AI-Driven Intrusion Detection in Industrial IoT: A Critical Survey of Computational Intelligence Approaches

HAITHAM F. ABDALLAH[1,*], MOHAMMAD HASSAN[2], AHMAD ELTAIEF[3]
[1]Department of Electrical Engineering,
Higher Institute of Engineering and Technology,
Kafrelsheikh, 33511,
EGYPT

[2]Lebara,
SAUDI ARABIA

[3]Department of Computer Science, Faculty of
Information Technology, Applied Science Private
University, Amman,
JORDAN

**Corresponding Author*

Abstract: - The Industrial Internet of Things (IIoT) revolutionizes industrial operations by integrating smart sensors, advanced analytics, and machine learning to optimize manufacturing and supply chain processes. However, the rapid growth of IIoT presents tremendous cybersecurity challenges, as the increasing number of connected devices makes these systems more susceptible to cyberattacks. Intrusion detection systems (IDS) are vital for monitoring network and system activity to prevent and identify malicious attacks in IIoT. This paper discusses the role of AI- and computational-intelligence-driven IDS in protecting IIoT networks, addressing crucial challenges such as device heterogeneity, real-time operational limitations, and the mission-critical nature of industrial systems. Traditional IDS solutions are challenged by high false positives and dynamic cyber threats, and AI-based methods are a necessary evolution. Machine learning (ML) and artificial intelligence (AI) improve the effectiveness of IDS by enhancing anomaly detection, minimizing false alarms, and learning new attack patterns in dynamic IIoT environments. The paper discusses current IDS technologies, emphasizing their strengths and weaknesses in IIoT security. AI-based IDS utilizes deep learning, hybrid models, and feature engineering to improve accuracy and efficiency. Furthermore, future directions in industrial cybersecurity include federated learning for decentralized security, lightweight AI models for resource-scarce IIoT devices, and blockchain for secure data exchange. AI-based IDS can protect critical infrastructure and provide reliable, safe, and resilient industrial processes in IIoT by addressing these issues. The present study highlights the growing role of AI and computational intelligence in making industrial cybersecurity resilient against evolving threats.

Key-Words:- IIoT, Computational Intelligence, Artificial Intelligence, Machine Learning, Intrusion Detection System, Deep Learning.

Received: June 9, 2025. Revised: August 24, 2025. Accepted: October 12, 2025. Published: June 23, 2026.

1 Introduction

The IoT is putting the internet to work in our daily lives, linking ordinary things together via the Internet. IoT finds wide applications in healthcare, agricultural, manufacturing and smart cities, innovating them and pumping the operational efficiency. Hence, IIoT is a new subclass of IoT designed for true industrial use. IIoT refers to the fusion of interconnected industrial devices, sensors, huge data analytics and machine learning that can improve manufacturing and supply chain operations, [1]. IIOT ecosystem is probably used in various applications such as predictive maintenance [2], asset tracking and management [3], energy management [4] smart manufacturing [5]. However, IIoT devices are inducing a different level of complexity in the security landscape. IIOT

Infrastructure remain under constant threat from Cybersecurity attacks such as data breaches, malware and DoS (Denial of Service) attacks. The growing trend of cybercrime Cybersecurity Ventures estimates that by 2025 the total annual cost of cybercrime will exceed USD 10.5 trillion, meaning there is an urgent need for robust security systems, [6]. In 2023, approximately 70% of industrial organizations were surveyed and reported experiencing a cyberattack within the previous year, [7]. This has provided IDS an important approach to these security issues. Integrated Detection System (IDS) monitors the datatraffic and system action in the network for any mishap or policy infraction. They are fundamental for the detection and responding to potential security threats and defending information systems against unauthorized use and attacks [8], as shown in Figure 1.

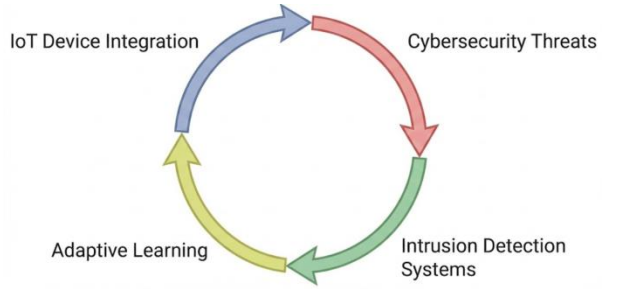

Fig. 1: IoT cycle and security touchpoints, illustrating the full lifecycle from device sensing through network transmission to cloud processing, with key attack surfaces and IDS monitoring positions annotated at each layer
Source: [8]

To address these problems, a IDS suitable for IIoT systems is necessary. Machine learning and AI-based next-generation IDS are IIoT centric as they focus on never being trained on data older than October, 2023 while having a higher detection percentage rate with complete elimination of false-positive, [9]. The impact of deploying IDS in IIoT environments can decrease cyber-attacks by 60% [10]. Moreover, IDS's capability of persistent monitoring and continuous learning ensures ongoing protection of industrial systems from evolving threats. Ongoing research and development on IDS technologies is essential to cope with the dynamics of cyber threats.

In this paper, the organization is designed to help IIoT deployment decision makers decide what suits them best. This paper is organized as follows: in Section 2, we review datasets and evaluation protocols with an emphasis on industrial realism. We also provide metrics that are not limited to just accuracy, including latency, size of the model and calibration in Section 3 Section 4 presents a survey of IDS families and their direct mapping to IIoT boundaries (edge/gateway/cloud) In Section 5, we distill insights on adversarial robustness and explainability. Section 6 provides actionable guidance on challenges and future directions.

Unlike existing surveys that focus primarily on benchmark accuracy, this survey makes three distinct contributions that set it apart from prior related work: (i) it systematically maps the reviewed IDS method families to practical deployment tiers (edge, gateway, or cloud) using reported latency, model size, and computational cost; (ii) it introduces calibration-aware evaluation as a required reporting dimension by recommending Expected Calibration Error (ECE) or Brier score alongside macro-F1 and accuracy; and (iii) it consolidates mathematical formulations, robustness considerations, and deployment-oriented comparisons into a unified reference framework for IIoT intrusion detection. All comparative tables in this survey are original compilations prepared by the authors from the cited studies and are not reproduced from a prior publication.

2 Background

2.1 IIoT System Architecture

The IIoT system architecture typically comprises multiple layers, as illustrated in Figure 2, each with distinct functions, thereby enhancing the overall efficiency and reliability of industrial processes. These are the perception, network, and application layers, [11].

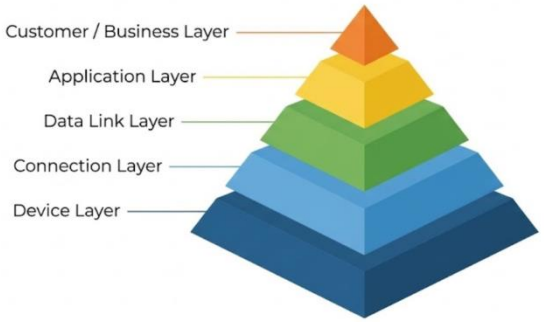

Fig. 2: Architecture of IDS
Source: [11]

Sensors and devices, which collect information from the outside world, form the perception layer. This data includes but is not limited to temperature, pressure, humidity, and other critical parameters needed to monitor industrial processes. This is based on wired or wireless communication technologies,

such as Ethernet, Wi-Fi, or 5G for secure and efficient data transmission, [12]. The processing and interpretation of data happen at this layer which enables actionable insights and informs the decision-making functions in industrial systems. The application layer leverages deep analytics, machine learning, and AI to turn raw data into just-in-time insights. Creation, efficiency improvement and cost reduction is the indispensable layer for application of industrial activity. The IIoT architecture uses the cloud for preprocessing data close to its source, thereby reducing latency and minimizing bandwidth usage, [13].

2.2 IIoT Datasets

Assessing IDS in IIoT security studies heavily depends on a few core datasets. These datasets vary in scope, the variety of simulated attacks they cover, and the targeted IIoT scenarios they encompass. Prominent examples include the Edge-IIoTset [14], NSL-KDD [15], UNSW-NB15 [16], IoT-Botnet [17], ToN-IoT [18], CICIDS2017 [19], DS2OS [20], WUSTL-IIoT-2021 [21], and CICIoT2023 [22] as in Table 1 (Appendix).

These datasets collectively provide a strong foundation for intrusion detection research across IIoT, IoT, and broader cyber-network environments. Edge-IIoTset and WUSTL-IIoT-2021 are especially valuable for IIoT-focused evaluation, while NSL-KDD and UNSW-NB15 remain widely used benchmark datasets for modeling network traffic and attack behavior. IoT-Botnet and ToN-IoT are particularly useful for studying IoT-specific threats such as botnets and device-level vulnerabilities. CICIDS2017 and DS2OS further expand coverage by including diverse cyberattacks and realistic threat scenarios. Together, they enable comprehensive, reliable, and comparative assessment of IDS performance.

3 Intrusion Detection Systems

IDS are essential elements in protecting IIoT environments. IDS can be categorized into Network-based IDS (NIDS) and Host-based IDS (HIDS). NIDS observes network traffic for malicious behavior, [23]. It also inspects traffic patterns to identify possible intrusions. HIDS, however, observes individual hosts or devices for indications of compromise. It inspects system logs, file integrity, and other host-based activities by [24], as in Figure 3.

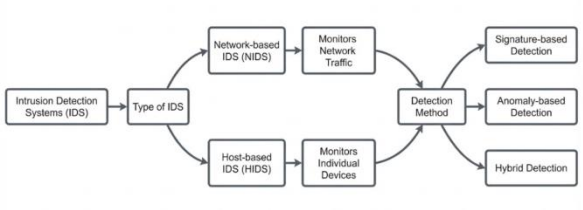

Fig. 3: IDS in IIoT Environment
Source: [24]

IDS utilizes several detection mechanisms to detect probable threats. Signature-based detection relies on predetermined patterns or signatures of established threats, [25]. It detects known attacks efficiently but is ineffective in detecting new or unknown attacks. Anomaly-based detection establishes a normal baseline of behavior and identifies any deviations from this established baseline. It detects unknown threats but has the possibility of higher false positives. Hybrid Detection blends signature-based and anomaly-based detection techniques, [26]. It offers a trade-off between detecting known threats and identifying new ones.

3.1 IDS Dataset

IDS for IIoT is based on some benchmark datasets to train and test detection models. The UNSW-NB15 dataset is commonly utilized to explore network attacks using ML and DL methods. This project contains datasets collected specifically for Threats in IIoT environment, WUSTL-IIoT-2021 and Edge-IIoTset2023 which showcase ensemble model based techniques or optimization methods used to improve detection accuracy. CICIDS2017 and ToN-IoT are common datasets that allows for the usage of deep learning models which can help with better intrusion detection, [9]. In addition, DS2OS and EdgeIIoT-2021 datasets accommodate hybrid neural networks and LSTM models solving scalability and security concerns within IIoT systems. According to Table 2 (Appendix), these datasets improve the knowledge and development of IDSs in IIoT by utilizing diverse and inclusive data for testing/evaluation purposes.

4 Advancements in IDS for IIoT

We then group the IDS solutions by their methodological family and link to possible IIoT deployment loci (edge/gateway/cloud). To facilitate comparisons and avoid redundancy across subsections, we capture the key trade-offs in a comparison matrix (Table 3 and Table 4 in Appendix as well as in Table 5 and Table 6 and

Table 7 in Appendix) and provide a synthesis at the end of each subsection that prioritizes deployment.

The introduction of IDS in IIoT, using advanced methods of artificial intelligence and machine learning has contributed a lot to improved security. Traditional IDS suffers from high false positive rates and low scalability, but advanced methods use deep learning, ensemble designs, and hybrid optimization to improve the detection of threats. As IIoT ecosystems expand, recent research aims to maximize computational efficiency, improve adaptability in the face of new attack vectors, and maintain lightweight deployment support for resource-constrained devices, as shown in Figure 4.

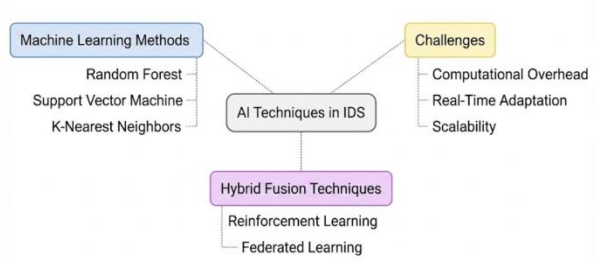

Fig. 4: IIoT Detection Systems
Source: created by the authors

4.1 Machine Learning and Deep Learning for Intrusion Detection in IIoT

[27] examined the security context of CPS under IIoT edge computing architecture and then identified outstanding DoS, ransomware, and MITM cybersecurity threats. Furthermore, the study emphasized the use of sophisticated security frameworks to protect the critical infrastructures of IIoT systems. In the proposed system, all researchers noted that IDS should be enhanced using ML, DL, FL, and blockchain technologies. Such security models may not be adopted in industrial settings, where timely decision-making is paramount.

[28] proposed a model that combines a CNN-IDS with a blockchain-based security model. The new combination was designed to enhance the security of SDN-based IIoT applications at both the network and application layers. Using CNN for sophisticated anomaly detection and blockchain for data integrity and secure communication, the authors sought to develop a more robust security solution against attacks such as rule manipulation and command injection.

A detailed study of IDS-IoT security using AI was conducted in [29], where training and deployment frameworks were compared against each other. They explored centralized, distributed, and federated learning approaches along with their advantages and limitations for IoT security. The authors highlighted that ML and DL can be an efficient approach for sIDS, especially in cloud, fog and edges of the IoT sys- tems which is becoming more important building blocks of IoT setups. A recent work [30] utilized various ML and DL methods to help improve anomaly detection, and cyber-attack resistance in IIoT systems by providing an overview of AI-based IDS. The authors showcased the potential of AI-based IDSs to detect a multitude of cyber threats by performing an extensive bibliographic review of deployed works on AI-based systems for intrusion detection and comparing these models based on some number of existing datasets. While the prospects for increasing IDS based on AI are very positive, specifying two aspects still need further research and development efforts: the false negative issue and optimization making it in real-time as illustrated through Figure 5.

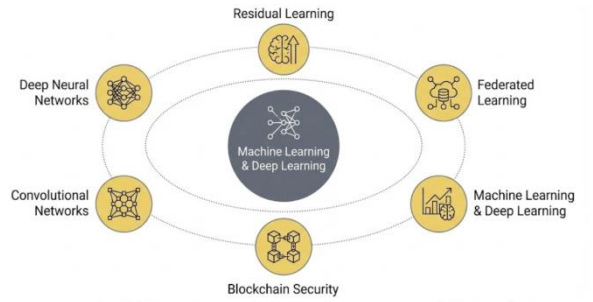

Fig. 5: IDS methods and deployment feasibility
Source: created by the authors

Comparative map of common ML/DL architectures (RF, SVM, CNN, LSTM, attention, GAN/VAE) with respect to IIoT constraints—icons indicate latency and model size; the same edge/gateway/cloud overlay as in Figure 4 is used; arrows show data/gradient flow where federated learning applies. While temporal models can help improve detection on sequential traffic, they typically go beyond edge latency/footprint without compression. To satisfy requirements from IIoT scenarios, such as unexplored device types with Deep Learning-based IDS designed by [31], and to be able to process data in a time-critically way. This method integrates deep neural network with adaptive schema to handle heterogeneous and non-stationary IIoT data.

This provides stronger detection capability in places where multiple types of devices reside. Thus, it makes the model more worthy of being trusted in IIoT systems. However, its biggest weakness is the computational overhead, as deep learning algorithms can overwhelm the processors in these

scenarios and require large amounts of memory, which impedes real-time adaptation, especially in industrial settings where space for overhead is limited.

[32] proposed Pelican, a deep residual network for NID, which employs residual learning principles to enhance detection performance and scalability. Nonetheless, the limitation of this model is that it requires high computational resources, making it challenging to deploy in real-time on devices with limited resources or in low-computational-power environments, such as edge or fog computing. A consolidated summary of these approaches is provided in Table 3 (Appendix).

End-to-end deep learning attains the highest macro-F1 on public datasets but incurs latency and memory that challenge edge nodes. Prefer compressed DL (pruning/quantization/distillation) or shallow back-ends at the edge, and keep full models on gateway/cloud where sub-50 ms latency is not mandatory.

4.2 Hybrid and Ensemble Learning Approaches for Improved IDS Performance

[33] also investigated the deployment of ensemble ML approaches for anomaly detection in IDS for IIoT. Ensemble learning involves integrating multiple machine learning models to enhance overall detection accuracy by leveraging the potential of individual classifiers. It enables the improvement of robustness and reliability by mitigating both false negatives and positives. The ensemble method integrates SVM, RF, and DT models to increase classification accuracy. [34] suggested a Hybrid Network Intrusion Detection System (NIDS), a combination of supervised machine learning algorithms, i.e., Random Forest and Neural Network, to overcome the drawbacks of traditional rule-based Security Information and Event Management (SIEM) systems. Following the preprocessing of network traffic data from the CICIDS2017 dataset, which included both attack and benign traffic data, feature extraction and classification using Random Forest and Neural Network classifiers were performed. The validation of the model using the CICIDS2017 dataset showed an increase in intrusion detection performance, with high precision, recall, and accuracy scores. The hybrid method has the added benefit of detecting subtle attack patterns that rule-based models might overlook.

A deep learning and reinforcement learning based Intrusion Detection System for IIoT systems is proposed in [35]. The model uses CNN for feature extraction from network traffic and reinforcement learning to improve the adaptivity of the model regarding dynamic threats. A two-step learning system, reinforcement learning continuously strengthens the model by updating detection methods based on feedback from previous detection experiences. This active learning in particular is well suited because attacks and other features can change significantly in dynamic IIoT systems. CICIoT2023 is a labeled network traffic dataset containing normal and attack data suitable for IIoT systems which was used to validate the model. At its core, it reinforces the promises of combining a strength of deep learning on different problems with reinforcement learning in improving threat detection. That being said, this hybrid model has an inherent drawback in terms of computational cost to train deep learning and reinforcement learning models. These models are extremely costly in terms of computing and memory resources, thereby restricting their deployment not only on resource-limited Industrial Internet of Things (IIoT) devices but also for real-time applications where low latencies are required.

[36] proposed an IDS that combines deep-learning methods with conventional rule-based detection to improve detection accuracy while reducing false positives. Similar to rule-based systems, it can be used not only for known attack detection (interpretable and fast) but can also learn complex patterns in data automatically. The described work validated the model against the CICIoT2023 dataset. The deep-learning part analyses network-traffics pattern,the rule-based part detects recognized attacks defined using pre-defined patterns or rules. Under the cited CICIoT2023 evaluation setting, this hybrid method reported a good detection accuracy (97.8%) and few false positives. An advantage with this model is that it can provide advanced detection capabilities while having the accuracy of rule-based models. The primary limitation is that predefined rules are unable to address new or unfamiliar types of attacks. Another consideration is the complication in training and deployment introduced by any deep-learning and rule-based components, which means that system design must thoroughly match target attack conditions.

[37] proposed LSTM-CNN-Attention for IIoT intrusion detection, combining LSTM, CNN, and attention mechanisms to improve cyberattack classification. The paper systematically compared various deep learning models, illustrating how the LSTM-CNN-Attention model outperformed other architectures consistently on the leading

performance indicators. The binary classification was almost perfectly correct, while the multi-class classification accurately classifies numerous cyber threats using a minimal loss value of 0.0220%. It results in enhanced threat detection.

[38] proposed a CNN-LSTM-IDS for IIoT usage, utilizing deep learning to enhance cyberattack detection accuracy. It learns from combining CNNs for feature learning and LSTM networks for sequential pattern learning, allowing efficient detection of cyber-attacks. The data used to train the model is based on Edge-IIoTset which contains live network traffic captured from IoT/IIoT services. In the future studies, learning adaptive or light deep-learned models for on-the-fly could be commended to exhibit enormous improvement with respect to adaptability of IIoT framework against arising digital dangers. Table 4 (Appendix) shows the summary of these works.

Macro-F1, and minority-class recall for hybrid and ensemble IDS frameworks are usually better by 5–10 % than the best performance of its individual learners. But these gains are so large that they need ~3× more FLOPs, larger models and also increased prediction latency, which makes it only feasible for gateway or cloud deployments.

4.3 Feature Selection and Optimization Techniques for IDS Enhancement

[39] proposed a novel approach to IIoT network security intrusion detection, which involved utilizing PSO and BA features for selection in conjunction with an RF classifier for intrusion identification. A new intrusion detection scheme exists for IIoT systems because it selects the optimal features and develops optimized classifier operating methods to enhance detection precision. Researchers validated BA as a strong tool for determining critical intrusion detection features in resource-constrained IIoT environments. The summary of these works is shown in Table 5.

[40] proposed an IDS feature selection model based on PSO, GWO, FFA, and GA. Hybridization of these algorithms is used to optimize the model during feature selection and maximizing corresponding detection ratios for critical features in dataset. The features selection process selects significant features and these optimization algorithms are used to improve the classification accuracy. Recently, an IDS based on combinatorial optimization techniques was proposed in [41]. Their model uses advanced optimization techniques they believe will improve the choice of important features, and therefore the performance of network intrusion detection.

Table 5. Summary of Feature Extraction and Optimization Approaches IDS

Author(s)	Method Used	Advantages	Limitations
[39]	PSO, BA for feature selection with RF classifier	Optimized feature selection enhances detection precision	High computational requirements affect real-time performance
[40]	PSO, GWO, FFA, GA for hybrid feature selection	Improved detection accuracy, minimized irrelevant features	Computational overhead due to multiple optimization algorithms
[41]	Combinatorial optimization techniques	Higher detection accuracy, effective feature selection	High computational cost reduces real-time performance
[42]	GA for feature selection, tree-based classifiers (DT, RF)	Enhanced precision adaptability to evolving IIoT threats	Computational burden in large-scale IIoT environments
[43]	SAPGAN with extended WSOA for feature selection	High detection accuracy lowers computational overhead compared to deep learning models.	Depending on large training datasets, the computational cost is still an issue.
[44]	Rule-based feature selection + Deep Feedforward Neural Network	High accuracy, effective threat detection in IIoT	Rule-based selection limits adaptability, high computational overhead

Source: created by the authors

Using both GA and algorithms, [42] put forth a new IDS in IIoT environments. This approach uses GA for feature selection, and tree-based algorithms such as the Decision Trees or Random Forests to enhance classification performance. This hybrid approach improves the detection performance by recognizing advanced attack patterns and dealing with dynamic IIoT environment changes.

[43] presented an enhanced intrusion detection framework employing SAPGAN for IoT security. The framework improves intrusion detection accuracy and less computational overhead compared to baseline deep learning IDS models. First, IoT network data is collected and preprocessed, during which missing values are regenerated with Local Least Squares. [44] introduced a combined rule-based features selection and a deep feedforward neural network model as a DL-IDS for IIoT. The problem to be solved was the detection of conventional and emerging cyberattacks that threaten the security of IIoT systems. Furthermore, the computational overhead of the deep feedforward neural network may be a challenge in real-time implementation, especially in resource-constrained IIoT environments. Future work should emphasize improving model flexibility to adapt to new threats and minimizing computational overhead for practical use.

Feature selection and optimization techniques reduce dimensionality and improve macro-F1 while

lowering FLOPs by approximately 20–35% across reported studies. These methods also strengthen calibration by removing noisy or redundant features, which stabilizes probability estimates. For time-critical IIoT settings, embedded or filter-based selection offers the most effective balance of model size, latency (typically <30 ms), and stable detection performance.

4.4 Federated Learning and Privacy-Preserving Techniques in IDS

[45] proposed an FL-IDS to enhance scalability and maintain privacy among IIoT devices. The model involves a distributed learning process, wherein several devices collectively train the model without exchanging sensitive information, thereby maintaining data privacy while reducing the need for centralized data storage. The summary of these works is shown in Table 6 (Appendix).

[46] presented a Federated learning (FL IDS) consisting of CNN, attention mechanisms, and variational autoencoders (VAEs) that are enriched to boost IIoT environments network security while keeping data privacy. Fl differs from the traditional model of IDS based on centralized datasets, in which several IIoT clients can be trained. It combats privacy threats in data collection by centralized gathering while still allowing for a cost-effective threat detection process.

[47] introduced an FTL method for IIoT NID using a CNN as the core mechanism for distributed learning. The approach entails dividing IIoT data among client and server devices, where local client models are individually trained, and their learned weights merge to update a global server model. [48] proposed an FL-based IDS to address the scalability and privacy concerns associated with centralized AI-based solutions for IoT networks. FL supports collaborative learning without compromising data privacy by distributing the training process among multiple clients, making it best suited for decentralized IoT systems. The work employed the TON-IoT dataset, which mimics realistic IoT network traffic and maps every IP address to a unique FL client, ensuring data locality. Several aggregation mechanisms and pre-training strategies were attempted to counter data heterogeneity, a common problem in distributed learning systems.

[49] proposed an FFL-IDS incorporating CNN to handle the scalability and privacy concerns in IIoT networks. This framework deploys fog computing for low-latency detection and preserves data privacy by distributing training models across IIoT nodes. The FFL-IDS was tested on two datasets containing diverse network attack scenarios. These all indicate that the presented FFL-IDS is an efficient, privacy-guaranteed, and high-performance technique to safeguard IIoT networks against sophisticated cyber threats.

[50] provide a FIDS for IIoT networks, proposing NIDS-FGPA. This aspect creates an agenda where further improvement revolves around reducing cost effects at encryption or aggregation. Federated learning achieves strong macro-F1 (≈ 95–97 %) while preserving data locality, but communication rounds add latency that can exceed cycle-time limits in fast-loop industrial plants. Model sizes and FLOPs grow with deeper CNN-based FL clients, restricting low-end edge participation. Calibration under non-IID data remains a challenge, as client drift produces inconsistent confidence scores without robust aggregation.

4.5 GANs, Transfer Learning, and Advanced Neural Networks for IDS

[51] examined using Generator–discriminator frameworks (GAN) for data anomaly detection in IIoT networks. Generators synthesize traffic, discriminators learn boundaries. Helpful for imbalance and hard negatives, but training is unstable and may break protocol semantics, validate on scenario splits. Under the G and D model architecture, artificial network traffic data is generated by G and then fed to D, which decides the real and the synthesized data. [52] proposed applying transfer learning to IDS in IIoT networks to strengthen the system's capability of detecting new and emerging types of attacks. The ability to leverage pre-trained knowledge in similar domains is achieved through transfer learning, which enables the model to generalize more efficiently to unseen attack scenarios in IIoT network traffic data.

[53] suggested an HDGAN for identifying malicious attacks in IoT networks. The model was designed to address the challenges posed by IoT's time-varying and resource-constrained environment, where there are insufficient resources, leading to high false detection rates. [54] explained that the Constrained Twin Variational Auto-Encoder (CTVAE) is a proposed deep neural network model designed to enhance IDS performance in resource-constrained IoT environments. The results indicate how CTVAE can improve accuracy and F1 by 1%.

[55] proposed an IDS that employs LSTM networks and attention mechanisms to detect time-changing attacks in IIoT networks. LSTM networks effectively capture time dependencies in network traffic, as is necessary for detecting attacks that build over time, e.g., APTs. The IDS model

achieved remarkable accuracy of 97.6%, detecting a broad spectrum of attack types with minimal delay (80/20 split). The strongest facet of the model lies in that it can supply real-time intrusion detection with low latency.

[56] proposed a DLNN for IIoT cyberattack detection, combining several learning approaches to improve the detection. The hybrid model combines random neural networks with other deep learning architectures, including CNNs and RNNs, to leverage the inherent strengths of each method. The most important benefit of the hybrid model is that it can unify the strength of multiple neural network types to improve detection performance. However, the model complexity is a drawback as the involvement of more than one neural network requires more computational effort and complicates the model's optimization and cost-effectiveness in real-time systems, particularly in resource-poor settings.

Generative, transfer-learning, and advanced sequence models improve zero-shot macro-F1 by 3–6 %, strengthening detection of unseen attacks. These approaches often rely on large models with high FLOPs and significant training time, resulting in latency unsuitable for edge deployment.

4.6 Lightweight, Resource-Efficient IDS for Real-Time IoT Security

[57] introduced a lightweight smart IDS for IIoT based on deep learning algorithms emphasizing efficiency and scalability. This IDS is intended to be lightweight to facilitate effective operation in environments with limited resources, including embedded systems and edge devices in IIoT networks. The system utilizes a lightweight deep learning architecture that balances computational efficiency and detection accuracy.

[58] presented a system to meet the requirements, including device diversity and time-critical data processing requirements in IIoT scenarios. We designed a deep learning-based IDS focused on custom-engineered IIoT scenarios to fit these requirements. The approach is based on deep neural networks combined with adaptive methods to control the heterogeneous nature of IIoT devices and data, as the latter is not steady. [59] proposed a lightweight IDS that integrates CNN and SVM for efficient intrusion detection in IIoT systems. The model employs CNNs and SVMs for feature classification as normal or malicious.

[60] presented a multi-layer perceptron (MLP)-designated IDS that incorporates feature selection mechanisms to optimize detection accuracy while reducing computational complexity.

[61] introduced a CNN-IDS for IoT networks, answering the demand for more efficient security measures against high-scale cyberattacks. In contrast to traditional time-sequential models such as LSTM, this paper utilizes CNNs to represent sensor data as convolutional operations. It enables the model to learn spatial and temporal patterns for identifying attacks.

[62] introduced a paradigm for tackling the OSR issue of NIDS for the IoT settings in this work. The OSR issue arises when IDSs find it challenging to recognize new attacks not encountered during training. The introduced technique uses image-based embeddings of network packet-level data and represents them spatially and temporally to exploit essential patterns to identify unknown or previously unseen attacks. The overall limitations are shown in Table 8 (Appendix).

Lightweight CNN, SVM, and MLP IDS designs achieve real-time detection (< 50 ms) with small model sizes (< 20 MB) and reduced FLOPs, making them practical for embedded edge devices. Their macro-F1 typically ranges from 88–92 %, slightly below full deep models but sufficient for rapid triage.

[63] proposed an ML-IDS built on the ToN-IoT dataset. Their methodology leverages a two-layer classification structure, one of which initially recognizes abnormal network behaviors and subsequently differentiates the targeted behaviors into varieties of attacks.

[64] introduced a sophisticated intrusion detection framework, Inception-CNN and BiGRU, to enhance detection accuracy in IIoT systems. Their methodology processes attack traffic using BiGRU to extract sequential features, which facilitate capturing temporal dependence in the data. Inception-CNN for feature representation works well with sophisticated and high-dimension input data.

[65] conducted a thorough examination of ML-IDS in IoT environments, with an emphasis on categorizing it into supervised, unsupervised, and hybrid learning methods. A major conclusion provided by their survey was the inherent ability of deep learning to detect new attacks and learn new vulnerabilities. It then noted that despite the promise of these advanced techniques, ML-based IDS also had a major limitation: its application on IoT devices with limited resources.

[66] created an expert system that can protect IIoT devices with high-impact consequences, such as power networks and transportation systems. The system used anomaly detection and RL to detect and prevent attacks like denial of service, data

tampering, device hijacking, and physical tampering. It consists of three main parts: a data collector, a data analyzer, and a data actuator, which collectively detect suspicious behavior. The summary of these works is shown in Table 8 (Appendix).

Beauty [67] developed an NIDS tailored for industrial IoT environments to solve some of the main shortcomings of standard IDS. The new model combines Autoencoders (AE) with Convolutional Neural Networks (CNN) to enhance intrusion detection effectiveness. [68] proposed an extension of the UNSW-NB15 dataset using deep learning methods to enhance the robustness and reliability of IDS in detecting advanced attacks. They rely on CNN and other deep networks to enrich the dataset in aspects such as quality and efficiency during feature extraction, addressing the complications of advanced pattern recognition in the attack and ultimately enhancing the accuracy levels of detection. [69] discussed the application of LSTM networks and ML models to mitigate future security threats to IIoT. The projected model, i.e., the ERT-based NIDS, leverages the effectiveness of LSTM and other machine learning approaches to identify and safeguard against cyberattacks on IIoT domains. [70] examine the application of NIDS in industrial and robotic systems, considering methods such as machine learning, deep learning, and hybrid approaches. To enhance intrusion detection quality, real-time responsiveness, and response efficiency. It highlights the inadequacies of IDS in handling the complexity of CPS and ICS. Different NIDS approaches are evaluated based on their ability to identify anomalies and cyberattacks in industrial settings.

They have presented an AttackNet model which is based on deep learning for botnet attack detection and classification from the adaptive CNN-GRU architecture, [71].

The GRU seems to fit very well for the time series problem of scanning multivariate and sophisticated botnet attacks, thereforethis model works CNN to extract features, then GRU used for learn the order. The model was extensively tested against the N_BaIoT dataset and real-time traffic from IoT and IIoT devices to show its robustness against evolving cyber threats.

To meet this demand for intelligent cyber defense systems, [72] comprehensively surveyed DRL-IDS in the IoT networks. A systematic analysis of the existing DRL-based IDS techniques was performed, and these techniques were classified into five different groups including DQN, healthcare, hybrid schemes, and other novel approaches in this research. The IDS frameworks use DRL (Deep Reinforcement Learning) to learn actively detect which identifying and defending malicious threats on IoT networks can limit human involvement. [73] proposed using the BO-GP framework with ensemble tree-based learning models to improve IIoT network intrusion detection. BO-GP is used to tune the hyperparameters of the ensemble learning model and improve the machine learning model performance for attack detection. This is particularly useful in IIoT networks, as security threats become more advancedIn other words, results which only ever involve a single dataset are likely to create some false sense of optimism; scenario splits (site, device, time) and at least one zero-shot test are essential for estimating out-of-domain performance.

We see that performance on any single dataset can vary greatly — macro-F1 variance across datasets often exceeds 15% suggesting that generalization cannot be guaranteed by high accuracy on a single benchmark. Abstract: Multi-class classification with large DL models is challenging, as the strong macro-F1 on in-dDataset tests do not translate to out-of-Distribution (OOD) data where calibration tends to drift under shifts like time, device or site. Due to model size and FLOPS, they are mostly limited to gateway/cloud tiers with little real-time edge deployment capabilities. The macro-F1 scores of lightweight models are below, but their latency and calibration of the across various environments is more stable. Inevitably, this means that any IIoT deployment must be evaluated at a scenario-split level; coupled with external validation and calibration reporting so monitored performance in the field can be relied on for making business critical decisions, [1].

4.7 Adversarial Robustness in IDS for IIoT

Periodic, protocol-constrained IIoT traffic with safety-critical control loops that are delay intolerant. Such properties enable sufficiently small, protocol-conforming perturbations to cross anomaly thresholds and pose practical risks in multi-vendor plants through poisoning at data collection (or training) time and model stealing. Although certified defenses currently lag in throughput, [74], they provide formal guarantees. One practical deployment pattern is to run a thin primary detector with rule-based guardrails at the edge, and having more heavyweight defenses (e.g., those based on adversarial training or certified smoothing) deployed within nodes in the gateway/cloud where latency budgets permit, [75].

For instance, early sequence-based IDS

frameworks such as RNNs model temporal dependencies in network traffic as shown by [76] and such structures were later previously studied for adversarial evasion, thus motivating robustness-aware training. [77] also evidence that adversarial training in industrial control systems preserves resilience, albeit at a cost of increased inference latency and computational load which demonstrates the robustity–utility trade-off which is critical to IIoT operations.

4.8 Explainable IDS for Operator Trust

Immediate and safe action on alerts is essential in safety-critical IIoT environments, which means having immediate visibility into why an alert was triggered by a detector. SHAP or LIME are examples of post-hoc explanation methods that yield feature-level attributions for tabular network flows. Gradient-based and attention roll-out techniques help interpret decisions in sequence models. Explanations are most beneficial when combined with operator playbooks mapping common attribution patterns to suggested next checks (e.g., packet capture scope, PLC tags for inspection, likely root causes), providing all the right context to minimize the mean time to investigation and alleviate alert fatigue, [78]. Deliver richer global dashboards at the gateway/cloud to inform policy changes, model refreshes, and post-incident retrospectives Calibrated scores and actionable explanations enable better operator trust, reduced false positive rates, and auditable decision-making aligned with safety, [79]. By employing sophisticated explainable AI methods for ICS intrusion detection, [80] pairs feature-level attributions with operator dashboards to enable transparency in the decision-making process and post-incident audits.

4.9 Domain Adaptation and Transfer Learning for IIoT

Reported gains on benchmark splits frequently collapse under cross-plant, cross-time, or cross-device shifts, reflecting real IIoT variability in traffic mix, firmware, and process schedules. Transfer learning and domain adaptation can mitigate accuracy loss under cross-plant or cross-time shifts, though negative transfer remains a risk as in [81]. These benefits come with risks: negative transfer is common when the source plant is poorly matched to the target; calibration often drifts even when accuracy holds; and naïve fine-tuning can overfit rare events or leak across sessions. For deployment, we recommend (i) scenario-driven evaluation splits by site/device/time that mirror rollouts; (ii) at least one zero-shot external dataset to estimate out-of-domain behavior; and (iii) reporting macro-F1 and calibration shift (ECE/Brier) alongside accuracy as in [82].

5 IDS Performance Evaluation Metrics in IIoT Networks

In addition to accuracy, we also report on macro-F1 (class balance), calibration (e.g., ECE or Brier score), inference latency (ms, batch = 1 on specified hardware), model size (MB/parameters) and estimate FLOPs as discussed in Table 9 (Appendix). For robustness, evaluation should also include attack success rate in both white-box and black-box settings, and provide a robustness–utility summary that connects the Latency and Throughput to the Data-Set Decision Accuracy (DDA). Footnote no statistical significance comparisons across papers are made, because surveyed results come from heterogeneous studies with different datasets, splits and protocols.

5.1 Mathematical Formulations of Key IDS Components

This subsection provides formal mathematical definitions and representative model formulations for the primary AI-based IDS components surveyed in this paper. A unified notation is used to improve rigor and comparability across methods.

5.1.1 Formal Definitions of Evaluation Metrics

Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively, over C attack classes. The key metrics are defined as follows:

Macro-F1 Score: For each class $c \in \{1, \ldots, C\}$, precision $Pc = TPc/(TPc + FPc)$ and recall $Rc = TPc/(TPc + FNc)$. The per-class F1 score is:

$$F1c = 2 \cdot Pc \cdot Rc / (Pc + Rc) \quad (1)$$

$$F1macro = (1/C) \cdot \Sigma F1c \quad (2)$$

Macro-F1 is preferred over accuracy for imbalanced IIoT datasets because rare attack classes should not be underweighted.

Expected Calibration Error (ECE): Let Bm denote the m-th confidence bin, conf(Bm) the mean predicted confidence, and acc(Bm) the empirical accuracy within Bm. With M bins:

$$ECE = \Sigma (|Bm|/n) \cdot |acc(Bm) - conf(Bm)| \quad (3)$$

where n is the total number of samples. Low ECE indicates that model confidence scores align well with empirical accuracy, which is essential for

trustworthy alert thresholding in IIoT dashboards.

Brier Score:

$$BS = (1/n) \cdot \Sigma (fi - oi)^2 \quad (4)$$

where fi is the predicted probability and oi ∈ {0,1} is the true outcome. The Brier score captures both calibration and discrimination; lower values indicate better performance.

5.1.2 Deep Learning Loss Functions for IDS

Binary Cross-Entropy (BCE) loss for binary normal/attack classification is given by:

$$LBCE = -(1/n) \cdot \Sigma [yi \log(pi) + (1 - yi) \log(1 - pi)] \quad (5)$$

where yi ∈ {0,1} is the true label and pi = σ(fθ(xi)) is the sigmoid-activated model output.

Categorical Cross-Entropy (CCE) loss for multi-class attack classification over C classes is:

$$LCCE = -(1/n) \cdot \Sigma \Sigma yi,c \log(pi,c) \quad (6)$$

where yi,c ∈ {0,1} is the one-hot class label and pi,c = softmax(fθ(xi))c. CNN-, LSTM-, and hybrid deep-learning IDS models are commonly trained by minimizing this loss with stochastic gradient descent or Adam.

This highlights an increasing focus on lightweight and explainable IDS architectures [61], [83]; federated and decentralized learning frameworks [31]; in addition, emerging self-supervised paradigms such as SimCLR-based models reflected a paradigm shift towards interpretable with resource-efficient industrial cybersecurity, [77].

In Table 10 (Appendix) we present a syncretized mapping of the preferred deployment tiers and matching computational–latency characteristics, within IIoT and ICS environments with classification families of artificial intelligence–based intrusion detection methods. The data are calculated directly from the studies cited and fit to ensure fact-based correlation with published experimental results. The table notes that CNN-based and hybrid deep learning approaches are generally employed at the cloud or gateway level, where sufficient compute resources enable complex model inference and deep packet inspection. Unlike deep learning networks, which target global variables and that may require expensive computational power to reach real-time analysis results in any kind of hardware layer, RNN- and GNN-based models have become more frequently adapted for near real-time fog or edge layers capable of computing temporal /spatial/relational patterns, [52].

Federated and self-supervised learning frameworks demonstrate versatility for cloud–edge collaboration, balancing privacy preservation and performance, while adversarial defense models emphasize robustness at the expense of higher training costs. Lightweight and explainable AI architectures, on the other hand, offer interpretable and resource-efficient solutions that are suitable for constrained edge devices. Collectively, the mapping in Table 10 (Appendix) clarifies deployment trade-offs across model families and supports practical selection of IDS approaches for heterogeneous IIoT.

6 Challenges and Future Directions

In these contexts, lightweight yet robust deep learning techniques should be integrated into IDS solutions, utilizing knowledge distillation and pruning to minimize computational overhead while maintaining high detection accuracy. To scale up, edge computing systems and distributed detection frameworks can also process data locally before merging the insights, thereby minimizing latency and further reducing bandwidth consumption. On the other hand, GANs for generating artificial data and dynamic resampling techniques help mitigate these data balancing problems, ultimately improving the detection rate of underrepresented classes.

7 Conclusion

The rapid growth of the IIoT presents unparalleled innovation and industrial efficiency opportunities. Yet, with the heightened attack surface and the mission-critical nature of industrial systems, formidable cybersecurity challenges are also presented. This paper discusses the role of AI-powered IDS in addressing these challenges, highlighting its significance in IIoT protection. AI-powered IDS utilizes sophisticated machine learning techniques and artificial intelligence to provide more accurate threat identification and reduce false positives. They make their operations more effective and ensure they can cope with the peculiar requirements of IIoT, including device diversity and real-time operational needs. Some areas remain for future research and development, given the advancements achieved. Improving the resilience of AI-based IDS against advanced attacks, enhancing their scalability for large-scale IIoT applications, and ensuring their reliability in various industrial setups are essential areas that should be explored. By developing these technologies and addressing their shortcomings, we can significantly enhance the

security and robustness of industrial systems, as well as the safe and efficient operation of IIoT systems in the future.

References:

[1] R. Ahmad *et al.*, "Enhancing database security through AI-based intrusion detection system," *Journal of Computing & Biomedical Informatics*, vol. 7, no. 02, 2024, [Online]. https://jcbi.org/index.php/Main/article/view/563 (Accessed Date: February 19, 2025).

[2] M. M. Alani and A. I. Awad, "An intelligent two-layer intrusion detection system for the internet of things," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 683–692, 2022.

[3] M. A. Alohali, F. N. Al-Wesabi, A. M. Hilal, S. Goel, D. Gupta, and A. Khanna, "Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment," *Cogn Neurodyn*, vol. 16, no. 5, pp. 1045–1057, Oct. 2022, doi: 10.1007/s11571-022-09780-8.

[4] B. Alotaibi, "A survey on industrial Internet of Things security: Requirements, attacks, AI-based solutions, and edge computing opportunities," *Sensors*, vol. 23, no. 17, p. 7470, 2023.

[5] K. Bansal and A. Singhrova, "Review on intrusion detection system for IoT/IIoT -brief study," *Multimed Tools Appl*, vol. 83, no. 8, pp. 23083–23108, Aug. 2023, doi: 10.1007/s11042-023-16395-6.

[6] I. Bibi, A. Akhunzada, and N. Kumar, "Deep AI-powered cyber threat analysis in IIoT," *IEEE Internet of Things Journal*, vol. 10, no. 9, pp. 7749–7760, 2022.

[7] D. Hamouda, M. A. Ferrag, N. Benhamida, and H. Seridi, "Intrusion detection systems for industrial internet of things: A survey," in *2021 International Conference on Theoretical and Applicative Aspects of Computer Science (ICTAACS)*, IEEE, 2021, pp. 1–8. https://doi.org/10.1109/ICTAACS53298.2021.9715177.

[8] S. Islam, D. Javeed, M. S. Saeed, P. Kumar, A. Jolfaei, and A. K. M. N. Islam, "Generative AI and Cognitive Computing-Driven Intrusion Detection System in Industrial CPS," *Cogn Comput*, vol. 16, no. 5, pp. 2611–2625, Sep. 2024, doi: 10.1007/s12559-024-10309-w.

[9] C. P. Kaliappan, K. Palaniappan, D. Ananthavadivel, and U. Subramanian, "Advancing IoT security: a comprehensive AI-based trust framework for intrusion detection," *Peer-to-Peer Netw. Appl.*, vol. 17, no. 5, pp. 2737–2757, Sep. 2024, doi: 10.1007/s12083-024-01684-0.

[10] G. Karacayılmaz and H. Artuner, "A novel approach detection for IIoT attacks via artificial intelligence," *Cluster Comput*, vol. 27, no. 8, pp. 10467–10485, Nov. 2024, doi: 10.1007/s10586-024-04529-w.

[11] V. Sobchuk, R. Pykhnivskyi, O. Barabash, S. Korotin, and S. Omarov, "Sequential intrusion detection system for zero-trust cyber defense of IOT/IIOT networks," *Advanced Information Systems*, vol. 8, no. 3, pp. 92–99, 2024.

[12] M. Srinivasan and N. C. Senthilkumar, "Intrusion Detection and Prevention System (IDPS) model for IIoT Environments Using Hybridized Framework," *IEEE Access*, 2025, Accessed: Feb. 19, 2025. https://doi.org/10.1109/ACCESS.2025.3538461.

[13] A. Serhane, E.-M. Hamzaoui, and K. Ibrahimi, "IA Applied to IIoT Intrusion Detection: An Overview," in *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, IEEE, 2023, pp. 1–6. Accessed: Feb. 19, 2025. [Online]. https://doi.org/10.1109/WINCOM59760.2023.10323032.

[14] M. A. FERRAG, "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications: Centralized and Federated Learning." IEEE, Jan. 17, 2022. https://dx.doi.org/10.21227/mbc1-1h68.

[15] H. zaib, "NSL-KDD", [Online]. https://www.kaggle.com/datasets/hassan06/nslkdd (Accessed Date: February 19, 2025).

[16] N. Moustafa and J. Slay, "The evaluation of Network Anomaly Detection Systems: Statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1–3, pp. 18–31, Apr. 2016, doi: 10.1080/19393555.2015.1125974.

[17] N. Koroniotis, N. Moustafa, E. Sitnikova, and J. Slay, "Towards Developing Network Forensic Mechanism for Botnet Activities in the IoT Based on Machine Learning Techniques," in *Mobile Networks and Management*, vol. 235, J. Hu, I. Khalil, Z. Tari, and S. Wen, Eds., in Lecture Notes of the Institute for Computer Sciences, Social

Informatics and Telecommunications Engineering, vol. 235. , Cham: Springer International Publishing, 2018, pp. 30–44. doi: 10.1007/978-3-319-90775-8_3.

[18] N. Moustafa, M. Ahmed, and S. Ahmed, "Data Analytics-Enabled Intrusion Detection: Evaluations of ToN_IoT Linux Datasets," in *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom)*, Guangzhou, China: IEEE, Dec. 2020, pp. 727–735. doi: 10.1109/TrustCom50675.2020.00100.

[19] Sharafaldin, "IDS 2017 | Datasets | Research | Canadian Institute for Cybersecurity | UNB", [Online]. https://www.unb.ca/cic/datasets/ids-2017.html (Accessed Date: February 19, 2025).

[20] FrancoisXA, "DS2OS traffic traces", [Online]. https://www.kaggle.com/datasets/francoisxa/ds2ostraffictraces (Accessed Date: February 19, 2025)

[21] Zolanvari, "WUSTL-IIOT-2021 Dataset for IIoT Cybersecurity Research" , [Online]. https://www.cse.wustl.edu/~jain/iiot2/index.html (Accessed Date: February 19, 2025).

[22] Neto, "IoT Dataset 2023 | Datasets | Research | Canadian Institute for Cybersecurity | UNB", [Online]. https://www.unb.ca/cic/datasets/iotdataset-2023.html (Accessed Date: February 19, 2025).

[23] N. Khan, K. Ahmad, A. A. Tamimi, M. M. Alani, A. Bermak, and I. Khalil, "Explainable AI-based Intrusion Detection System for Industry 5.0: An Overview of the Literature, associated Challenges, the existing Solutions, and Potential Research Directions," Jul. 21, 2024, *arXiv*: arXiv:2408.03335. doi: 10.48550/arXiv.2408.03335.

[24] C. Park, J. Lee, Y. Kim, J.-G. Park, H. Kim, and D. Hong, "An enhanced AI-based network intrusion detection system using generative adversarial networks," *IEEE Internet of Things Journal*, vol. 10, no. 3, pp. 2330–2345, 2022.

[25] Y. Otoum, "AI-Based Intrusion Detection Systems to Secure Internet of Things (IoT)," PhD Thesis, Université d'Ottawa/University of Ottawa, 2022, [Online]. https://ruor.uottawa.ca/handle/10393/44077 (Accessed Date: February 19, 2025).

[26] S. Hallaq and P. Kulrujiphat, "A Survey of AI-Based Attack Detection Models on the Edge-IIoTset Dataset," in *2024 8th International Conference on Business and Information Management (ICBIM)*, IEEE, 2024, pp. 127–131.

[27] T. Zhukabayeva, L. Zholshiyeva, N. Karabayev, S. Khan, and N. Alnazzawi, "Cybersecurity Solutions for Industrial Internet of Things–Edge Computing Integration: Challenges, Threats, and Future Directions," *Sensors*, vol. 25, no. 1, Art. no. 1, Jan. 2025, doi: 10.3390/s25010213.

[28] S. K. Poorazad, C. Benzaıd, and T. Taleb, "Blockchain and Deep Learning-Based IDS for Securing SDN-Enabled Industrial IoT Environments," Dec. 31, 2023, *arXiv*: arXiv:2401.00468. doi: 10.48550/arXiv.2401.00468.

[29] S. K. R. Mallidi and R. R. Ramisetty, "Advancements in training and deployment strategies for AI-based intrusion detection systems in IoT: a systematic literature review," *Discov Internet Things*, vol. 5, no. 1, p. 8, Jan. 2025, doi: 10.1007/s43926-025-00099-4.

[30] A. Serhane, E.-M. Hamzaoui, and K. Ibrahimi, "IA Applied to IIoT Intrusion Detection: An Overview," in *2023 10th International Conference on Wireless Networks and Mobile Communications (WINCOM)*, Oct. 2023, pp. 1–6. doi: 10.1109/WINCOM59760.2023.10323032.

[31] S. Soliman, W. Oudah, and A. Aljuhani, "Deep learning-based intrusion detection approach for securing industrial Internet of Things," *Alexandria Engineering Journal*, vol. 81, pp. 371–383, Oct. 2023, doi: 10.1016/j.aej.2023.09.023.

[32] Wu, "Network Anomaly Detection With Temporal Convolutional Network and U-Net Model," *ResearchGate*, Dec. 2024, doi: 10.1109/ACCESS.2021.3121998.

[33] Jeffrey, "Using Ensemble Learning for Anomaly Detection in Cyber–Physical Systems." https://doi.org/10.3390/electronics13071391.

[34] Uccello, "Towards Hybrid NIDS: Combining Rule-Based SIEM with AI-Based Intrusion Detectors" in *ResearchGate*, Dec. 2024. doi: 10.1007/978-3-031-56950-0_21.

[35] Khacha, "Hybrid Deep Learning-based Intrusion Detection System for Industrial Internet of Things," in *ResearchGate*, 2022. doi: 10.1109/ISIA55826.2022.9993487.

[36] Sk. T. Mehedi, A. Anwar, Z. Rahman, and K. Ahmed, "Deep Transfer Learning Based Intrusion Detection System for Electric

Vehicular Networks," *Sensors*, vol. 21, no. 14, p. 4736, Jul. 2021, doi: 10.3390/s21144736.

[37] A. Gueriani, H. Kheddar, and A. C. Mazari, "Adaptive Cyber-Attack Detection in IIoT Using Attention-Based LSTM-CNN Models," Jan. 21, 2025, *arXiv*: arXiv:2501.13962. doi: 10.48550/arXiv.2501.13962.

[38] A. Khacha, R. Saadouni, Y. Harbi, and Z. Aliouat, "Hybrid Deep Learning-based Intrusion Detection System for Industrial Internet of Things," in *2022 5th International Symposium on Informatics and its Applications (ISIA)*, 2022, pp. 1–6. doi: 10.1109/ISIA55826.2022.9993487.

[39] Gaber, "Industrial Internet of Things Intrusion Detection Method Using Machine Learning and Optimization Techniques - Gaber - 2023 - Wireless Communications and Mobile Computing - Wiley Online Library." https://doi.org/10.1155/2023/3939895.

[40] O. Almomani, "A Feature Selection Model for Network Intrusion Detection System Based on PSO, GWO, FFA and GA Algorithms," *Symmetry*, vol. 12, no. 6, Art. no. 6, Jun. 2020, doi: 10.3390/sym12061046.

[41] A. Nazir and R. A. Khan, "A novel combinatorial optimization based feature selection method for network intrusion detection," *Computers & Security*, vol. 102, p. 102164, Mar. 2021, doi: 10.1016/j.cose.2020.102164.

[42] Kasongo, "An advanced Intrusion Detection System for IIoT Based on GA and Tree based Algorithms", IEEE Access, Vol. 9, pp. 113199-113212, https://doi.org/10.1109/ACCESS.2021.3104113.

[43] V. Kantharaju, H. Suresh, M. Niranjanamurthy, S. I. Ansarullah, F. Amin, and A. Alabrah, "Machine learning based intrusion detection framework for detecting security attacks in internet of things," *Sci Rep*, vol. 14, no. 1, p. 30275, Dec. 2024, doi: 10.1038/s41598-024-81535-3.

[44] A. Bamidele, C. Chakraborty, and E. Adeniyi, "Intrusion Detection in Industrial Internet of Things Network-Based on Deep Learning Model with Rule-Based Feature Selection," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–17, Sep. 2021, doi: 10.1155/2021/7154587.

[45] R. Lazzarini, H. Tianfield, and V. Charissis, "Federated Learning for IoT Intrusion Detection," *AI*, vol. 4, no. 3, pp. 509–530, Jul. 2023, doi: 10.3390/ai4030028.

[46] J. Huang, Z. Chen, S.-Z. Liu, H. Zhang, and H.-X. Long, "Improved Intrusion Detection Based on Hybrid Deep Learning Models and Federated Learning," *Sensors (Basel)*, vol. 24, no. 12, p. 4002, Jun. 2024, doi: 10.3390/s24124002.

[47] L. T. Rajesh, T. Das, R. M. Shukla, and S. Sengupta, "Give and Take: Federated Transfer Learning for Industrial IoT Network Intrusion Detection." 2023. [Online]. https://arxiv.org/abs/2310.07354

[48] O. Belarbi, T. Spyridopoulos, E. Anthi, I. Mavromatis, P. Carnelli, and A. Khan, "Federated Deep Learning for Intrusion Detection in IoT Networks." 2023. [Online]. https://arxiv.org/abs/2306.02715.

[49] T. Rehman, N. Tariq, F. A. Khan, and S. U. Rehman, "FFL-IDS: A Fog-Enabled Federated Learning-Based Intrusion Detection System to Counter Jamming and Spoofing Attacks for the Industrial Internet of Things," *Sensors*, vol. 25, no. 1, 2025, doi: 10.3390/s25010010.

[50] J. Wang, K. Yang, and M. Li, "NIDS-FGPA: A federated learning network intrusion detection algorithm based on secure aggregation of gradient similarity models," *PLOS ONE*, vol. 19, no. 10, pp. 1–33, Oct. 2024, doi: 10.1371/journal.pone.0308639.

[51] F. Ullah, "A Framework for Anomaly Detection in IoT Networks Using Conditional Generative Adversarial Networks," *ResearchGate*, 2022, doi: 10.1109/ACCESS.2021.3132127.

[52] Y. Wu, H.-N. Dai, and H. Tang, "Graph Neural Networks for Anomaly Detection in Industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9214–9231, Jun. 2022, doi: 10.1109/JIOT.2021.3094295.

[53] S. Balaji, G. Dhanabalan, C. Umarani, and J. Naskath, "A GAN-based Hybrid Deep Learning Approach for Enhancing Intrusion Detection in IoT Networks," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 6, 2024, doi: 10.14569/IJACSA.2024.0150637.

[54] P. V. Dinh, Q. U. Nguyen, D. T. Hoang, D. N. Nguyen, S. P. Bao, and E. Dutkiewicz, "Constrained Twin Variational Auto-Encoder for Intrusion Detection in IoT Systems." 2023. [Online]. https://arxiv.org/abs/2312.02490.

[55] Chen, "Intrusion Detection in Industrial Internet of Things Network-Based on Deep Learning Model with Rule-Based Feature Selection," *ResearchGate*, 2023, doi:

10.1155/2021/7154587.

[56] Z. E. Huma *et al.*, "A Hybrid Deep Random Neural Network for Cyberattack Detection in the Industrial Internet of Things," *IEEE Access*, vol. 9, pp. 55595–55605, 2021, doi: 10.1109/ACCESS.2021.3071766.

[57] Mendonca, "A lightweight intelligent intrusion detection system for industrial internet of things using deep learning algorithms - Mendonça - 2022 - Expert Systems - Wiley Online Library." https://doi.org/10.1111/exsy.12917.

[58] Sinha, "Development of lightweight intrusion model in Industrial Internet of Things using deep learning technique | Discover Applied Sciences." https://doi.org/10.1007/s42452-024-06044-4.

[59] Lansky, "Deep Learning-Based Intrusion Detection Systems: A Systematic Review," *ResearchGate*, Dec. 2024, doi: 10.1109/ACCESS.2021.3097247.

[60] Martinez, "Feature Selection and Ensemble-Based Intrusion Detection System: An Efficient and Comprehensive Approach." https://doi.org/10.3390/sym13101764.

[61] Ebrahimi, F., Javidan, R., Akbari, R. and Hosseini, Y., 2025. Intrusion detection in the internet of things using convolutional neural networks: an explainable AI approach. Cybersecurity, 8(1), p.66.

[62] Y. A. Farrukh, S. Wali, I. Khan, and N. D. Bastian, "Detecting Unknown Attacks in IoT Environments: An Open Set Classifier for Enhanced Network Intrusion Detection." 2023. [Online]. https://arxiv.org/abs/2309.07461.

[63] R. Aljohani, A. Bushnag, and A. Alessa, "AI-Based Intrusion Detection for a Secure Internet of Things (IoT)," *J. Netw. Syst. Manage.*, vol. 32, no. 3, Jun. 2024, doi: 10.1007/s10922-024-09829-5.

[64] yang and wang, "An improved intrusion detection method for IIoT using attention mechanisms, BiGRU, and Inception-CNN | Scientific Reports." https://doi.org/10.1038/s41598-024-70094-2.

[65] B. R. Kikissagbe and M. Adda, "Machine Learning-Based Intrusion Detection Methods in IoT Systems: A Comprehensive Review," *Electronics*, vol. 13, no. 18, Art. no. 18, Jan. 2024, doi: 10.3390/electronics13183601.

[66] G. Karacayılmaz and H. Artuner, "A novel approach detection for IIoT attacks via artificial intelligence," *Cluster Comput*, vol. 27, no. 8, pp. 10467–10485, Nov. 2024, doi: 10.1007/s10586-024-04529-w.

[67] J. A. Beauty Angelin and C. Priyadharsini, "Deep Learning based Network based Intrusion Detection System in Industrial Internet of Things," in *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, Jan. 2024, pp. 426–432. doi: 10.1109/IDCIoT59759.2024.10467510.

[68] A. M. Aleesa, M. Younis, A. A. Mohammed, and N. M. Sahar, "Deep-Intrusion Detection System With Enhanced Unsw-Nb15 Dataset Based On Deep Learning Techniques," vol. 16, 2021.

[69] Ramaiah, "Securing the Industrial IoT: A Novel Network Intrusion Detection Models," in *ResearchGate*, Aug. 2024. doi: 10.1109/AIIoT58432.2024.10574728.

[70] R. Holdbrook, O. Odeyomi, S. Yi, and K. Roy, "Network-Based Intrusion Detection for Industrial and Robotics Systems: A Comprehensive Survey," *Electronics*, vol. 13, no. 22, p. 4440, Nov. 2024, doi: 10.3390/electronics13224440.

[71] H. Nandanwar and R. Katarya, "Deep learning enabled intrusion detection system for Industrial IOT environment," *Expert Syst. Appl.*, vol. 249, no. PC, Sep. 2024, doi: 10.1016/j.eswa.2024.123808.

[72] A. Gueriani, H. Kheddar, and A. C. Mazari, "Deep Reinforcement Learning for Intrusion Detection in IoT: A Survey," in *2023 2nd International Conference on Electronics, Energy and Measurement (IC2EM)*, IEEE, Nov. 2023. doi: 10.1109/ic2em59347.2023.10419560.

[73] M. Injadat, "Optimized Ensemble Model Towards Secured Industrial IoT Devices." 2024. [Online]. https://arxiv.org/abs/2401.05509.

[74] J. Vitorino, I. Praça, and E. Maia, "Towards adversarial realism and robust learning for IoT intrusion detection and classification," Ann. Telecommun., vol. 78, no. 7, pp. 401–412, 2023.

[75] M. Rigaki and S. Garcia, "Bringing a GAN to a knife-fight: Adapting malware communication to avoid detection," in 2018 IEEE Security and Privacy Workshops (SPW), 2018, pp. 70–75, doi: 10.1109/SPW.2018.00019.

[76] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," in Proc. ACM Int. Conf. Inf. Syst. Security and

Privacy (ICISSP), Porto, Portugal, 2017, pp. 1–6. [Online]. Available: https://dl.acm.org/doi/10.1145/3056732.3056737.

[77] Li, C., Li, F., Zhang, L., Yang, A., Hu, Z. and He, M., 2023. Intrusion detection for industrial control systems based on improved contrastive learning SimCLR. Applied Sciences, 13(16), p.9227.

[78] S. Misra, C. Roy, T. Sauter, A. Mukherjee, and J. Maiti, "Industrial Internet of Things for safety management applications: A survey," IEEE Access, vol. 10, pp. 83415–83439, 2022.

[79] F. Reegu, W. Z. Khan, S. M. Daud, Q. Arshad, and N. Armi, "A reliable public safety framework for industrial Internet of Things (IIoT)," in Proc. 2020 Int. Conf. Radar, Antenna, Microwave, Electron., and Telecommun. (ICRAMET), Nov. 2020, pp. 189–193.

[80] F. Aprilia, M. Wahid, and A. Setiawan, "Explainable artificial intelligence for industrial control system intrusion detection," IEEE Access, vol. 10, pp. 3150146–3150159, 2022, doi: 10.1109/ACCESS.2022.3150146.

[81] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," IEEE Trans. Emerg. Topics Comput. Intell., vol. 2, no. 1, pp. 41–50, Feb. 2018, doi: 10.1109/TETCI.2017.2772792.

[82] R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in Proc. IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, 2018, pp. 29–35, doi: 10.1109/SPW.2018.00013.

[83] Patil, S., Varadarajan, V., Mazhar, S.M., Sahibzada, A., Ahmed, N., Sinha, O., Kumar, S., Shaw, K. and Kotecha, K., 2022. Explainable artificial intelligence for intrusion detection system. Electronics, 11(19), p.3079.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.

APPENDIX

Table 1. Datasets: Key Attributes Comparison

Dataset Name	Number of Features	Number of Instances	Number of Attacks
Edge-IIoTset	50	1,000,000	10
NSL-KDD	41	125,973	22
UNSW-NB15	49	2,540,044	10
IoT-Botnet	30	500,000	8
ToN-IoT	115	450,000	13
CICIDS2017	79	14,851,719	15
DS2OS	50	1,000,000	12
WUSTL-IIoT-2021	60	800,000	15
CICIoT2023	85	2,000,000	20

Source: created by the authors

Table 2. Comparative Analysis of IDS Methods for IIoT

Dataset	Methods	Results	Advantages	Limitations
UNSW-NB15	PSO, GWO, FFA, GA	Accuracy = 86.3% (UNSW-NB15; split not stated; results may not transfer across plant/site/time)	Optimized feature selection	High computational complexity
UNSW-NB15	Random Forest	Accuracy = 83.12% (UNSW-NB15; transferability uncertain)	Improved detection accuracy	Requires significant computing
UNSW-NB15	GA + tree-based	Accuracy = 87.61% (UNSW-NB15; transferability uncertain)	High detection rate	Computational overhead
WUSTL-IIoT-2021	ML + optimization	Accuracy = 99.99% (WUSTL-IIoT-2021; results may not transfer across plant/site/time)	High accuracy	Integrating multiple techniques is complex
Edge-IIoTset (2022)	Ensemble learning	+4–7% vs. baselines (Edge-IIoTset; transferability uncertain)	Enhanced robustness	Higher compute
NSL-KDD; UNSW-NB15	Deep residual network	99.21% (NSL-KDD), 86.64% (UNSW-NB15) (splits not stated; transferability uncertain)	High accuracy on NSL-KDD	High compute
UNSW-NB15	Deep learning	Accuracy = 85.42% (split not stated; transferability uncertain)	Improved robustness	Increased complexity
UNSW-NB15; DS2OS	Hybrid neural network	99% (UNSW-NB15), 98% (DS2OS) (transferability uncertain)	High accuracy	Model complexity
CICIDS2017	DL + sparse training	+6.25% vs. baseline (CICIDS2017; transferability uncertain)	Lightweight/efficient	Possible accuracy trade-offs
ToN-IoT	LSTM (DL)	Accuracy = 99.99% (results may not transfer across plant/site/time)	High accuracy	Computationally heavy
DS2OS	DL + Archimedes opt.	F-measure = 98.97% (transferability uncertain)	Scalability/efficiency	Needs more training data
CICIDS2017	RF; NN	67% (RF), 73% (NN) (transferability uncertain)	Simple baselines	Single-dataset limitation
Proprietary/Collected	Neural networks	99.3% (single collected dataset; results may not transfer)	High accuracy	Non-public dataset
Edge-IIoTset (EdgeIIoT-2021)	LSTM	99.93% (Edge-IIoTset; results may not transfer)	Outperforms compared methods	Requires more data

Source: created by the authors

Table 3. Summary of ML and DL IDS

Author(s)	Method Used	Advantages	Limitations
[27]	FL, ML, DL, Blockchain for IDS	Secures IIoT, handles big data, decentralized decision-making	High computation, complexity, and weak privacy measures
[28]	CNN-IDS + Blockchain	High-accuracy, secure, tamper-proof logs	Memory-intensive, real-time inefficiency
[29]	AI-powered IDS (ML, DL)	Adapts to threats, scalable in cloud & fog	Imbalanced data, real-time adaptation issues
[30]	AI-based IDS	Detects advanced threats in real-time	High false positives, scalability issues
[31]	DL-based IDS	High precision, handles diverse IIoT devices	Computationally heavy, real-time challenges
[32]	Pelican (Deep Residual Network)	High accuracy, detects complex attacks	Resource-intensive, unsuitable for low-power devices

Source: created by the authors

Table 4. Summary of Ensemble Learning Approaches IDS

Author(s)	Method Used	Advantages	Limitations
[33]	Ensemble ML (SVM, RF, DT) on Edge-IIoTset2023	High accuracy, reduced false positives & negatives	High computational demand limits real-time performance
[34]	Hybrid NIDS (RF, Neural Network) on CICIDS2017	Improved intrusion detection, detects subtle attacks	Requires high-quality labeled data, high computational power
[35]	CNN + Reinforcement Learning on CICIoT2023	Adaptive to dynamic threats, multi-step learning	High computational complexity, not ideal for low-latency applications
[36]	DL + Rule-Based IDS on CICIoT2023	High detection accuracy, reduced false positives	Pre-defined rules limit unknown attack detection, complex deployment
[37]	LSTM-CNN-Attention on IIoT dataset	High accuracy, effective cyberattack classification	Computationally complex, less effective in real-time deployment
[38]	CNN-LSTM IDS on Edge-IIoTset	Better feature extraction generalizes well to attack types	Needs optimization for real-time and adaptive learning

Source: created by the authors

Table 6. Comparative Analysis of Federated Learning-Based IDS Approaches in IIoT Security

Author(s)	Key Contributions	Limitations	Similarities to Other Studies	Differences from Other Studies
[45]	FL-IDS with distributed learning	High accuracy preserves data privacy and is scalable across IIoT devices.	Limited model integration due to data heterogeneity, high computational and communication overhead	Shares privacy-preserving benefits with Rajesh et al. and Rehman et al.
[46]	FL-IDS with CNN, attention mechanisms, and VAEs	Improves network security, cost-effective threat detection, privacy preservation	Computationally expensive, high communication overhead limiting scalability	Similar to Rajesh et al. in using CNN-based models
[47]	FTL-based IDS with CNN	Ensures data privacy, high detection accuracy, and effectiveness in real-time cybersecurity	Communication overhead, model synchronization challenges, high resource consumption	Uses CNN-based IDS like Huang et al.
[48]	FL-based IDS using the TON-IoT dataset	Collaborative learning addresses scalability and privacy concerns	Performance affected by data heterogeneity requires improved aggregation techniques	Similar privacy-preserving focus as Lee et al. and Rehman et al.
[49]	FFL-IDS with CNN and fog computing	Low-latency detection, privacy preservation, high accuracy	Computational limitations in fog nodes, latency in data aggregation	Shares fog computing approach with Wang et al.
[50]	NIDS-FGPA with federated learning and homomorphic encryption	High security, effective in detecting sophisticated attacks, reduces unnecessary communication overhead	High computational and communication costs, encryption-related model sluggishness	Similar to Rehman et al. in reducing unnecessary communication

Source: created by the authors

Table 7. Summary of weight and resource-efficient models in IDS

Author	Method Used	Advantages	Limitations
[57]	Lightweight deep learning-based IDS	High efficiency, real-time deployment on resource-constrained IIoT devices	The simplified model may reduce the accuracy of complex attacks
[58]	Deep neural networks with adaptive methods	High detection accuracy, adaptable to diverse IIoT devices	High computational load, large memory requirement
[59]	CNN-SVM-based IDS	Low computational overhead, efficient for IIoT with limited resources	With limited flexibility for unknown attacks, CNN may miss sophisticated attack behaviors
[60]	MLP with feature selection	Optimized detection accuracy, reduced computational complexity	May lose important features, less effective for unseen attacks
[61]	CNN-based IDS (ResNet, EfficientNet)	Strong feature extraction, improved TPR and FPR	Computationally intensive, struggles with dynamic attack patterns
[62]	Image-based embeddings with sub-clustering	High detection rate for unknown attacks, strong adaptability	Needs further testing on diverse attack types and IoT environments

Source: created by the authors

Table 8. Comparative Analysis of ML-Based IDS Approaches for IIoT Security

Authors	Methods	Dataset Used	Key Findings	Limitations
[68]	CNN-based deep learning model for IDS	Extended UNSW-NB15	Improved attack detection accuracy, including unknown attacks	High computational cost in real-time applications
[69]	LSTM and ML-based ERT-NIDS for IIoT security	EdgeIIoT-2021	High detection rates with real-time threat prediction	Computational overhead is not suitable for resource-limited environments
[70]	NIDS for industrial and robotic systems using ML, DL, and hybrid approaches	Various industrial datasets	Identified promising techniques like federated learning and blockchain for IDS	Dataset scarcity, high false positive rates, and integration issues
[71]	CNN-GRU-based AttackNet for botnet detection	N_BaIoT	Classifies botnet attacks into ten categories with high accuracy	High computational intensity may not generalize to unknown networks
[72]	DRL-based IDS for IoT security	Various IDS datasets	Enhanced detection rates with reduced false alarms	High training time, poor generalizability to new attack types
[73]	BO-GP optimized ensemble tree-based IDS	IIoT network datasets	Improved detection performance through hyperparameter tuning	High computational complexity, limited scalability in real-time IIoT settings
[63]	ML-IDS with a two-layer classification structure	ToN-IoT	Improved accuracy and efficacy in detecting IoT network anomalies; better differentiation of attack types	Struggles with imbalanced datasets and scalability in larger networks
[64]	Inception-CNN and BiGRU with hybrid sampling and denoising	Edge-IIoTset, CIC-IDS2017, CIC-IoT2023	Enhanced intrusion detection through sequential and spatial feature extraction	High computational overhead limits real-time applicability in IIoT environments
[65]	Supervised, unsupervised, and hybrid ML approaches for IDS	Extended UNSW-NB15	Highlights deep learning's adaptability for detecting emerging security threats	IoT devices have limited computational resources, requiring optimized models.
[66]	Expert system using anomaly detection and RL for IIoT security	Modbus, MQTT protocols on a PLC testbed	High accuracy, low latency, and low overhead in detecting cyber threats	Fixed rule-based reasoning limits adaptability to new attack vectors

Source: created by the authors

Table 9. Comparative Summary of AI-Based Intrusion Detection Systems (IDS) in IIoT and ICS Security

Ref.	Method	Evaluation Protocol	Result Source	Deployment Context
[6]	Multi-layer CNN–LSTM hybrid for cyber-threat analysis	UNSW-NB15 and NSL-KDD datasets; Acc = 99.12 %, F1 = 98.85 % (80/20 split)	Reported (original experiments)	Cloud-centric threat analytics and anomaly detection in IIoT networks
[10]	CNN–BiLSTM ensemble model for attack detection	Edge-IIoTset dataset; improved latency by 18 % (70/30 split)	Reported (original)	Edge-level AI-based IDS in IIoT environments
[24]	GAN–Autoencoder hybrids (G-DNNAE, G-CNNAE, G-LSTM)	NSL-KDD, UNSW-NB15, IoT-23 datasets; train–test split (70/30 for all reported metrics)	Reported (original data)	Gateway/edge-level IDS for IIoT and enterprise networks
[31]	Federated CNN–LSTM model for distributed IDS	BoT-IoT dataset; Acc = 98.7 % (10 federated rounds), communication cost reduced by 23 %	Reported (original)	Privacy-preserving distributed IDS for IIoT cloud–edge integration
[52]	Semi-supervised Graph Neural Network (GNN) for anomaly detection	ToN-IoT and UNSW-NB15 datasets; 10-fold cross-validation (10-fold CV for metrics)	Reported	Edge-layer IDS with relational learning for IIoT systems
[56]	Hybrid Deep Random Neural Network (HDRaNN) integrating RaNN and MLP	DS2OS and UNSW-NB15 datasets; Accuracy = 98–99 % (75/25 split + 10-fold validation)	Reported (original)	Lightweight deployable IDS for IIoT edge devices (ARM/Raspberry Pi)
[61]	Hybrid 1D-CNN with SHAP (global) + LIME (local) interpretability	ToN-IoT dataset; 60/20/20 split with SMOTE balancing and 10× run averaging	Reported	Explainable, resource-efficient IDS for IoT edge devices
[77]	Improved SimCLR contrastive learning with dual ResNet50 encoders	SWaT and MSU datasets; Acc = 98.9 %, F1 = 97.3 %(10 % labeled fine-tuning with SMOTE + PCA normalization)	Reported (original)	Self-supervised contrastive learning–based IDS for ICS anomaly detection

Source: created by the authors

Table 10. Deployment Suitability and Resource Latency Characteristics of Major AI-Based IDS Method Families in IIoT and ICS

Method Family	Representative Studies	Recommended Deployment Tier	Typical Latency / Resource Range (from literature)	Key Characteristics and Suitability
CNN-based Models	[6], [61], [83]	Cloud / Gateway	100–300 ms; requires GPU or multi-core CPU for inference	Strong spatial-feature extraction; well suited for centralized analytics and deep packet inspection.
RNN / LSTM-based Models	[6], [10], [31]	Edge / Cloud-integrated	10–100 ms; moderate memory; higher energy demand	Captures sequential flow dynamics; effective for real-time traffic analysis and adaptive IIoT detection.
GNN-based Models	[52]	Fog / Edge	50–150 ms; moderate compute and RAM	Learns relational device–topology patterns; scalable for distributed IIoT telemetry.
Hybrid / Ensemble Models (e.g., CNN–BiLSTM, GAN–AE, HDRaNN)	[10], [24], [56]	Edge / Gateway	20–120 ms; moderate computational footprint	Combines complementary learners; balances accuracy and latency; moderate implementation complexity.
Federated Learning-based Models	[31]	Cloud–Edge Collaborative	100–500 ms per aggregation round; distributed compute	Enables privacy-preserving training; reduces bandwidth needs; suitable for multi-site IIoT infrastructures.
Adversarial Defense / Robust Learning Models	[24], [77]	Gateway / Cloud	50–200 ms; high compute cost during training	Improves robustness against adversarial perturbations; integrates with GANs or contrastive learning.
Self-Supervised / Contrastive Learning Models	[77]	Edge / Cloud-assisted	< 50 ms inference after training; moderate GPU for pretraining	Leverages unlabeled data; allows lightweight fine-tuning; supports evolving IIoT behavioral patterns.
Explainable AI (XAI) Models	[61], [83]	Edge / Cloud	≤ 100 ms; lightweight interpretable models	Provides transparency via SHAP or LIME; supports trust, auditing, and compliance in industrial environments.
Lightweight Random / Probabilistic Neural Networks	[56]	Edge	< 20 ms; low power (ARM/Raspberry Pi class)	Compact and energy-efficient; ideal for constrained embedded IIoT devices.

Source: created by the authors