ResearchGate. We focused only on articles written in English and published between 2016 and 2024. To narrow down the results, we used a set of key terms such as "artificial intelligence," "Machine Learning," "deep learning," "intrusion detection," and "cybersecurity" which we applied to titles, abstracts, and keywords during the search process.

After collecting the initial results, we first screened abstracts to remove papers that didn't clearly match our topic. For the remaining studies, we reviewed the full texts to check whether they met the following inclusion criteria:

- The paper must be at least four pages long.
- It must present original research published in a peer-reviewed journal, conference, or workshop.
- The research must directly focus on intrusion detection in IoT environments.

We excluded editorials, prefaces, tutorials, panel summaries, posters, and any work that lacked clear relevance to our focus. Studies that addressed cybersecurity in general without a specific IoT context were also filtered out.

By applying this review process, we were able to select a set of studies that provided a solid foundation for our analysis and helped us capture current trends in the use of AI for IoT security.

3. BACKGROUND

The Internet's rapid growth has fueled a huge rise in connected devices across different industries but it has also made systems more vulnerable to cyberattacks [7]. Intrusion, meaning unauthorized access to digital systems, has become a major global concern because the Internet lets attackers cross borders easily [8].

To help manage modern cybersecurity risks, Network Intrusion Detection Systems (NIDS) have become a central part of many defense strategies. These tools monitor traffic on a network and raise alerts often in real time when they detect suspicious behavior or signs of a policy breach [9]. In general, NIDS fall into two main categories: signature-based systems, which search for known attack patterns, and anomaly-based systems, which try to spot unusual behavior that doesn't match expected norms [10].

Anomaly-based methods have gained more traction in recent years, largely because they can detect previously unseen threats including zero-day attacks without needing a known signature. NIDS can also be categorized based on where they operate: at the network level (N-NIDS), on individual hosts (H-NIDS), or in wireless environments (W-NIDS). Depending on the setup, they can help detect everything from phishing attempts and malware infections to large-scale DDoS attacks and long-term persistent intrusions.

As attacks grow more sophisticated, researchers have turned to Machine Learning (ML) to strengthen intrusion detection. By analyzing past traffic data, ML models can pick up on unusual trends that signal a possible attack. These models are useful both for offline analysis and real-time monitoring. In fast-moving environments like the IoT, the need for adaptable, intelligent detection tools is more urgent than ever.

4. RELATED WORK

Researchers have applied a variety of artificial intelligence techniques to improve intrusion detection in IoT systems, addressing challenges that range from feature selection to real-time classification.

For instance, Kumar et al. [11] worked on optimizing feature selection by combining Random Forest ranking with correlation analysis and gain ratio scoring. Their hybrid method led to better classification results and more efficient use of input features. In a different context, Jeyanthi et al. [12] focused on IoT applications in healthcare, developing a CNN-based intrusion detection system that could automatically extract relevant patterns from traffic data. Their model achieved an average accuracy of over 95% across various IoT-specific attack types, demonstrating reliable generalization. Looking at urban IoT systems, Rashid et al. [13] explored multiple Machine Learning approaches for anomaly detection in smart cities. Their study tested models like KNN, Logistic Regression, Random Forest, SVM, and Decision Trees, and also evaluated ensemble methods such as bagging and boosting. These ensemble combinations improved robustness, with accuracy and F1-score consistently exceeding 94% in experimental evaluations.

Moving toward decentralized approaches, Guha Roy and Srirama [14] introduced an IDS framework that combines blockchain technology with Software-Defined Networking (SDN). Their solution emphasized both secure event logging and flexible traffic management, addressing scalability concerns in large IoT networks.

In terms of performance optimization, Heidari et al. [7] explored how GPU-accelerated CNNs could speed up intrusion detection without compromising accuracy, aiming for real-time deployment even under resource constraints. Elsisi and Tran [15] pursued a similar goal but focused on lightweight Deep Neural Network (DNN) architectures specifically for protecting Automated Guided Vehicles (AGVs) in industrial IoT systems.

Finally, Saharkhizan et al. [16] proposed an ensemble model combining Long Short-Term Memory (LSTM) networks with Decision Trees, capturing both temporal and structural patterns in IoT traffic, this ensemble achieved high detection accuracy, with F1-scores above 95%, making it suitable for handling the evolving complexity of modern cyberattacks.

Altogether, these studies show a clear trend: while Machine Learning and deep learning methods have significantly advanced IoT intrusion detection, challenges remain around building generalizable models, improving interpretability, and deploying solutions in resource-constrained environments.

Altogether, these studies show a clear trend: while Machine Learning and deep learning methods have significantly advanced IoT intrusion detection, challenges remain around building generalizable models, improving interpretability, and deploying solutions in resource-constrained environments. Notably, several works evaluated their approaches using benchmark datasets like NSL-KDD, CICIDS2017, or custom IoT traffic datasets. For example, the CNN-based model by Jeyanthi et al. [12] outperformed traditional methods like Decision Trees and SVM on IoT-specific traffic data, achieving over 95% accuracy, while Rashid et al. [13] reported similar performance (F1-scores > 94%) using ensemble methods on smart city datasets. In contrast, the LSTM-based ensemble proposed by Saharkhizan et al. [16] achieved F1-scores above 95% while capturing temporal dependencies, highlighting the advantage of deep temporal models for evolving traffic patterns. Compared to lightweight DNNs explored by Elsisi and Tran [15], which prioritized fast inference on AGVs, GPU-accelerated CNNs from Heider et al. [7] demonstrated better scalability for real-time detection but required higher computational resources. This contrast illustrates the trade-offs between model complexity, inference speed, and detection accuracy, especially under deployment constraints. Although each method addresses different facets of the intrusion detection problem, a standardized evaluation framework on common datasets is still lacking, which makes direct comparison and generalization challenging.

5. MOTIVATION

As cyber threats become more advanced, traditional security tools are struggling to keep up especially in fast-changing environments like the IoT. With more devices connecting every day, network vulnerabilities are expanding as well. Manual, rule-based intrusion detection systems (IDS) often can't match the speed and complexity of today's attacks, making it critical to move toward smarter, adaptive defense strategies.

5.1 The role of Machine Learning in intrusion detection

Machine Learning (ML), a subfield of Artificial Intelligence

(AI), offers adaptable methods for identifying threats in evolving environments. ML provides powerful tools for dealing with large, evolving cybersecurity threats. Unlike traditional signature-based systems, ML models learn from data directly, allowing them to spot malicious behavior and unknown attack patterns without needing a full library of signatures [17]. Models like Decision Trees, Support Vector Machines (SVMs), Random Forests, and clustering algorithms have all shown strong detection performance across different types of intrusion scenarios [18]. Another advantage of ML-based IDS is that they can automate much of the threat analysis process, cutting down the need for constant human supervision. In IoT environments where data streams are continuous, diverse, and often unstructured supervised, semi-supervised, and unsupervised ML models have proven especially useful for tasks like anomaly detection, spam filtering, malware identification, and behavior tracking [19]. By learning from both labeled and unlabeled data, these systems can react faster and help maintain better awareness of evolving threats.

5.2 The rise of deep learning

Deep learning (DL), a specialized branch of ML, has pushed intrusion detection even further by allowing systems to automatically pull complex features out of massive datasets. Architectures like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) networks are particularly good at recognizing spatial and time-based patterns hidden within network traffic [20].

DL models have made real progress in spotting new types of attacks, handling encrypted traffic, and keeping detection rates high even in dynamic IoT settings [21]. Thanks to advances in GPU and TPU hardware, it's now possible to deploy deep learning-based IDS models close to real-time even in environments with limited computing power.

Of course, DL approaches still come with challenges. Interpretability, high computational costs, and vulnerability to adversarial attacks remain open issues. Researchers are actively working on solutions, from explainable AI (XAI) frameworks to lightweight deep learning architectures and more resilient training methods [22]. As DL continues to evolve, its role in smart, adaptive intrusion detection systems for IoT will likely grow even more important.

6. ENHANCING CYBERSECURITY THROUGH ML TECHNIQUES

As cyberattacks continue to grow more sophisticated, static, rule-based intrusion detection systems are struggling to keep up especially in dynamic environments like IoT. Machine Learning (ML) offers a better alternative by adapting to new patterns, scaling data growth, and identifying threats that haven't been seen before. Because of these strengths, ML techniques are becoming an essential part of modern intrusion detection.

By analyzing historical network behavior, ML-based systems can spot previously unknown attacks, lower false alarm rates, and strengthen real-time defenses. In the next part of this work, we review a range of ML algorithms that have been applied to intrusion detection in IoT networks, looking at both their advantages and their limitations.

6.1 Summary of ML techniques for IoT security

Many studies have explored how ML models can improve intrusion detection. Table 1 provides a summary of selected research efforts, outlining the methods used, the problems they addressed, and the key results achieved.

These findings show that no single model always performs best in every situation. How well an ML method works often depends on the specific dataset, how features are selected, and the nature of the attacks being targeted.

Table 1. Summary of AI-based intrusion and malware

Article	Dataset	Purpose	Techniques	Results
Javaheri et al. [23]	- Emulated Traffic: Simulated RoQ attack + legitimate traffic - Real Traffic: Real-world traffic with M-RoQ attacks - Training Data: Mixed attack and legitimate traffic.	To identify DDoS and other threats using simulated and real network traffic.	MLP, KNN, SVM, Fuzzy Logic, Euclidean Distance, MNB	TMLP provided the most effective classification; F1-score reached 98% on emulated data and exceeded 99% on real data.
Li et al. [24]	- KDDCup'99: Simulated - NSL-KDD: An improved and more balanced version of KDD - CIC-IDS2017: Realistic modern dataset created by the Canadian Institute for Cybersecurity.	To design a robust intrusion detection system using a streamlined classifier.	SVM	Achieved high accuracy of 98.62%, highlighting strong generalization.
Chen et al. [25]	- A real alarm dataset collected from a real network environment. - A synthetic testbed dataset constructed in a simulated network environment.	To enhance alert filtering with a knowledge-driven method.	KNN	Achieved 93.2% accuracy and 91.8% F-measure, Marking it as the most robust filter, with 93.2% accuracy and an F-measure of 91.8%.
Mahindru and Sangal [26]	- Over 500,000 real-world Android applications, including both benign and malicious apps, collected for dynamic analysis.	To detect malware in large-scale Android apps using. feature extraction.	Deep Learning (DL), FFC, YMLP, DT	Accuracy surpassed 98% for real-world Android threat detection.
Zuhair and Selamat [27]	- Dataset comprising multiple ransomware families and nine commonly exploited traits.	To recognize ransomware through a multi-stage detection framework.	Naïve Bayes, Decision Tree	Delivered an average accuracy of 96.27% with a real-time error margin of just 1.32%.
Gharbi et al. [28]	- High-dimensional ransomware dataset with 30,000 features, from which five were selected.	To predict ransomware presence using a rich	SVM	Classification accuracy was over 88%, indicating solid performance

		attribute dataset.		for ransomware detection.
Guezzaz et al. [29]	- The KDD99 dataset Is used, containing labeled network traffic data representing various types of attacks (DoS, Probe, U2R, R2L).	To monitor systems for threats and detect malicious activity.	Decision Tree (Binary Split)	Accuracy exceeded 99% when classifying intrusion attempts.
Song et al. [30]	- KDD'99 dataset, which contains labeled network traffic data used to evaluate misuse, anomaly, and hybrid intrusion detection systems.	To refine Random Forest for misuse and anomaly detection.	Enhanced RF frameworks	Achieved reliable detection, with precision and recall values exceeding 95% in identifying abnormal network behavior.
Shamim et al. [31]	- Network traffic capturing and simulation of TCP SYN flood DDoS attacks in an SDN environment, with labeled features extracted from P4-enabled switches to enable real-time Machine Learning-based attack detection.	To detect DDoS attacks in SDN environments.	KNN, RF, SVM, and ANN	Accuracy surpassed 98%, confirming the viability of traditional classifiers in modern architectures.

6.2 Detecting intrusions with naïve bayes classifier

Naïve Bayes (NB) classifiers are simple yet powerful tools built on probabilistic reasoning and Bayes' theorem. Thanks to their speed and ability to work well with high-dimensional data, they've been widely used for intrusion detection. Shamim et al. [31] designed an NB-based intrusion detection system that achieved a 96.45% detection rate on datasets like KDDCup'99. Because NB models are lightweight, they're a time intrusion detection, especially in resource-limited IoT environments.

6.3 Evolutionary fuzzy neural networks in cybersecurity

Hybrid approaches that bring together fuzzy logic, neural networks, and evolutionary algorithms offer new ways to deal with uncertainty and nonlinear traffic behavior. In this area, Verma et al. [32] showed that evolving fuzzy neural networks (EFuNNs) can adapt dynamically as they encounter new traffic patterns, outperforming more static detection models. By combining the flexibility of neural networks with the reasoning power of fuzzy systems, EFuNNs achieved improved detection rates, outperforming traditional models with accuracy levels above 95% on evolving traffic patterns.

6.4 Support vector machine for anomaly detection

Support Vector Machines (SVMs) are widely used for binary classification, especially when dealing with high-dimensional data. In their work, Shamim et al. [31] introduced a hybrid intrusion detection framework based on SVMs, optimized with a modified K-means clustering method. Their model reached 95.75% accuracy on the KDDCup'99 dataset, demonstrating SVM's ability to separate normal and malicious traffic with an accuracy of 95.75% on the KDDCup'99 dataset even in complex IoT environments.

6.5 Proactive defense of CAVS

As Connected and Autonomous Vehicles (CAVs) become more common, they also create new cybersecurity risks. Bakker et al. [33] addressed this by developing the Knowledge Enhanced Machine Learning Pipeline (KEMLP), a framework that adds domain-specific knowledge to ML models. Their approach enhanced IDS reliability, enabling real-time detection with improved precision in fast-changing vehicular network conditions in vehicle networks, allowing for real-time threat detection even under fast-changing conditions.

6.6 Combatting DDoS attacks with ML algorithms

Distributed Denial of Service (DDoS) attacks are a serious risk for IoT systems because of their ability to overwhelm networks. Al-Yaseen et al. [34] tested several ML classifiers, including SVM, Decision Trees, and KNN, to detect DDoS activity. Using the ISCX dataset, they found that SVM performed the best, with detection rates between 92% and 93%. Their findings highlight how ML classifiers such as SVM achieved detection rates between 92% and 93%, reinforcing their utility for early DDoS detection and help limit large-scale disruptions.

6.7 Behavioral analysis for botnet detection

Detecting botnets remains a tough challenge because of how decentralized and adaptive they are. Bakker et al. [33] built a behavior-based detection system using Machine Learning models, reaching accuracy rates from 99.23% to 99.86%. SVM achieved the highest accuracy (up to 99.86%) across multiple test cases, confirming its robustness across different test cases, pointing to how important feature selection and model tuning are for spotting botnet traffic effectively.

6.8 Utilizing particle swarm optimization with one-class SVM

To improve feature selection and anomaly detection, Gürel et al. [35] proposed a hybrid method that combines Particle Swarm Optimization (PSO) with One-Class Support Vector Machines (OC-SVM). Their model, tested on the UNSW-NB15 dataset, achieved 86.68% accuracy. The results suggest that this approach works well for detecting new types of attacks, especially in IoT environments where labeled training data is limited.

6.9 Real-World deployment challenges

Despite promising results in experimental settings, deploying AI-based intrusion detection systems in real-world IoT environments presents numerous challenges. Many IoT devices operate under strict resource constraints, limited memory, processing power, and battery life which makes it difficult to run complex ML or DL models locally. Additionally, real-world network traffic is often noisy, imbalanced, and unpredictable, reducing the generalizability of models trained on clean benchmark datasets. Furthermore, latency, privacy concerns, and unreliable connectivity complicate data collection and decision-making. To address these obstacles, future research must focus on lightweight, adaptive, and privacy-preserving AI models tailored to the realities of IoT ecosystems.

7. ENHANCING CYBERSECURITY THROUGH DEEP LEARNING

Deep learning (DL), a specialized area within Machine Learning (ML), has become an important tool for handling large, complex cybersecurity datasets. Its ability to automatically pull structured features from raw input makes it especially good at detecting sophisticated attacks, including zero-day exploits and stealth intrusions in IoT environments [36].

Traditional ML models still work well for simpler or lower-dimensional problems. But when it comes to dealing with encrypted traffic, analyzing sequential behavior, or managing noisy data, DL architectures like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders (AEs), and Deep Neural Networks (DNNs) are often a better fit [37]. Advances in GPU and TPU technology have also made it much easier to deploy DL models in real-time cybersecurity applications.

That said, deep learning isn't without its challenges. Problems like limited interpretability, vulnerability to adversarial attacks, and high resource demands are still open research areas. Explainable AI (XAI) has emerged as a critical field aimed at making deep learning models more transparent and trustworthy. In the context of IoT security, where decisions can affect real-world systems like healthcare, smart homes, or industrial controls, understanding why a model flagged an intrusion is essential. XAI methods allow researchers and practitioners to trace, interpret, and justify model outputs, helping to identify false positives or uncover biases. This transparency builds trust in automated systems and supports better decision-making by human analysts. As deep learning continues to evolve, the integration of XAI into intrusion detection systems will be vital to ensure responsible and interpretable deployment in sensitive IoT environments, helping DL systems become safer and more practical for intrusion detection [38].

The next sections will look at some of the main deep learning architectures used to boost IoT cybersecurity.

7.1 Multi-layer perceptron

The Multi-Layer Perceptron (MLP) is one of the earliest and simplest deep learning models. Structured as a feedforward neural network, an MLP stacks multiple layers of neurons to learn complex nonlinear patterns [39]. In intrusion detection, MLPs often perform well on structured, tabular data. However, they aren't the best choice for capturing spatial or time-based patterns, where models like CNNs and RNNs tend to do better.

7.2 Artificial neural network

Artificial Neural Networks (ANNs) build on the basic ideas behind MLPs but offer more flexible designs suited to different types of data and applications. Some studies have reported that well-tuned ANN models can hit near-perfect detection rates on certain cybersecurity datasets [40]. Their ability to adapt and keep learning over time makes them a strong option for dynamic environments where intrusion patterns are constantly changing.

7.3 Conventional neural network

Convolutional Neural Networks (CNNs) were originally designed for computer vision but have been successfully adapted to cybersecurity tasks too especially for analyzing network traffic at the packet or flow level [41]. CNNs are great at pulling out spatial features automatically, which helps when classifying different types of cyberattacks like malware, botnets, and DoS incidents. Deeper CNNs and hybrid models have further boosted detection performance, particularly in industrial IoT and smart grid security settings [42].

7.4 Recurrent neural network

Recurrent Neural Networks (RNNs) add memory to deep learning models, which allows them to capture patterns over time [43]. This makes RNNs especially useful for cybersecurity problems, where many attacks don’t happen at once but unfold across a sequence of events. Improved variants like Long Short-Term Memory (LSTM) networks [44] and Gated Recurrent Units (GRUs) [45] help solve common training problems like the vanishing gradient issue and have shown strong results in time-series-based intrusion detection.

7.5 Deep neural network

As deep learning models grow in size, they can sometimes suffer from performance drops during training. Residual Networks (ResNets) [46] fix this by adding shortcut connections that help gradients move easily through deeper layers. In cybersecurity tasks, ResNet-based models have delivered impressive results, caught complex attacks while stayed efficient and stable during training. Their ability to learn from complicated data makes them a strong fit for intrusion detection in IoT systems.

Table 2. Comparison of DL architectures used in IDS

DL Architecture	Strengths	Best Use Cases in IDS	Limitations
MLP	Learns complex nonlinear patterns; easy to implement	Detecting intrusions in structured, tabular datasets	Not suitable for spatial or temporal data
ANN	Flexible architecture; adaptable to changing environments	Evolving intrusion patterns in dynamic systems	Needs careful tuning; can overfit on small data
CNN	Automatically extracts spatial features; scalable to large inputs	Packet- and flow-level traffic classification; malware, DoS	Not effective for sequential or time-series data
RNN	Captures temporal patterns; memory of past inputs	Time-dependent attacks; sequential analysis	Training challenges like vanishing gradients (mitigated in GRU/LSTM)
DNN	Learns from high-dimensional, complex data; stable training with ResNet	Detecting sophisticated or blended threats in IoT systems	Computationally intensive; low interpretability
AE	Unsupervised anomaly detection; handles noise in input	Spotting novel or subtle attacks in noisy IoT data	May reconstruct benign-looking attacks too well

7.6 Auto encoder

Autoencoders (AEs) are a type of unsupervised model that compresses data and then tries to reconstruct it. They're widely used for detecting anomalies, since anything that can't be reconstructed well is likely to be an unusual event [47]. Denoising Autoencoders (DAEs) [48] take this a step further by learning to clean up noisy input data before reconstruction. In intrusion detection systems, high reconstruction errors often flag abnormal behavior, helping AEs and DAEs spot subtle attacks even when the data is messy, like in many IoT environments.

To provide a clearer overview of how different deep learning models contribute to intrusion detection systems, Table 2 presents a comparative summary of their strengths, use cases, and limitations in cybersecurity contexts.

8. CONCLUSION

AI has attracted growing interest from both researchers and industry due to its potential to improve the security of IoT systems. In the context of intrusion detection, AI offers new ways to identify threats quickly and adapt to evolving attack patterns. Over the past few years, it has become a major focus for improving how IoT networks are monitored and protected in real time.

This paper presents a detailed review of AI-driven methods for detecting intrusions in IoT environments. It categorizes approaches based on the type of learning used supervised, semi-supervised, or unsupervised and highlights the strengths and limitations of each. One of the key challenges discussed is the heavy reliance on clean, diverse training data, which is not always available or representative of real-world conditions. While many systems report accuracy values ranging from 90% to 99% on benchmark datasets, their performance in real deployments can vary significantly due to noise, class imbalance, and limited generalization.

Still, the research shows that AI-powered intrusion detection systems hold strong potential to improve IoT security and respond more effectively to modern threats. The review also outlines areas where more work is needed, such as improving model efficiency for resource-limited devices and testing on more realistic datasets. Moving forward, continued research and practical validation will be essential to realize the full benefits of AI for securing IoT systems.

However, several challenges remain. Most AI-based IDS models are evaluated on benchmark datasets that may not reflect the complexity or noise of real-world IoT environments. This gap leads to poor generalization when deployed outside the lab. Moreover, many models assume abundant clean training data, which is rarely available in IoT contexts. Resource constraints such as limited processing power, memory, and battery further restrict the implementation of deep learning solutions on edge devices. Future work must address these issues by prioritizing lightweight, adaptive, and robust models capable of handling real-world variability.

REFERENCES

[1] Azeez, N.A., Bada, T.M., Misra, S., Adewumi, A., Van der Vyver, C., Ahuja, R. (2020). Intrusion detection and prevention systems: An updated review. Advances in Intelligent Systems and Computing, 1042: 685-696. https://doi.org/10.1007/978-981-32-9949-8_48

[2] Sarker, I.H. (2023). Machine learning for intelligent data analysis and automation in cybersecurity: Current and future prospects. Annals of Data Science, 10(6): 1473-1498. https://doi.org/10.1007/s40745-022-00444-2

[3] Sarker, I.H. (2022). AI-based modeling: Techniques, applications and research issues towards automation, intelligent and smart systems. SN Computer Science, 3(2): 158. https://doi.org/10.1007/s42979-022-01043-x

[4] Sarker, I.H. (2021). Machine learning: Algorithms, real-world applications and research directions. SN Computer Science, 2(3): 160. https://doi.org/10.1007/s42979-021-00592-x

[5] Ge, M., Fu, X., Syed, N., Baig, Z., Teo, G., Robles-Kelly, A. (2019). Deep learning–based intrusion detection for IoT networks. In 2019 IEEE 24th Pacific Rim International Symposium on Dependable Computing (PRDC), Kyoto, Japan, pp. 256-25609. https://doi.org/10.1109/PRDC47002.2019.00056

[6] Qazi, E.U.H., Zia, T., Faheem, M. H., Shahzad, K., Imran, M., Ahmed, Z. (2024). Zero-Touch Network Security (ZTNS): A network intrusion detection system based on deep learning. IEEE Access, 12: 141625-141638. https://doi.org/10.1109/ACCESS.2024.3466470

[7] Heidari, A., Jabraeil Jamali, M.A. (2023). Internet of Things intrusion detection systems: A comprehensive review and future directions. Cluster Computing, 26: 3753-3780. https://doi.org/10.1007/s10586-022-03776-z

[8] Khraisat, A., Alazab, A. (2021). A critical review of intrusion detection systems in the Internet of Things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges. Cybersecurity, 4(1): 18. https://doi.org/10.1186/s42400-021-00077-7

[9] Rahman, M.M., Shakil, S.A., Mustakim, M.R. (2025). A survey on intrusion detection system in IoT networks. Cyber Security and Applications, 3: 100082. https://doi.org/10.1016/j.csa.2024.100082

[10] Al-Haija, Q.A., Droos, A. (2024). A comprehensive survey on deep learning-based intrusion detection systems in Internet of Things (IoT). Expert Systems, 42(2). https://doi.org/10.1111/exsy.13726

[11] Kumar, P., Gupta, G.P. Tripathi, R. (2021). Toward design of an intelligent cyber attack detection system using hybrid feature reduced approach for IoT networks. Arabian Journal for Science and Engineering, 46(4): 3749-3778. https://doi.org/10.1007/s13369-020-05181-3

[12] Jeyanthi, D.V., Indrani, B. (2022). IoT based intrusion detection system for healthcare using RNN-BILSTM deep learning strategy with custom features. Soft Computing, 27(16): 11915-11930. https://doi.org/10.21203/rs.3.rs-2302072/v1

[13] Rashid, M.M., Kamruzzaman, J., Hassan, M.M., Imam, T., Gordon, S. (2020). Cyberattacks detection in IoT-based smart city applications using machine learning techniques. International Journal of Environmental Research and Public Health, 17(24): 9347. https://doi.org/10.3390/ijerph17249347

[14] Guha Roy, D., Srirama, S.N. (2021). A blockchain-based cyber attack detection scheme for decentralized Internet of Things using software-defined network. Software: Practice and Experience, 51(7): 1540-1556. https://doi.org/10.1002/spe.2972

[15] Elsisi, M., Tran, M.Q. (2021). Development of an IoT architecture based on a deep neural network against cyber attacks for automated guided vehicles. Sensors, 21(24): 8467. https://doi.org/10.3390/s21248467
[16] Saharkhizan, M., Azmoodeh, A., Dehghantanha, A., Choo, K.K.R., Parizi, R.M. (2020). An ensemble of deep recurrent neural networks for detecting IoT cyber attacks using network traffic. IEEE Internet of Things Journal, 7(9): 8852-8859. https://doi.org/10.1109/JIOT.2020.2996425
[17] Bharadiya, J.P. (2023). Machine learning in cybersecurity: Techniques and challenges. European Journal of Technology, 7(2): 1-14 https://doi.org/10.47672/ejt.1486
[18] Rani, D., Kaushal, N.C. (2020). Supervised machine learning based network intrusion detection system for Internet of Things. In Proceedings of the 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, pp. 1-7. https://doi.org/10.1109/ICCCNT49239.2020.9225340
[19] Selem, M., Jemili, F., Korbaa, O. (2025). Deep learning for intrusion detection in IoT networks. Peer-to-Peer Networking and Applications, 18(2): 22. https://doi.org/10.21203/rs.3.rs-4306367/v1
[20] Deshmukh, A., Ravulakollu, K. (2024). An efficient CNN-based intrusion detection system for IoT: Use case towards cybersecurity. Technologies, (12): 203. https://doi.org/10.3390/technologies12100203
[21] Hammood, B.A.K., Sadiq, A.T. (2024). Deep learning approaches for IoT intrusion detection systems. Iraqi Journal of Science, 65(11). https://doi.org/10.24996/ijs.2024.65.11.36
[22] Halbouni, A., Gunawan, T.S., Habaebi, M.H., Halbouni, Kartiwi, M., Ahmad, R. (2022). Machine learning and deep learning approaches for cybersecurity: A review. IEEE Access, 10: 19572-19585. https://doi.org/10.1109/ACCESS.2022.3151248
[23] Javaheri, D., Gorgin, S., Lee, J.A., Masdari, M. (2023). Fuzzy logic-based DDoS attacks and network traffic anomaly detection methods: Classification, overview, and future perspectives. Information Sciences, 626: 315-338. https://doi.org/10.1016/j.ins.2023.01.067
[24] Li, Y., Xia, J., Zhang, S., Yan, J., Ai, X., Dai, K. (2012). An efficient intrusion detection system based on support vector machines and gradually feature removal method. Expert Systems with Applications, 39(1): 424-430. https://doi.org/10.1016/j.eswa.2011.07.032
[25] Chen, J.S., Kuo, C.M. (2024). An efficient GNSS coordinate classification strategy with an adaptive KNN algorithm for epidemic management. Mathematics, 12(4): 536. https://doi.org/10.3390/math12040536
[26] Mahindru, A., Sangal, A.L. (2021). MLDroid—framework for Android malware detection using machine learning techniques. Neural Computing and Applications, 33(10): 5183-5240. https://doi.org/10.1007/s00521-020-05309-4
[27] Zuhair, H., Selamat, A. (2019). RANDS: A machine learning–based anti-ransomware tool for Windows platforms. In Advancing Technology Industrialization Through Intelligent Software Methodologies, Tools and Techniques, pp. 573-587. https://doi.org/10.3233/FAIA190081
[28] Gharbi, I., Belaoued, M., Derhab, A., Barkaoui, K. (2025). Exploring the landscape of IoT ransomware prediction through machine learning techniques: A comprehensive survey. SN Computer Science, 6(3): 264. https://doi.org/10.1007/s42979-025-03765-0
[29] Guezzaz, A., Benkirane, S., Azrour, M., Khurram, S. (2021). A reliable network intrusion detection approach using decision tree with enhanced data quality. Security and Communication Networks. 2021(1): 1230593. https://doi.org/10.1155/2021/1230593
[30] Song, B. (2024). Random forest based intrusion detection system. In Proceedings of the 2024 Asian Conference on Communication and Networks (ASIANComNet), Bangkok, Thailand, pp. 1-4. https://doi.org/10.1109/ASIANComNet63184.2024.10811056
[31] Shamim, N., Asim, M., Baker, T., Awad, A.I. (2023). Efficient approach for anomaly detection in IoT using system calls. Sensors, 23(2): 652. https://doi.org/10.3390/s23020652
[32] Verma, P., Pateriya, R.K., Baghel, S., Mehta, N., Chaurasia, N., Bharot, N. (2025). Securing cloud networks: A hybrid intrusion detection approach using fuzzy C-means clustering and decision tree classification. In 2025 IEEE International Conference on Communication and Information Technology (ICCIT), Tabuk, Saudi Arabia, pp. 568-573. https://doi.org/10.1109/ICCIT63348.2025.10989339
[33] Bakker, J.N., Ng, B., Seah, W.K. (2018). Can machine learning techniques be effectively used in real networks against DDoS attacks? In 2018 27th International Conference on Computer Communication and Networks (ICCCN), Hangzhou, China, pp. 1-6. https://doi.org/10.1109/ICCCN.2018.8487445
[34] Al-Yaseen, W.L., Othman, Z.A., Ahmad Nazri, M.Z. (2017). Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system. Expert Systems with Applications, 67: 296-303. https://doi.org/10.1016/j.eswa.2016.09.041
[35] Gürel, N.M., Qi, X., Rimanic, L., Zhang, C., Li, B. (2021). Knowledge enhanced machine learning pipeline against diverse adversarial attacks. arXiv preprint arXiv:2106.06235. https://doi.org/10.48550/arXiv.2106.06235
[36] Janati, M., Messaoudi, F. (2025). Intrusion Detection System–based Network Behavior Analysis: A systematic literature review. International Journal of Advanced Computer Science and Applications, 16(3): 793-802. https://doi.org/10.14569/IJACSA.2025.0160378
[37] Dasgupta, D., Akhtar, Z., Sen, S. (2022). Machine learning in cybersecurity: A comprehensive survey. The Journal of Defense Modeling and Simulation, 19(1): 57-106. https://doi.org/10.1177/1548512920951275
[38] Alaghbari, K.A., Lim, H.S., Saad, M.H.M., Yong, Y.S. Deep Autoencoder-Based Integrated Model for Anomaly Detection and Efficient Feature Extraction in IoT Networks. IoT 2023, 4: 345-365. https://doi.org/10.3390/iot4030016
[39] Liu, X., Ahmad, S.F., Anser, M.K., Ke, J., Irshad, M., Ul-Haq, J., Abbas, S. (2022). Cyber security threats: A never-ending challenge for e-commerce. Frontiers in Psychology, 13: 927398. https://doi.org/10.3389/fpsyg.2022.927398
[40] Qaddoura, R., Al-Zoubi, A.M., Faris, H., Almomani, I.

(2021). A multi-layer classification approach for intrusion detection in IoT networks based on deep learning. Sensors, 21(9): 2987. https://doi.org/10.3390/s21092987

[41] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778.

[42] Hanif, S., Ilyas, T., Zeeshan, M. (2019). Intrusion detection in IoT using artificial neural networks on UNSW-15 dataset. In Proceedings of the 2019 IEEE 16th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT and AI (HONET-ICT), Charlotte, NC, USA, pp. 152-156. https://doi.org/10.1109/HONET.2019.8908122

[43] Nguyen, G.N., Le Viet, N.H., Elhoseny, M., Shankar, K., Gupta, B.B., Abd El-Latif, A.A. (2021). Secure blockchain enabled cyber–physical systems in healthcare using deep belief network with ResNet model. Journal of Parallel and Distributed Computing, 153: 150-160. https://doi.org/10.1016/j.jpdc.2021.03.011

[44] Habibi, M.R., Baghaee, H.R., Dragičević, T., Blaabjerg, F. (2020). Detection of false data injection cyber-attacks in DC microgrids based on recurrent neural networks. IEEE Journal of Emerging and Selected Topics in Power Electronics, 9(5): 5294-5310. https://doi.org/10.1109/JESTPE.2020.2968243

[45] Lin, T.N., Giles, C.L., Horne, B.G., Kung, S.Y. (1997). A delay damage model selection algorithm for NARX neural networks. IEEE Transactions on Signal Processing, 45(11): 2719-2730. https://doi.org/10.1109/78.650098

[46] Xu, Y., Zheng, D., Shao, C., Zheng, S., Gu, H., Chen, H. (2023). Real-time diagnosis of structural damage based on NARX neural network with dynamic response. Mathematics, 11(6): 1281. https://doi.org/10.3390/math11061281

[47] Song, Y., Hyun, S., Cheong, Y.G. (2021). Analysis of autoencoders for network intrusion detection. Sensors, 21(13): 4294. https://doi.org/10.3390/s21134294

[48] Jia, L., Du, X. (2021). Rolling bearing fault classification based on stacked denoising auto encoders. IOP Conference Series: Earth and Environmental Science, 769(4): 042085. https://doi.org/10.1088/1755-1315/769/4/042085