

Systematic Review

A Systematic Literature Review of Intrusion Detection and Prevention Frameworks for Industrial Communication Protocols Using ML and DL

Khawla Al-Tarawneh ^{1,2,*}, Ahmad Sharieh ^{1,3}  and Sherenaz Al-Haj Baddar ¹

¹ Department of Computer Science, King Abdullah II School of Information Technology, The University of Jordan, Amman 11942, Jordan; sharieh@ju.edu.jo or a.sharieh@aum.edu.jo (A.S.); s.baddar@ju.edu.jo (S.A.-H.B.)

² Department of Computer Science, Faculty of Information Technology, Mutah University, Karak 61710, Jordan

³ Department of Intelligent Robotics System Engineering, Faculty of Engineering, American University of Madaba, Madaba 11821, Jordan

* Correspondence: khawla_t@mutah.edu.jo or kol9220471@ju.edu.jo

Abstract

This systematic literature review examines 31 peer-reviewed articles released in 2021–2025. It offers a coherent summary of intrusion detection and prevention systems based on machine learning and deep learning of industrial communication protocols. The review categorizes the studies depending on research focus, experimental setup, datasets, and analytical methods. According to the quantitative analysis results, the most suitable model for use in this case is the hybrid deep learning architecture, which includes the combination of Transformer-LSTM models and MODLSTM models, with 29% of the reviewed studies using these models and achieving detection rates of over 99%. Federated learning was mentioned in about 9.7% of the studies, and for 67% of them, real-world data was not available, indicating a lack of access to real-world data. These models are prevalently implemented to identify Denial-of-Service, Man-in-the-Middle, and data injection attacks. The results show that Modbus/TCP is the most studied protocol, which indicates how common it is in industrial systems. Meanwhile, other more recent protocols like MQTT and OPC UA are gaining momentum. Another insight revealed by this review is the tendency towards the use of more realistic validation techniques. Hardware-in-the-loop simulations and physical testbeds are in use in many studies. Integrated solutions which comprise a combination of edge, fog, and cloud computing are gaining popularity. Federated learning (utilized in 6.45% of the selected corpus) and software-defined networking are two emerging directions. Although these developments have taken place, there are still critical gaps, including the scarcity of real-world datasets combined with a lack of robust approaches to address scalability and privacy complications. Furthermore, recent IIoT protocols have not been thoroughly evaluated. The study highlights the need for adaptive and lightweight frameworks and the importance of implementing mechanisms that ensure privacy. There is also a need to have standardized evaluation criteria. These factors combined are instrumental for creating secure, resilient, and interoperable industrial networks during the Industry 4.0 period.



Academic Editor: Ricardo Augusto Rabelo Oliveira

Received: 21 May 2026

Revised: 13 June 2026

Accepted: 15 June 2026

Published: 1 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons](#)

[Attribution \(CC BY\)](#) license.

Keywords: industrial control systems; industrial internet of things; machine learning and deep learning; intrusion detection and prevention; Modbus/TCP; industrial communication protocols

1. Introduction

The digital transformation of industrial operations has introduced unprecedented efficiency, connectivity, and automation across critical infrastructures such as energy, water, manufacturing, and transportation. At the heart of this transformation lie industrial control systems (ICSs) and Supervisory Control and Data Acquisition (SCADA) networks, which coordinate physical and cyber processes in real time. However, this growing interconnection between operational technology (OT) and information technology (IT) has simultaneously exposed these systems to escalating cybersecurity risks [1]. Many industrial environments still depend on legacy protocols like Modbus/TCP, DNP3, and PROFINET. These protocols were never designed with security features such as authentication or encryption, leaving them highly susceptible to modern cyber threats [2,3].

Recent years have witnessed a surge in cyberattacks targeting industrial communication networks, including Denial-of-Service (DoS), Man-in-the-Middle (MITM), and false data injection (FDI) attacks that disrupt or manipulate control processes [4,5]. These threats have exposed the limitations of traditional rule-based intrusion detection systems, which struggle to handle high-dimensional, time-dependent industrial data. Consequently, researchers have increasingly turned to machine learning (ML) and deep learning (DL) techniques as powerful tools to detect, prevent, and mitigate these evolving threats. Such data-driven models can identify complex attack patterns, learn from historical behaviors, and adapt to new threats in near real time [6,7].

Numerous ML/DL models have demonstrated remarkable promise. For instance, LSTM and CNNs have achieved high accuracy in identifying spoofing and DoS attacks on Modbus-based systems [8,9]. Meanwhile hybrid models combining Transformers and LSTMs have reached detection rates as high as 99.8% in air-gapped nuclear control systems [10]. Similarly, integrating Elliptic Curve Cryptography (ECC) with Modbus TCP communications proved effective in securing resource-constrained SCADA devices [11]. Yet, these advancements are not without challenges, most notably, the reliance on synthetic datasets, small-scale simulations, and a narrow focus on specific protocols, all of which limit generalizability and real-world application [12,13].

Federated learning is recognized as one of the emerging techniques, with 3 out of 31 studies ($\approx 9.7\%$) using this technique, namely in the distributed energy resource networks in [14] and in the IIoT environment in [15]. The frameworks represent a shift from reactive to intelligent, proactive security systems and a move towards operating environments that maintain operational resilience in complex industrial settings.

In addition to federated learning, digital twins, blockchain-based security infrastructures, and new developments in large model agents offer novel opportunities for performing autonomous security operations. The idea of large language model (LLM)-powered agents has gained increased attention in recent years in order to create an intelligent agent that is able to perceive, reason, plan, and engage in collective decision-making. Large Model-Based Agents, Embodied Intelligence, and Internet of Agents are new concepts that build on the principles of AI and allow several intelligent agents to coordinate and exchange contextual data to respond to the dynamic operating environment. Even though these approaches have not been extensively applied in the area of industry-based intrusion detection and prevention, they offer interesting perspectives for further exploration in autonomous cybersecurity operations [16–18].

In addition to these advances, ICS security and new Industry 4.0 paradigms are creating a new threat and defense landscape. It has been suggested that federated learning in 6G-enabled IIoT environments is a key enabler in distributed, privacy-preserving intelligence in the next generation of industrial networks, to enable real-time anomaly detection with minimal data centralization [19]. At the same time, digital twin technology

is emerging as a powerful framework for intrusion detection in SCADA, delivering high-fidelity cyber–physical replication environments that enable operational-aware anomaly detection which network-only IDSs fail to provide [20].

In addition, blockchain-based systems can provide trustworthy, tamper-proof audit trails and distributed IDS architectures that mitigate the trust and scalability issues of centralized IDSs in industrial IoT networks [21]. These cross-cutting paradigms increase the attack surface and the security toolkit as well, and therefore must be taken into consideration in any extensive study on ICS security in Industry 4.0.

Given this rapidly evolving research landscape, the present study conducts a systematic literature review (SLR) to consolidate and critically examine the most recent advancements in intrusion detection and prevention frameworks for industrial communication protocols utilizing ML and DL. The review covers studies published between 2021 and 2025, aiming to capture the latest scientific contributions, research trends, and practical developments in this domain. Accordingly, this study seeks to address the following research questions (RQs):

- RQ1: What ML/DL techniques are used to detect or prevent intrusions in industrial communication protocols?
- RQ2: Which industrial protocols and datasets are most studied?
- RQ3: What are the typical architectures of intrusion detection and prevention frameworks/filters (IDPFs) in terms of feature extraction, placement, latency, and response mechanisms?
- RQ4: How are methods evaluated in terms of metrics, testbeds, and real-world deployments?
- RQ5: What are the key open challenges and future research directions?

By addressing these questions, this study provides a structured and comprehensive overview of the state of the art in AI-driven intrusion detection and prevention for industrial networks. It not only highlights the dominant methodologies and experimental designs found in the recent literature but also identifies critical gaps that need to be bridged to ensure the secure, scalable, and resilient operation of industrial control systems in the era of Industry 4.0 [22,23].

Compared to general reviews on IoT anomalies (e.g., [24]), which cover high-level network layouts and classical algorithms, our review provides three distinct scientific contributions: (1) it focuses strictly on native operational technology (OT) protocol frames (Modbus, S7Comm, IEC 60870 [25]); (2) it maps out real physical deployments across hardware-in-the-loop (HIL) simulators, physical testbeds, and edge–fog topologies; (3) it explicitly benchmarks deep learning models capable of handling non-linear time-series industrial telemetry.

The choice of the 2021–2025 publication window was strategically selected to isolate and synthesize modern AI advancements, following a massive paradigm shift away from simple classical machine learning toward advanced deep learning (e.g., Transformers, hybrid temporal networks) and advanced distributed deployment models like federated learning. While foundational intrusion detection principles were established prior to 2021, this review specifically captures contemporary architectures designed to handle the high throughput and telemetry demands of modern Industry 4.0 environments.

Figure 1 illustrates the high-level architecture of the systematic review. The figure summarizes the main stages of the review process, the key analytical dimensions examined throughout the study, and the subsequent synthesis of findings that lead to the identification of research challenges, future directions, and overall conclusions. This visual representation serves as a roadmap for understanding how the different components of the review collectively address the research objectives and questions.

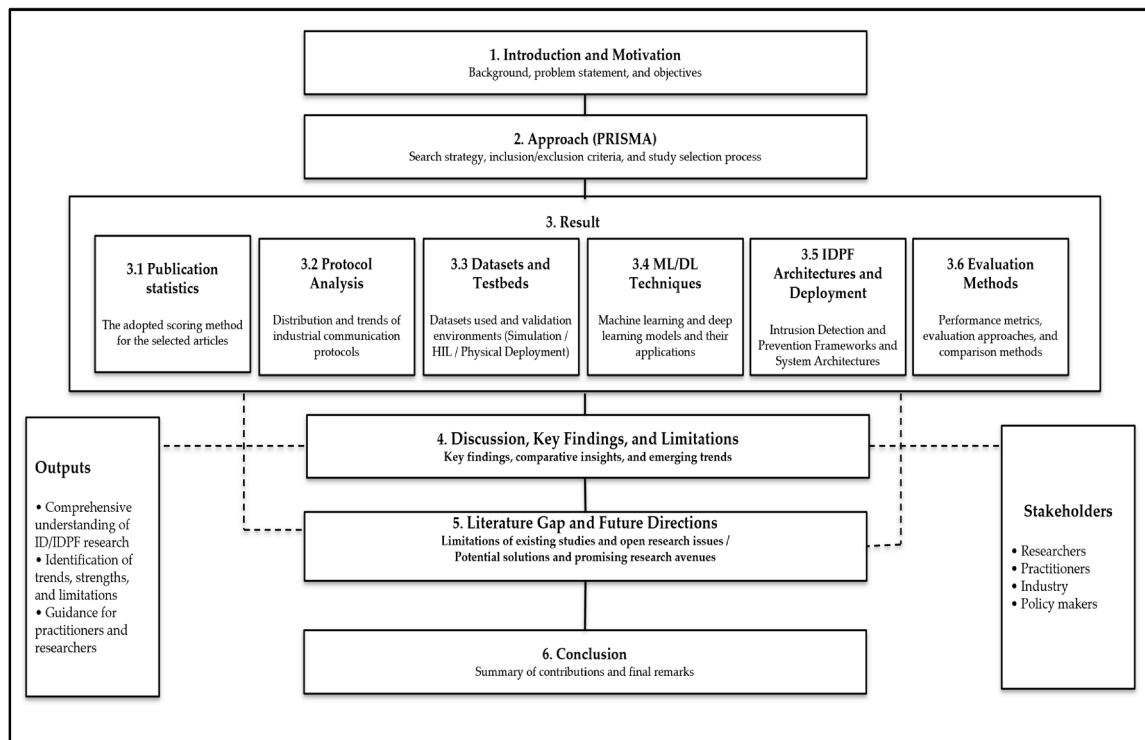


Figure 1. Survey structure.

2. Approach

Here we outline how we used a thorough review of the literature on intrusion detection and prevention in industrial control systems using machine and deep learning. This was aimed at streamlining the process to be systematic, transparent and reproducible to ensure the findings are reliable as well as meaningful. Our search strategy consisted of designing a search to look into several popular academic databases, including studies published within a selected timeframe. To capture the main areas of interest, the keywords and search terms were chosen very carefully to include industrial cybersecurity, relevant control system protocols, and other machine learning methods. After collecting the first round of the studies, we used some criteria to select or filter out papers, and only the most relevant and high-quality research was considered. These papers were then analyzed and synthesized to identify patterns, challenges, and gaps in intrusion detection in industrial settings.

2.1. Search Strategy

We conducted a structured search across multiple reputable academic databases. The selected databases included IEEE Xplore, ACM Digital Library, Scopus, Web of Science, SpringerLink, and arXiv. These cover a wide spectrum of research in industrial cybersecurity and machine learning (ML) applications. We developed Boolean search queries combining key terms for intrusion detection (e.g., “intrusion detection,” intrusion detection system (IDS), anomaly detection), industrial environments (e.g., ICS, SCADA, OT), industrial protocols (e.g., Modbus, Distributed Network Protocol version 3 (DNP3), Process Field Net (PROFINET)), and machine learning methods (e.g., neural networks (NNs), support vector machine (SVM), random forest (RF), long short-term memory (LSTM), convolutional neural network (CNN), and auto encoders (AEs)). Each database required slight adjustments to accommodate its specific search syntax and field options. Table 1 summarizes the exact search strings used for each database, along with practical notes on syntax adjustments and search fields.

Table 1. Database search strings and execution parameters.

Database	Query/Search String	Notes/Adjustments
SpringerLink	("Intrusion detection" OR "intrusion prevention" OR IDS OR IDPS OR "anomaly detection") AND (industrial OR ICS OR SCADA OR OT) AND ("Modbus" OR DNP3 OR PROFINET OR "Ethernet/IP" OR "IEC 61850 [26]" OR "industrial protocol") AND ("machine learning" OR "deep learning" OR "neural network" OR SVM OR "random forest" OR LSTM OR CNN OR autoencoder)	Use Advanced Search, then select "All Fields" or "Abstract & Keywords"
IEEE Xplore	((("intrusion detection" OR "intrusion prevention" OR IDS OR IDPS OR "anomaly detection") AND (industrial OR ICS OR SCADA OR OT) AND ("Modbus" OR DNP3 OR PROFINET OR "Ethernet/IP" OR "IEC 61850" OR "industrial protocol") AND ("machine learning" OR "deep learning" OR "neural network" OR SVM OR "random forest" OR LSTM OR CNN OR autoencoder)))	Use Advanced Search → select "All Metadata" or "Abstract" fields.
ACM DL	"Intrusion detection" OR "intrusion prevention" OR IDS OR IDPS OR "anomaly detection" AND industrial OR ICS OR SCADA OR OT AND "Modbus" OR DNP3 OR PROFINET OR "Ethernet/IP" OR "IEC 61850" OR "industrial protocol" AND "machine learning" OR "deep learning" OR "neural network" OR SVM OR "random forest" OR LSTM OR CNN OR autoencoder	Use "Advanced Search" → set fields to "All Metadata" or "Abstract/Title." ACM does not always support nested parentheses perfectly.
Scopus	TITLE, ABS, KEY(("intrusion detection" OR "intrusion prevention" OR IDS OR IDPS OR "anomaly detection") AND (industrial OR ICS OR SCADA OR OT) AND ("Modbus" OR DNP3 OR PROFINET OR "Ethernet/IP" OR "IEC 61850" OR "industrial protocol") AND ("machine learning" OR "deep learning" OR "neural network" OR SVM OR "random forest" OR LSTM OR CNN OR autoencoder))	Use TITLE, ABS, KEY to search in title, abstract, and keywords. Scopus supports nested Boolean operators.
Web of Science	TS = (("intrusion detection" OR "intrusion prevention" OR IDS OR IDPS OR "anomaly detection") AND (industrial OR ICS OR SCADA OR OT) AND ("Modbus" OR DNP3 OR PROFINET OR "Ethernet/IP" OR "IEC 61850" OR "industrial protocol") AND ("machine learning" OR "deep learning" OR "neural network" OR SVM OR "random forest" OR LSTM OR CNN OR autoencoder))	TS = Topic (title, abstract, author keywords). It supports nested Boolean logic.
arXiv	all: "intrusion detection" OR all: "intrusion prevention" OR all: IDS OR all: IDPS OR all: "anomaly detection" AND all: industrial OR all: ICS OR all: SCADA OR all: OT AND all: "Modbus" OR all: DNP3 OR all: PROFINET OR all: "Ethernet/IP" OR all: "IEC 61850" OR all: "industrial protocol" AND all: "machine learning" OR all: "deep learning" OR all: "neural network" OR all: SVM OR all: "random forest" OR all: LSTM OR all: CNN OR all: autoencoder.	arXiv search uses all: prefix for fields. Boolean operators must be capitalized. Large queries may need splitting for better accuracy.

The study focused on key IEEE journals such as *IEEE Transactions on Industrial Informatics*, *IEEE Internet of Things Journal*, *IEEE Transactions on Information Forensics and Security*, and *IEEE Access*, which are all important for publishing high-quality papers on security and intrusion detection in ICSs using ML.

2.2. Inclusion and Exclusion Criteria

This study reviews the most recent literature about industrial control systems (ICSs) and SCADA networks, with emphasis on Modbus-based protocols (TCP, RS-485). It considers publications in 2021–2025 to capture innovative solutions through a critical review addressing Denial-of-Service (DoS) or Man-in-the-Middle (MITM) attacks, including subtypes (e.g., field flooding, ARP spoofing, false data injection). The inclusion criteria include:

1. Empirical Validation

Studies must implement and evaluate proposed solutions in realistic operational environments to ensure applicability beyond theoretical models. Validation settings include:

- Physical and hardware testbeds (e.g., PLCs, microgrids, nuclear plant replicas).
- Cyber-physical co-simulations (e.g., RTDS integrated with NS-3).
- Virtual environments (e.g., Node-RED, Conpot).
- Real-world datasets (e.g., ICS honeynets, industrial network traffic captures).

Rationale: This ensures practical feasibility and realistic performance assessment of the proposed security measures.

It should be noted that for this criterion, software-only simulation environments (such as evaluation frameworks in Python 3.10 or virtual networks in GNS3 2.2.x) are accepted as valid empirical settings with measurable, reproducible results reported based on defined datasets and performance metrics. This position is informed by the reality of physical ICS testbeds which are expensive, scarce and difficult to access for researchers; software simulation with realistic industrial traffic and/or protocol emulation is a valid empirical contribution. Works based purely on theoretical analysis, mathematical proofs or protocol assessment (descriptive) were excluded.

2. Quantifiable Outcomes

Studies must report measurable and reproducible performance indicators to enable objective evaluation. Reported metrics include:

- Detection performance metrics (e.g., accuracy, precision, recall, F1-score).
- Computational efficiency metrics (e.g., latency, memory consumption, processing overhead).
- Vulnerability assessment metrics (e.g., CVSS scores, exploitability measures).

Rationale: This enables systematic cross-study benchmarking and transparent performance comparison.

3. Solution-Centric Approach

Studies must focus on implementable detection, mitigation, or prevention mechanisms rather than descriptive vulnerability analysis alone. Proposed solutions typically involve:

- Machine learning- and AI-based models (e.g., LSTM, Transformers, federated learning).
- Cryptographic enhancements (e.g., ECC, lightweight encryption schemes).
- Protocol hardening mechanisms (e.g., stateful monitoring, access control).

Rationale: This prioritizes actionable and deployable defensive strategies over purely theoretical or descriptive research.

References [1,25], although published before 2021, are cited as foundational background works that motivated this review rather than as part of the analyzed study corpus.

Studies were excluded if they were purely analytical, or theoretical, as this review does not focus on purely abstract or theoretical frameworks. We recognize, however, that theory can provide complementary insights into issues of protocol vulnerabilities, and therefore mention relevant theory in the Discussion where applicable.

The selected timeframe (2021–2025) is a more recent period to showcase the latest developments in intrusion detection for industrial protocols using ML/DL methods, because of the substantial advancements in federated learning, transformer-based models, and threats specific to IIoT. Although there are some foundation papers published prior to 2021, they are made available as background information and not as the main analyzed papers.

2.3. Selection Process

In this study, the three stages depicted in Figure 2 were employed to select the publications included in developing this systematic literature review (SLR) process.

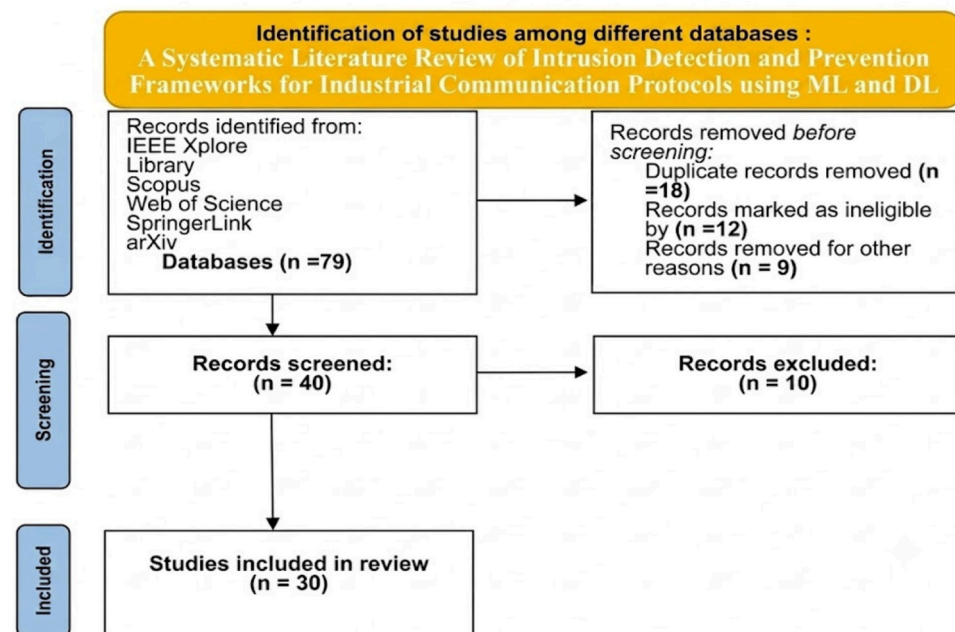


Figure 2. Publication selection process.

A total of 18 records were removed because they were duplicated across different databases, meaning they did not add new value to the review process. Twelve records were excluded after screening the titles and abstracts because they did not meet the inclusion criteria, either because they did not focus on intrusion detection and prevention frameworks in industrial communications or because they relied on other, not directly relevant research areas. Additionally, nine records were removed for various reasons, such as a lack of underlying data, unavailability of the full text, and poor content quality that prevented reliable scientific analysis. Following this selection process, 31 publications that fitted the inclusion criteria were included in this review. The next subsection justifies their inclusion.

The selection process was done according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines to guarantee rigor and reproducibility. The titles and abstracts were screened independently by 2 reviewers with conflicts resolved by discussion. To ensure consistency of interpretation, a pilot test of the inclusion/exclusion criteria was carried out on a random sample of 10 papers prior to full application.

2.4. Information Extraction

This section highlights selected studies from the systematic review. Table 2 presents the first and final studies, outlining their objectives, methodologies, and experimental settings.

It should be noted that references [1,2,9,22–31] serve as complementary background references and are not part of the 31 primary studies analyzed in this review. The 31 analyzed studies correspond to references [2–15,27,32–48], as detailed in Table 2.

Table 2 presents the studies included in this review and visualizes the extracted data to illustrate the bibliometric patterns aligned with the research objectives. The reviewed works emphasize the growing cyber threats targeting industrial control systems, particularly those relying on legacy protocols such as Modbus/TCP, with Denial-of-Service and Man-in-the-Middle attacks shown to cause severe operational disruptions, including PLC crashes and power grid destabilization.

Table 2. Reviewed research area.

Reference	Purpose of Study	Methodology/ Approach	Findings/ Results	Simulator/ Testbed Used	Protocol Studied	Dataset	Testbed/ Simulation
[6]	Omni intrusion detection for SCADA networks	Deep learning ensemble of FNNs and LSTM	Ensemble improved detection of temporally correlated and uncorrelated attacks	SCADA network sim	SCADA	Power system disturbance data (custom simulated)	SCADA network simulation (OPNET, OMNeT++)
[2]	Investigate Modbus protocol vulnerabilities to Denial-of-Service (DoS) attacks and propose mitigation strategies	Virtual testbed simulation of DoS attacks on Modbus; traffic analysis and rule-based mitigation	Successfully demonstrated DoS attack impacts on Modbus; proposed effective mitigation using virtual testbed	Virtual Modbus testbed (software-defined)	Modbus/ TCP	Simulated DoS attack traffic (custom dataset)	Virtual software-defined testbed
[32]	Develop an AI-based IDS to detect cyberattacks in critical infrastructure ICS networks using the Modbus protocol	ML + DL + Ensemble methods combined; custom Wireshark post-dissector developed, feature selection applied; implemented in Python	High detection accuracy with reduced feature set; Ensemble methods outperformed standalone models; surpassed prior literature benchmarks	Python-based evaluation environment; custom Wireshark post-dissector	Modbus	Hybrid dataset (normal + attack Modbus traffic)	Software-based (Python), no dedicated hardware, SCADA testbed
[8]	Detect smart meter Modbus RS-485 spoofing attacks	LSTM-based IDS, dataset created for RS-485 spoofing	Effective detection of spoofing attacks using deep learning	Simulated RS-485 network	Modbus RS-485	Dataset created for RS-485 spoofing attack (unnamed)	Simulated RS-485 smart meter network
[33]	Evaluate automotive IDS techniques for ICS (Modbus/TCP)	Multi-modal IDS using timing, address probing, and data fuzzing	100% detection, <1% false positive on simulated data	Pre-recorded Modbus data + test network	Modbus/ TCP	Pre-recorded Modbus data (e.g., “Modbus Dataset” from UNB/ISCX)	Test network (Southwest Research Institute)
[3]	Recognize DoS attacks on Modbus/TCP using neural networks	MODLSTM neural network (preprocessing, feature recording, classification)	Improved accuracy by 0.74%, fine-grained DoS detection with high precision	Industrial control network traffic	Modbus/ TCP	Industrial control network traffic (e.g., HIL-based Augmented ICS)	Hardware-in-the-loop (HIL) testbed
[12]	Automated detection-in-depth method for ICS flooding	Data acquisition (PLC logs, network), NetFlow preprocessing, unsupervised ML	High-precision detection of flooding attacks outperforms single-source methods	Factory automation dataset	Modbus	Factory automation dataset (unnamed), SWaT dataset	iTrust SWaT testbed (physical testbed)
[34]	Analyze off-the-shelf cyber threats and propose simple protection methods	Penetration testing & experimental validation with PLC/HMI	Demonstrated MITM and other attacks, proposed PLC-based integrity checks and signal validity monitoring	Real compressor station ICS	PLC/HMI proprietary	Operational data from real compressor station (proprietary)	Real compressor station ICS setup
[35]	Enhance Modbus/TCP security protocol with authentication	Node-RED-based implementation of access control with MITM component	Authentication and authorization methods are effective in preventing unauthorized access	WWTP physical/hybrid testbed	Modbus/ TCP	Industrial process data (WWTP cycle metrics)	Physical testbed with Moxa devices

Table 2. Cont.

Reference	Purpose of Study	Methodology/ Approach	Findings/ Results	Simulator/ Testbed Used	Protocol Studied	Dataset	Testbed/ Simulation
[4]	Detect and localize false data injection attacks in power grids	ML-based anomaly detection integrating network & grid measurement data	Successful detection and localization of MITM-based false data injection attacks	HIL real-time simulation	Power grid protocols	Power system models (e.g., IEEE bus models)	HIL real-time simulation testbed (RTDS)
[36]	Implement stateful monitoring in NIDS for industrial protocols	UML state chart-based protocol modeling integrated with NIDS, tested on Modbus TCP and IEC 60870-5-104	Stateful monitoring enhanced attack detection capabilities, including replay and privileged requests	Integrated with existing NIDS	Modbus/ TCP & IEC 60870	Public ICS datasets (NGIDS, DS or CICIDS2017)	Integrated with NIDS (software testbed)
[11]	Integrate ECC with Modbus TCP SCADA communication	Experimental setup implementing ECC-based encryption and authentication	ECC is suitable for resource-constrained SCADA devices, it improves security with low overhead	Experimental SCADA test setup	Modbus/ TCP (SCADA)	Public ICS datasets (NGIDS, DS or CICIDS2017)	Integrated with NIDS (software testbed)
[37]	Explore vulnerabilities and attacks in the industrial control systems of water treatment plant	Penetration testing on the Wastewater Treatment Plant (WWTP) testbed, multiple attack types analyzed	Demonstrated DoS, buffer overflow, and privilege escalation, security recommendations	WWTP ICS testbed	Water Treatment Plant (ICS protocols, Modbus/ SCADA)	SWaT or WADI dataset (iTrust)	Water/ Wastewater Treatment Plant (WWTP) ICS testbed
[27]	Enhance Modbus TCP protocol security using OS-level tools	Fuzzing framework based on information theory to identify PLC protocol vulnerabilities.	Improved security with minimal impact on existing systems	PLC fuzzing environment (Modbus TCP and S7Comm)	Modbus TCP, S7Comm	Generated fuzzing test cases and protocol traffic.	Software-based protocol parsing and fuzzing execution
[38]	Detect cyberattacks using multi-modal data in ICS	Deep multi-modal model combining network + sensor data	Precision 0.99, recall 0.98, F-measure 0.98 on the SWaT dataset	Secure Water Treatment (SWaT) system	ICS SWaT dataset (multi-modal protocols)	SWaT dataset (iTrust Labs)	Secure Water Treatment (SWaT) system (testbed)
[39]	Fuzzing technique based on information theory for industrial proprietary protocols (Modbus, S7Comm)	Protocol parsing via information theory and genetic algorithm-based fuzzing case generation	Discovered two DoS vulnerabilities, including a zero-day in proprietary PLC protocols	Proprietary PLC protocol emulations	Modbus & S7Comm (PLC protocols)	Proprietary industrial protocol traces generated for fuzzing	Software-based emulation of PLC communication for vulnerability testing
[7]	Deep learning-based cyberattack detection for IoT/IIoT devices using Modbus protocol	Neural network architectures with embedding layers, binary & multi-class classification	CNNs, FCNNs, and LSTM achieved up to 99.97% accuracy, the embedding layer improved attack classification	Public Modbus datasets	Modbus (IoT/IIoT)	Public Modbus datasets (e.g., ICS Network Traffic from Malmö University)	Software-based Modbus/ TCP simulation (e.g., using pymodbus, Python)
[40]	Anomaly detection in a smart microgrid system	Autoencoder neural network on network traffic and time-series data	High accuracy in detecting DDoS and false data injection attacks	MCAS Miramar smart microgrid	Modbus/ TCP (SCADA)	Microgrid load/gen data (custom)	MCAS Miramar (physical testbed)
[41]	Detect anomalies in Modbus/TCP using entropy and classification	Statistical entropy analysis + classification (J48, random forest)	High accuracy for DoS (97.87%); cluster-based entropy is effective	Network traffic dataset	Modbus/ TCP	WUSTL IIoT or Gas Pipeline Dataset	Software-defined or physical testbed

Table 2. Cont.

Reference	Purpose of Study	Methodology/ Approach	Findings/ Results	Simulator/ Testbed Used	Protocol Studied	Dataset	Testbed/ Simulation
[5]	Detect/mitigate field flooding DoS on oil and gas infrastructure	Experimental attack on a real industrial testbed, XGBoost ML detection	Field flooding causes PLC memory overflow; ML detects with 99% accuracy	Real industrial CPS testbed	Oil & gas ICS (proprietary)	Process control data (custom)	Real industrial CPS testbed
[42]	Model DoS attacks in IIoT networks, map to MITRE framework	Simulation platform including physical processes and controllers	Standardized attack modeling and forensic artifact extraction	IIoT network sim platform	IIoT protocols (DoS modeling)	MQTT/IoT datasets (N-BaIoT)	IIoT simulation platform (NS-3, Cooja)
[10]	Detect anomalies in air-gapped nuclear plant ICS safety shutdown systems	Hybrid Neural Network: Transformer + LSTM	99.8% detection accuracy for DoS and MITM attacks; white-box model for transparency	ICS testbed replicating nuclear plant components	Nuclear plant ICS (proprietary protocol)	Custom sensor/actuator data simulating nuclear processes	ICS testbed (e.g., with PLCs and HMIs)
[13]	Detect and mitigate ARP spoofing in Modbus TCP/IP communications for Siemens S7 PLCs	Lightweight hardware device, experimental testbed with Siemens S7-1200/1500 PLCs	Effective detection and mitigation of ARP spoofing, maintaining stable communication	Siemens S7 PLC experimental testbed	Modbus TCP/IP (Siemens S7 PLCs)	Custom-generated traffic attack (Siemens S7 Comm protocol)	Siemens S7-1200/1500 PLC experimental testbed
[14]	Federated learning anomaly detection system for DER networks to improve security	Vertical and horizontal federated learning models trained on DER datasets	Improved accuracy and F1-scores over baseline; preserves data privacy during collaborative learning	DER datasets (real-world representative)	DER networks (distributed energy resources)	IEEE PES Cyberattack Dataset (Power system scenarios)	RTDS Simulator or OPAL-RT (Common for DER simulations)
[15]	Enhance cybersecurity of IIoT environments using Modbus/TCP by integrating AI- and SDN-based mitigation	Threat modeling (STRIDE, CVSS) & intrusion detection via ResNet50 CNN with SDN with Transfer & Active Learning–SDN-based mitigation with Thompson Sampling	Detection accuracy > 98% with low false positives; SDN effectively isolated threats in real time	Two VMs: Ryu SDN controller & IDPS host	Modbus/TCP (IIoT)	Custom SDN flow data (generated via scripts)	Two VMs (Ryu SDN controller + IDPS), Binvis & scurre
[43]	Map ICS threats to defensive controls (systematic analysis)	MITRE ATT&CK framework applied to ICS testbeds	Identified key controls and future testbed needs	Review + analysis	Modbus/TCP (IIoT)	Custom SDN data	Ryu SDN + IDPS VMs
[44]	Scanner–Hunter system to identify ICS scanning groups using honeynet data	Traffic representation + self-expanding multi-class classification (SEMCC) applied to 6 years of data	Effectively identified scanning groups, evaluated unknown scanners, and uncovered groups with APT characteristics.	Real-world honeynet deployment.	Modbus	6 years of real-world honeynet datasets (ICS request packets).	Prototype implementation evaluated against popular ML methods and SOTA scanning identification methods.

Table 2. Cont.

Reference	Purpose of Study	Methodology/ Approach	Findings/ Results	Simulator/ Testbed Used	Protocol Studied	Dataset	Testbed/ Simulation
[45]	Simulate a secure SCADA/IoT hardware test bench deploying lightweight ciphers (PCHC and Ascon-ACR).	Performance evaluation of lightweight ciphers on Raspberry Pi compared with the AGA-12 protocol standard.	Identified vulnerabilities (DoS, MiTM) and demonstrated effective execution speed and memory utilization.	Raspberry Pi hardware test bench.	Modbus, MQTT	Custom SCADA/IoT traffic (e.g., servo motor control, water distribution, voltage indicators)	Hardware-in-the-loop simulation simulating a PLC environment using Raspberry Pi.
[46]	Detect MITM attacks in Advanced Metering Infrastructure (AMI)	Real-time AMI testbed; ML model training including random forest + Hybrid Bat Optimization	The hybrid model outperforms others in detecting MITM attacks	AMI testbed with Raspberry Pi, sensors	AMI (Advanced Metering Infrastructure, smart grid protocols)	Smart meter power consumption data (custom)	AMI testbed with Raspberry Pi and sensors
[47]	Analyze cyberattacks on the Modbus protocol in smart grids via a virtual environment	Virtual testbed using open-source software emulating Modbus/TCP network	Detailed cyberattack cycle analyzed; testbed improvements proposed	Virtual Modbus/TCP network testbed	Modbus/TCP (smart grid)	Modbus/TCP packet captures (custom attack traffic)	Virtual Modbus/TCP network testbed (GNS3/Eve, NG)
[48]	Proposed VOTNet framework: A virtual OT network simulating PLCs, network devices, and power systems for cybersecurity assessment.	Cybersecurity and scalability assessment under multiple attacks (DoS, Modbus scanning/data reading, data manipulation/injection, session hijacking).	Demonstrated risk assessment, mitigation strategies, and situational awareness/coordination under cyberattacks.	Virtual Operational Technology Network (VOTNet) environment.	Modbus TCP & Industrial communication protocols.	Simulated/emulated industrial network traffic and control center data.	Virtual testbed combining simulated power systems and emulated PLCs/network devices.

Note on Reference Organization and Corpus Synthesis: To ensure methodological transparency, a clear distinction is maintained between the primary systematically reviewed corpus ($n = 31$ studies) and the supplementary background literature used to anchor the theoretical context. The primary reviewed papers are specifically concentrated within references [2–15,19,27,32–49], representing the exact 31 articles extracted under our strict PRISMA criteria. Conversely, references such as [1,16,20–26,28–31,50,51] serve exclusively as complementary background material, such as systematic reviews, line studies, simulation validations, and theoretical frameworks, and are treated independently of the core statistical datasets of the reviewed corpus.

2.5. Quality Assessment

To evaluate the methodological quality and reliability of the selected studies, a scoring system provided by [16] was applied. Each study was assessed against five predefined criteria:

- Clarity of research objectives (0.5–2 points);
- Methodological rigor (0.5–2 points);
- Experimental validation using datasets/testbeds (0.5–2 points);
- Reproducibility and transparency of results (0.5–2 points);
- Contribution and relevance to intrusion detection and prevention in industrial protocols (0.5–2 points).

Each criterion was scored from 0.5 (partially addressed) to 2 (fully addressed). The total score for a paper could therefore range from 0 to 10 points, where

- 0–4.9 = Low quality (insufficient methodological support; often discarded).
- 5–6.9 = Moderate quality (partially meets requirements; considered cautiously).
- 7–8.9 = Good quality (well-conducted and relevant; included in synthesis).
- 9–10 = High quality (rigorous, reproducible, and highly relevant; strong inclusion).

This scoring allowed for a fair and transparent distribution of quality levels across the reviewed studies.

2.6. Data Synthesis Approach

In view of the significant methodological differences between the studies reviewed regarding the protocol, type of attacks, ML/DL architectures, characteristics of datasets, and experimental environments, a formal quantitative meta-analysis was not feasible. The reviewed studies used a variety of and unstandardized outcomes (accuracy, F1-score, precision, recall, latency and memory overhead), which were measured under different experimental environments (physical testbeds, software simulations and hardware-in-the-loop environments), and it is not possible to statistically validate the aggregation of all these metrics.

Instead, a structured narrative synthesis was used, structured around the five predefined research questions (Section 1). Data from the data extraction table (Table 2) were summarized quantitatively to support the narrative analysis, such as percentage distributions of protocols, ML techniques, deployment types, and validation approaches. Furthermore, a novel taxonomy of IDPFs (Table 3) was built to offer a structured classification, which not only summarizes, but also classifies, the reviewed literature into four dimensions: ML/DL technique, attack type, protocol coverage, and deployment architecture. The taxonomy is a methodological contribution of this review.

Table 3. Selected papers' quality assessment results.

Result (Score)	No. of Papers	Paper Source
5	2	[6,19]
5.25	1	[35]
5.5	3	[36,39,47]
6	4	[7,8,37,42]
6.5	2	[45,46]
7	5	[3,12,32,34,44,48]
7.5	3	[4,41,43]
8	4	[5,11,13,49]
8.5	3	[14,38,40]
9	2	[10,15]
10	1	[33]

3. Results

This section describes and discusses the results of the 31 studies identified in the systematic review process outlined in Section 2. The findings are presented in a manner that directly answers the five research questions (Section 1). Specifically, RQ1 (ML/DL techniques used in IDP) is discussed in Section 3.4, RQ2 (most studied protocols and datasets) is discussed in Sections 3.2 and 3.3, RQ3 (typical IDPF architectures including feature extraction, placement and response mechanisms) is discussed in Section 3.5, RQ4 (evaluation methods, metrics, testbeds and real-world deployments) is discussed in Section 3.6, and RQ5 (open challenges and future directions) is synthesized in Sections 4 and 5. The

subsections move from macro-level bibliometric and quality indicators (Section 3.1), to the analysis of protocol and the landscape of datasets, and then to technical synthesis of methods and architectures leading up to the critical gap analysis in the Discussion.

3.1. Publication Statistics

The adopted scoring method for the selected articles, as outlined in Table 2, considering the inclusion criteria, ensured a fair and transparent distribution of quality levels across the reviewed studies. The results of this assessment are summarized in Table 4.

Table 4. The assessment of the criteria.

Criterion	Description	Max Score
Research Objectives	Clarity and specificity of stated goals	2
Methodology	Appropriateness of ML/DL technique used	2
Experimental Setup	Realism of testbed or simulation	2
Reproducibility	Dataset availability and metric reporting	2
Significance	Contribution to the field	2
Total		10

The resulting scores ranged between 5.0 and 10.0, with most studies clustered between 6.0 and 8.5. Using the weighted average Formula (1) [28]

$$W = \frac{\sum_{i=1}^n (w_i \cdot x_i)}{\sum_{i=1}^n w_i} \quad (1)$$

where x_i represents the score and w_i the number of papers in each score group, the overall quality level across all studies was found to be approximately 7.2. This indicates a consistently strong methodological foundation among the reviewed literature, reflecting both the increasing maturity of research in AI-driven intrusion detection for industrial systems and a balanced methodological quality across publications. Only one paper achieved the maximum score of 10, underscoring its exceptional rigor and reproducibility. To position and contextualize our study within the existing body of literature, we analyze several bibliometric indicators that reflect the structure and evolution of the research field. The results presented in these figures are derived from a bibliometric analysis conducted using the bibliometrix package in R and its Biblioshiny interface, applied to the final corpus of studies selected for this review [28]. The objective is to identify the leading contributors in the field, highlight the most influential institutions and countries, and examine the overall structure of the research community. This analysis provides insights into the maturity, concentration, and research leadership of intrusion detection and prevention systems for industrial communication protocols based on machine learning and deep learning techniques.

Figure 3 shows that publication activity is concentrated among a few key researchers. Corchado Emilio leads with eight papers, followed by Boutaba Raouf and Herrero Alvaro with seven each. Authors like Biswas Santosh, Subba Basant, and Wan Ming contributed six papers, while Ahmad Iftikhar, Ghosh Partha, Jiang Yaping, and Karmakar Sushanta each published five. This highlights a small group of highly prolific authors, alongside several consistent contributors, who drive research in this field.

Based on the analysis, we can conclude that AI-based intrusion detection and prevention approaches for industrial communication protocols have become rather mature due to the consistency in methodology quality shown by most papers, as well as the appearance of multiple highly active research teams. The fact that the average quality score reached 7.2 proves that most research efforts feature proper design of the experiment, significant outcomes, and valuable contributions. Yet, taking into account the variance in the scores,

one should ensure that aspects such as reproducibility, availability of the datasets used in studies, and the realism of environments chosen for verification purposes are given enough attention. Furthermore, it is worth mentioning that the high concentration of published papers within just several research teams indicates that this sphere has yet to fully develop.

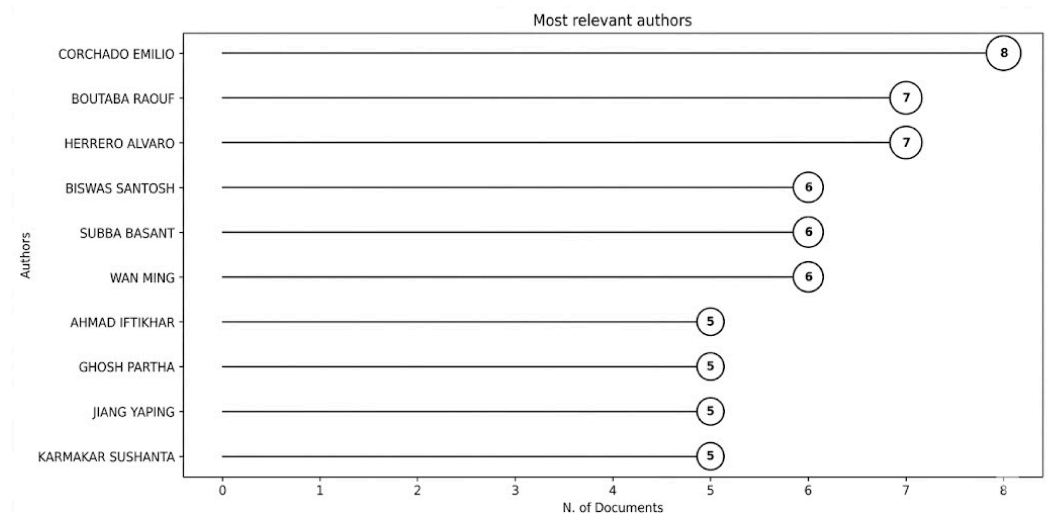


Figure 3. Most relevant authors.

3.2. Protocols Studied

Addressing RQ2, this subsection analyzes the distribution of industrial communication protocols across the 31 reviewed studies.

A review of the literature, illustrated in Table 2, reveals that most studies on intrusion detection and prevention systems in industrial control systems (ICSs) have focused heavily on the Modbus/TCP protocol, as it is one of the most widely used protocols in SCADA systems and industrial control circuits. Other protocols, such as MQTT, S7Comm, IEC 60870-5-104, IEC 61850, DNP3, and OPC UA, are also present, but they are still relatively less addressed than Modbus. They are often associated with specific sector applications such as smart grids, oil and gas infrastructure, and water plants. This focus reflects the importance of Modbus as an older protocol that remains widely used despite its security vulnerabilities. On the other hand, it highlights a research gap in the literature regarding other protocols specific to the new generation of IIoT and smart grid systems, which are characterized by greater interconnectivity and more complex protocol diversity.

Hence, the current systematic study seeks to evaluate existing frameworks using machine learning and deep learning techniques to detect and prevent security threats, with a focus on the inclusion of various industrial protocols and not being limited to traditional protocols only.

Figure 4 shows the distribution of communication protocols examined in the 31 reviewed studies. The majority (16 out of 31 studies, representing 51.6%) focus on the Modbus and Modbus/TCP protocols, reflecting their dominant role in industrial communication and SCADA networks. Other protocols such as SCADA variants, S7Comm, and MQTT appear less frequently but are gaining attention due to their increasing use in IIoT and smart grid systems. Only a few studies examined IEC 60870-5-104, DNP3, OPC UA, or DER-related protocols, indicating that these areas remain underexplored. This concentration highlights the continued relevance of Modbus and the existing research gap concerning newer IIoT communication protocols.

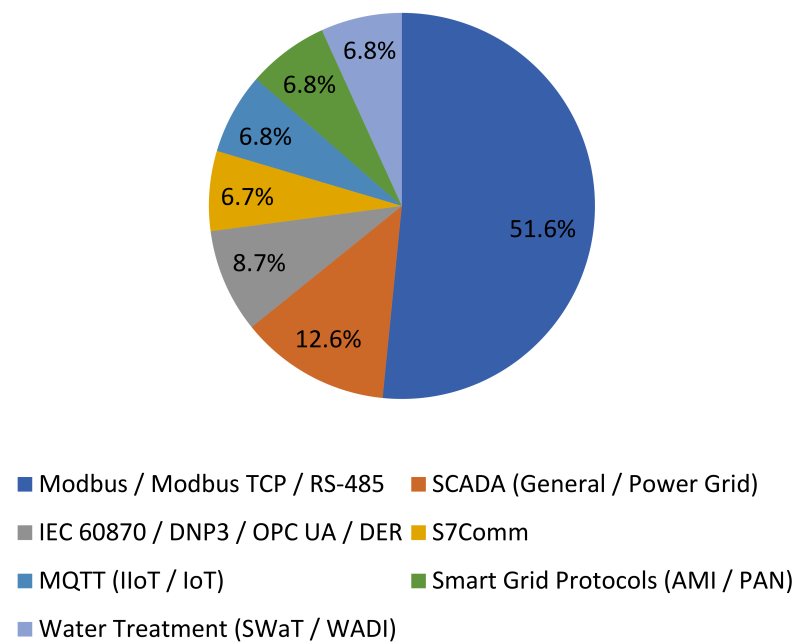


Figure 4. Distribution of industrial communication protocols among the 31 reviewed studies.

From the analysis of the protocol, one can observe a clear tendency for researchers to focus more on Modbus/TCP, which results from its frequent use in industry, as well as from its recognized security threats [2,3,5]. Although the focus allows researchers to make great progress in developing intruder detection and prevention methods for a particular protocol, it can also be viewed as an indication of the fact that the bulk of the existing research base still relies on legacy industrial communication infrastructures. At the same time, the development of protocols such as MQTT, OPC UA, IEC 60870-5-104, DNP3, and IEC 61850 has been neglected, even though they have become increasingly popular in the domain of IIoT, smart grids, and cyber-physical systems [15,36,42]. Therefore, the issue of protocol diversity and interoperability needs to be taken into account when developing a modern cybersecurity framework. One of the major areas of research to be addressed in the near future concerns the need for evaluating artificial intelligence-based approaches to ensure security in a variety of protocols used in industry.

3.3. Datasets and Testbeds

Also addressing RQ2, this subsection highlights the datasets and testbeds employed in recent studies, as summarized in Table 2 which shows that researchers have relied on a variety of datasets and testbeds to evaluate intrusion detection and prevention frameworks for industrial communication protocols. Public datasets such as the SWaT, WADI, and ICS have been widely used, along with custom datasets created specifically to simulate specific attacks, such as RS-485 spoofing or false data injection attacks in electrical networks. Some studies have relied on software simulation environments such as NS-3, Node-RED, and GNS3, while others have focused on more realistic physical environments such as PLC systems, water plants, and smart power plants. This diversity reflects the need to strike a balance between the ease of use and reproducibility provided by software simulations and the realism and complexity provided by physical testbeds. However, it is noted that many studies still rely on traditional or synthetic datasets, highlighting a research gap related to the lack of recent data representing evolving cyber threats in IoT and smart grid environments. Hence, this systematic review highlights the role of machine learning and deep learning in testing the effectiveness of protection systems, focusing on the diversity and completeness of the data sources used.

Analysis of the datasets and testbeds suggests that scientists understand the significance of validating their frameworks for IDSs/IPSS in real-life settings in the industrial sector. In addition to using existing datasets like SWaT, WADI, and ICS, there are a large number of studies where scientists have used customized datasets, hardware-in-the-loop (HIL), and real-world industrial testbeds [4,5,12,34]. This implies that there is an understanding that although good results can be achieved with some dataset or testbed, they may not be transferable to real-life settings in the industry. However, the usage of custom datasets also indicates another problem that scientists are facing at the moment—the lack of availability, comparability, and reproducibility in the datasets used by different researchers [12,13]. Based on this, perhaps the largest scientific need highlighted in this review is to come up with industrial cybersecurity datasets that are publicly available and realistic.

3.4. Machine and Deep Learning

This subsection addresses RQ1 by examining the ML/DL techniques applied across the reviewed studies. This part illuminates the way researchers are addressing cybersecurity issues in industrial control systems and IoT. In general, their strategies can be divided into three categories: developing intelligent detection models based on machine learning and deep learning, evaluating the resulting models in real-world or simulated physical/physical settings, and enhancing the security of the underlying communication protocols. Combined, these measures demonstrate the development of the field towards improved detection, prevention, and response to cyber threats in critical infrastructure.

Scholars have ventured in several interesting directions to address the multifaceted issue of cybersecurity in industrial control and in IoT space. The approaches they follow usually hinge on whether they are about mode of attack detection, implementation of secure systems, or the analysis of vulnerabilities.

Some researchers [24] conducted a systematic literature review of 124 peer-reviewed papers on ML-based anomaly detection for IoT security, following PRISMA guidelines and Kitchenham's methodology. The study classified IoT attacks across a four-layer architecture and evaluated classical ML models such as random forest, LSTM, and CNN, which achieved high detection rates on datasets like BoT-IoT and NSL-KDD but struggled with scalability and generalization. The authors also examined Quantum Machine Learning (QML) approaches, where the QAE-QkNN framework achieved an F1-score of 98.26%, though practical deployment remains limited by hardware constraints. Future directions included integrating Explainable AI, strengthening federated learning security, and developing adaptive learning models. This study is particularly relevant as it addressed industrial protocols such as Modbus and DNP3, SCADA and IIoT attacks, and industrial datasets including Edge-IIoT and ICS-SCADA.

Creating smart detection systems is a very intensive task. Some of the researchers used ML and/or DL models, which they trained on the datasets to detect the small indicators of an attack. Indicatively, other groups got excellent results by developing specialized neural networks. In [3] the authors were specifically interested in identifying DoS attacks on the Modbus/TCP systems using their MODLSTM network, whereas [10] developed a hybrid Transformer-LSTM model that successfully identified DoS and MITM attacks with an impressive accuracy of 99.8% in a simulated nuclear plant environment. Some, including [41], have been successful in pursuing a combination of different methods, with statistical entropy analysis alongside traditional classifiers, such as random forest, detecting anomalies.

Nonetheless, an intelligent algorithm will not be very helpful unless it has been tested in real-world conditions. This is where the second key approach comes in: physically and

virtually constructing testbeds. Such arrangements play a significant role in the transfer of theory to practice. In one example, researcher in [25] used a real-time cyber-physical testbed system using OPAL and RT simulators and real SEL-351S relays to experiment with the effects of false data injection (FDI) and DoS attacks on power grid stability, along with testing their system adaptive response strategies. Many other researchers, including the authors of [34], who used a real compressor station, and [46], who used Raspberry Pi devices, used their own environment to simulate threats and implement defenses in a safe environment.

There is also a major undertaking besides detection and testing to strengthen the protocols themselves. As industrial systems are based on protocols such as Modbus TCP which were not developed with security in mind, numerous researchers are striving to overcome this situation. Research works such as the one of [47] carefully examine the ways such protocols are used, charting the whole cyberattack process. The solutions are creative and feasible. Researchers in [11] was able to incorporate Elliptic Curve Cryptography (ECC) directly into Modbus TCP communications, making it encrypted and not overloading the devices. The interesting project of [26] was based on the use of eBPF, which is a technology of Linux kernel used to transparently insert a layer of security into the traffic of the Modbus TCP. In the meantime, in [15] bigger-picture networks were seen, integrating AI-based detection with software-defined networking (SDN), not only to identify threats but also to automatically isolate them on-the-fly.

The realm of new defensive technologies goes beyond, however, and includes three paradigms which are becoming increasingly relevant to the security of ICSs. First, digital twin frameworks are becoming process-aware IDS platforms, which allow for detection of attacks that sneak through the ICS network, such as false data injection or other attacks that alter the status of the process without generating a traditional network anomaly alert. Recent studies have shown F1-scores of 96.3% against FDI and DoS attacks in a digital twin-based SCADA environment with false positive rates less than 2.5% [50]. Lin et al. then applied the same method to a hybrid deep learning model combined with a digital twin of a SCADA system, resulting in real-time anomaly detection and low latency [51]. Secondly, blockchain-based IDS frameworks overcome the drawbacks of centralized IDSs in distributed IIoT systems, providing tamper-proof logs and distributed threat intelligence sharing among organizations [21]. Third, among the analyzed corpus, federated learning (FL) appeared as a specialized primary architecture in 6.45% of the selected papers ($n = 2$), highlighting its emerging but highly localized role in decentralized data privacy protection across distributed energy and smart industrial environments. FL combined with 6G communication infrastructure is a very promising route for IIoT security, allowing for decentralized model training with low latency across geographically distributed resources while preserving privacy, without the need to send sensitive operational data to a central location.

The analyzed articles clearly indicate the progression from the use of standard machine learning algorithms towards more advanced deep learning models and hybrid systems that are able to detect and understand intricate temporal, behavioral, and contextual patterns of industrial communications [3,10,24,41]. While the use of classical machine learning algorithms remains efficient enough to guarantee accurate results, modern studies often emphasize the benefits of using hybrid systems with multiple learning processes involved in order to enhance the accuracy of attack detection and increase the level of resilience to attacks. At the same time, the appearance of such trends as federated learning, digital twins, blockchain-based cybersecurity solutions, and intelligent software-defined networks also indicates an ongoing transition towards intelligent cybersecurity environments [14,15,21,50]. Nonetheless, while modern models demonstrate impressive results in

detecting cyberattacks, some issues of explainability, processing costs, and applicability of machine learning systems still require additional attention in further research.

3.5. Architectures and Deployment

This subsection addresses RQ3 by analyzing the architectures and deployment strategies identified in the reviewed studies, and looks at how cybersecurity solutions are designed and deployed in real industrial settings. Researchers use everything from physical testbeds that mimic real infrastructure to smart adaptive systems that combine AI and network controls. Some approaches focus on securing legacy protocols without major changes, while others integrate detection and automated response to tackle threats in real time. Overall, the goal is to create solutions that are both effective and practical for complex industrial environments.

When it comes to moving from theory to practice, the architecture of a cybersecurity solution and how it is deployed are just as critical as the detection algorithm itself. The research reveals a fascinating spectrum of deployment strategies, each tailored to different industrial environments and constraints.

For validating attacks and defenses under the most realistic conditions possible, many researchers invest in building specialized physical testbeds. These are not just simulations: they are functional replicas of critical infrastructure. For instance, in [25] a sophisticated real-time cyber-physical testbed was constructed using OPAL, RT hardware and real SEL-351S relays to study how FDI attacks could destabilize an entire power grid. Similarly, the authors in [32] took their work directly to the field, conducting penetration testing on a real compressor station's ICS to propose simple but effective PLC-based integrity checks. These physical deployments, from [46] a Raspberry Pi-based AMI testbed to [5] a real industrial CPS setup, are invaluable. They provide undeniable proof of vulnerability or a solution's effectiveness in an environment that mirrors the complexities of a live industrial site.

A significant architectural challenge lies in securing legacy protocols never designed for a connected world. Here, research has focused on protocol-level enhancements and integrations. Some solutions add a security layer directly onto existing systems. Ref. [11], for example, successfully integrated ECC directly into Modbus TCP communications, demonstrating that strong security is possible even for resource-constrained devices. In a clever twist, other authors in [27] leveraged eBPF technology to transparently enforce security rules at the operating system level for Modbus TCP, requiring no changes to the legacy applications themselves. Other architectures, such as the stateful monitoring NIDS proposed by [36], work by deeply understanding industrial protocol semantics to catch malicious commands that would otherwise slip past a standard firewall.

Perhaps the most forward-looking architectural trend is the move towards intelligent and adaptive security frameworks. This involves integrating various technologies into a cohesive, automated defense system. The work by [15] is a prime example. They built an architecture that integrates an AI-based ResNet50 CNN for detection with a software-defined networking (SDN) controller for mitigation, allowing their system to not only find threats but also automatically isolate them in real time. This philosophy of integrated, intelligent defense is also clear in the federated learning system for DER networks by [14], which preserves data privacy while enabling collaborative learning across different organizations.

Ultimately, the chosen architecture, whether a physical testbed, a cryptographic add-on, or an AI SDN hybrid, is a direct response to the unique challenges of the industrial environment it aims to protect. The best solutions are those that are not only effective but also deployable within the rigid constraints of real-world industrial control systems.

It could be noted from the architecture analysis that the evolution of cybersecurity systems for the industrial domain not only involves intrusion detection systems but tends towards advanced security systems based on technologies such as artificial intelligence-based security systems, encryption methods, software-defined networks, and distributed learning algorithms [11,14,15,27,36]. As could be seen from the analyzed literature, it becomes clear that a good security architecture system does not only provide intrusion detection capabilities, but also provides scalability and automation capabilities and is compatible with the current industrial architecture. In addition, the use of testbeds, protocol-aware methods, and artificial intelligence-based security solutions show the need for their implementation instead of performance evaluation. Nevertheless, some challenges to the implementation of cybersecurity architectures in industrial settings include interoperability problems, legacy system support, privacy preservation, and complexity of the architecture.

3.6. Evaluation

This part discusses the evaluation of cybersecurity solutions that go beyond detection accuracy. It emphasizes the significance of realistic performance, such as velocity, effectiveness of resources and real-world implementation. The emphasis is placed on demonstrating not only high accuracy but also operational viability and resilience in the industrial setting.

To assess the effectiveness of these cybersecurity strategies, it is necessary to consider more than the detection rate. The questions asked by the researchers are: Does this solution work in practice, and what does it cost to operate? The results uncover an emphasis on demonstrating real-world viability and gauging the practical tradeoffs.

The results can be impressive when it comes to pure detection power. Numerous studies have documented very high accuracy rates, often surpassing 97% to 99% of attacks. An example is that when detecting attacks on a nuclear plant safety system, Ref. [10] achieved a remarkable 99.8% accuracy, whereas when detecting the types of attacks, they detected 86.13% as compared to 99.97%. But pure precision is not all. Other teams, such as the one cited in [19], also focused on reducing false alarms, achieving a false positive rate of less than 1 percent, which is critical to get operators to trust an automated system. These figures suggest that on a pure performance basis, machine learning models are becoming very competent in detecting malicious intent in industrial traffic.

But an accurate, but slow or resource-heavy, solution is of no use on a factory floor. It was here that the issue of measuring computational and resource efficiency became important in 2021. This is one of the key components of the results of several studies, which is essential to such devices such as Raspberry Pi devices. Equally, in [11] discovered that their Elliptic Curve Cryptography (ECC) implementation offered stronger security with low overhead which is a critical finding for legacy devices with limited processing capability. Such considerations are indicative of a well-rounded perspective that a defense should not merely be powerful, but also practical and cost-effective to execute.

Lastly, the strongest assessment consists of actual, real-life validation. It is one thing to test on a curated dataset and another to observe a system in realistic conditions. By implementing attacks and defenses on real industrial systems, such as a real compressor station and a real CPS testbed, research teams, including [34,41], gained invaluable experience. The importance of their findings, which revealed tangible vulnerabilities such as PLC memory overflow and effective mitigation measures, are weighty since they were not obtained in pure simulation. Such a transition of theoretical performance to established practical effectiveness is, perhaps, the most significant development in the context of assessing ICS security solutions today.

Finally, the evaluation blueprint implies that the profession is growing up. Being able to score highly on accuracy is no longer good enough. The most effective research currently

also shows that a solution is efficient, resilient and indeed capable of withstanding the challenges of a real industrial environment.

4. Discussion, Key Findings, and Limitations

This section discusses the key findings and limitations of current ICS and IIoT security research. It highlights major trends such as the dominance of Modbus/TCP studies, the rise of advanced AI-driven detection models, and the growing shift toward realistic testbeds and real-world validation. At the same time, it underscores persistent challenges, limited datasets, scalability issues, underexplored protocols, and a lack of integrated remediation strategies that must be addressed for solutions to truly succeed in practice.

Whereas the described achievements can be perceived as individual breakthroughs in intrusion detection technologies, the reviewed studies have shown that the area has undergone major changes. It has moved from being focused on protocol-oriented approaches to implementing intelligent adaptive protective systems able to function in complex cyber-physical conditions [14,15,20,50]. Such evolution has been driven by the increased convergence of industrial control systems, IIoT infrastructure, and artificial intelligence [1,22,23].

The first noteworthy observation that comes to light upon examination of the reviewed literature relates to the dominance of the Modbus/TCP protocol, which continues to be utilized extensively in experiments and assessments [2,3,5,7,8]. This phenomenon can be attributed to the prevalence of traditional industrial infrastructures, although it also points to the fact that much of today's expertise is still focused on legacy communication scenarios. On the other hand, the gradual proliferation of MQTT, IEC 60870-5-104, DNP3, OPC UA, and software-defined industrial networking proves that industrial communication systems have become increasingly diverse [15,36,42,44]. Therefore, the design of next-generation IDS/IPS tools will require cross-platform support for a variety of protocols [21,23].

It is important to highlight that there is evidence of the emergence of a significant methodological shift from classical machine learning methods towards hybrid architectures based on deep learning. In previous works, there was a tendency to use classification-based detection models together with classical machine learning techniques [32,41,46], while modern research tends to use architectures like Transformer-LSTM, MODLSTM, and deep ensembles that take into account temporal aspects of attacks, contextual information, and multi-modality [3,10,38]. This might be due to understanding that static detection is not enough against more advanced cyberattacks.

The other emerging trend relates to the growing importance of operational realism. From the literature, one can observe that there has been a steady shift from simulations based solely on software to hardware-in-the-loop simulators, cyber-physical systems, smart grid systems, water treatment plants, and pilot deployments in industry [4,5,10,12,34,37]. This shift indicates the realization of the fact that merely being able to provide high-accuracy results is not enough for the suggested methods, if they cannot perform well in realistic industrial scenarios [5,15,34].

Further analysis of the selected research studies highlights critical issues with respect to the tradeoffs between various validation environments that have been employed for evaluating intrusion detection/prevention architectures. The approach based on software-based simulations is still one of the most common validation techniques, which can be explained by the low cost of experiments and controllable environment [2,32,35,42,47]. This technique allows researchers to create extensive sets of attack data and analyze a variety of attack scenarios using similar conditions. It is therefore common for simulation experiments to yield impressive results in terms of the accuracy of detection and classification. However, the findings may be somewhat skewed owing to the absence of delay and other real-world factors.

HIL setups serve as a medium-level validation technique that entails the integration of physical control units with simulations of industrial process [3,4,14,48]. As opposed to pure simulation techniques, HIL systems possess the capability of providing more realistic conditions without sacrificing experimental control. These setups can prove particularly useful when assessing the impact of latency attacks and studying the interaction of intrusion detection techniques and control process operations. Nonetheless, HIL systems necessitate costly hardware and labor-intensive development.

The realistic nature of physical deployment and testbeds for industrial operations ensures that the most reliable and accurate information is obtained for applicability [5,12,34,37,38,40,46]. Experiments on water treatment systems, smart grids, compressor stations, and industrial cyber–physical systems have illustrated the effects of communication overhead, sensing error, operational variance, and resource constraints on the performance of detecting techniques. Although the accuracy levels might sometimes be lower compared to simulations, they still offer more reliable information about feasibility.

From a practitioner perspective, the reviewed literature suggests that reported performance metrics should always be interpreted in the context of the validation environment. High accuracy achieved within simulation environments does not necessarily guarantee equivalent performance in operational deployments. As validation environments become increasingly realistic, additional factors such as latency, computational overhead, scalability, robustness, and resilience become equally important evaluation criteria. Therefore, future studies should adopt multi-stage validation strategies that combine simulation, HIL testing, and physical deployment to provide a more comprehensive assessment of intrusion detection and prevention frameworks.

All these aspects, taken together, clearly point to a trend in which cybersecurity ecosystems that utilize artificial intelligence, distributed computing, and industry requirements are being considered. Moreover, recent advances in the use of federated learning approaches [14,15], software-defined networking [15], digital twins [50], blockchain-based security measures [21], and edge–fog–cloud computing systems [20] confirm this trend. Consequently, future research in this area should not only aim at enhancing the efficiency of attack detection, but also consider development of cybersecurity solutions that are trustworthy, understandable, interoperable, privacy-preserving, and deployable [20,21,50].

The findings indicate a paradigm shift towards increasing the reality of the evaluation. Initial work was either based on real-world data, as in [25], or used virtual simulation environments. More recent studies have shifted to physical testing conditions, such as models of nuclear plants [10] and actual compressor stations. Moreover, some studies have used integrated hardware–software systems [9,35], and joint cyber–physical simulation environments. Meanwhile, the detection methods have moved beyond the traditional statistical methods of detection [41] towards more advanced AI- and deep learning-based approaches. These are hybrid models such as those that incorporate Transformers and LSTMs and achieve 99.8% accuracy [10], federated learning frameworks that are developed to protect the privacy of data and achieve 99.8% accuracy [10], and lightweight hardware-based solutions such as ECC encryption [11].

Important results and significant innovations are being developed. Neural networks, including CNNs, LSTM, and autoencoders, are proving to be superior to traditional methods, with accuracy rates of over 97 percent in detecting DoS and MITM attacks [7,48]. With regard to the nature of the attacks, research papers like [49] have shown that DoS attacks are very effective and can be detected with a high accuracy of up to 99% using XGBoost, whereas MITM attacks require a multi-layered mitigation strategy, such as having dedicated hardware that can combat an address spoofing attack [13]. The use of lightweight cryptographic algorithms and protocol-specific integrity checks to protect legacy devices

has also been demonstrated in some works; however, such methods are not comprehensive enough to protect all devices in the field of information technology (IT). The emphasis on post-attack response has been moved to proactive prevention by frameworks such as MITRE ATT&CK [22] and honeynet threat tracking techniques. Although there has been a lot of development, there are still obstacles that hinder the effectiveness of research. There are very few large-scale real-world datasets of a wide variety of attacks and environment operational scenarios [12]. The challenge of obtaining publicly accessible synthetic data also contributes to the lack of reproducibility of the results. It also overemphasizes detection accuracy, as it was found that a high accuracy in the lab does not correlate to a strong model in real-world conditions and there is a lack of studies examining how models behave when subjected to advanced attacks or malware [31]. Other issues are associated with scalability and response time, as some deep learning-based models can only be run on the computational resources available, possibly not including those offered by limited-power controllers or sensors [44].

Moreover, the existing protocols like OPC UA, PROFINET, and IEC 61850 are not as popular as the Modbus/TCP protocol and have not yet been thoroughly studied (although they are gaining more popularity) [23]. Moreover, the integrated response mechanisms have not been developed yet, and a lot of research is being conducted on the detection aspect without going into remedial actions or adaptive control in the event of an attack [29]. In this regard, the need to develop integrated frameworks, which combine detection, remediation, and recovery, emerges, such as using software-defined networks (SDNs) to isolate attacks rapidly, and applying adaptive control to maintain operational continuity. Enhancing security in the design of new protocols such as OPC UA and IEC 61850 using formal verification methods and adding advanced security features is a key step [23].

The findings also underscore the priority of developing standard and comprehensive datasets that include network traffic, sensor data, and controller logs, with broad coverage of various attack types. Collaborative initiatives between industry and universities could contribute to the creation of repositories like SWaT and WADI, dedicated to the energy, oil and gas, and manufacturing sectors. The development of flexible and adaptive detection models is also an urgent necessity, including attack-resistant learning, online learning, and incremental learning, as well as the application of reinforcement learning techniques to build proactive defenses [27]. Furthermore, lightweight security technologies are required to accommodate the capabilities of resource-constrained devices, such as legacy controllers and sensors in IIoT environments [11]. With the growing interest in learning in the IIoT, it is becoming essential to ensure privacy protection when sharing data between parties using techniques such as differential privacy and secure aggregation [14].

There is also a need for in-depth studies that link technical solutions to human and organizational factors, including operator behavior, socio-technical risks, and supply chains [1]. Finally, establishing standard performance metrics, in terms of response time, false alarm rate, and cost, is essential to facilitate comparison between studies and identify the most appropriate solutions for practical application.

Given the systematic analysis conducted in this review, Table 5 introduces a new taxonomy for IDP frameworks for industrial communication protocols based on four different domains: ML/DL technique, attack type, protocol coverage, and deployment architecture.

To complement the descriptive taxonomy in Table 5 and provide actionable guidance for researchers and practitioners, Table 6 presents a structured gap analysis mapping identified research limitations to concrete future directions.

Through this review, it becomes clear that the most common and dangerous attacks on ICS networks include DoS, command or data injection, MITM attacks, reconnaissance attacks, replay attacks, impersonation, and eavesdropping. These attacks are of particular

importance because they represent the focus of detection and prevention efforts, providing opportunities for in-depth exploration of the latest protection methods and identifying solutions capable of preventing these threats or at least mitigating their potential impact.

Table 5. Classification of selected studies.

Category	Sub-Category	Studies	Example Refs
ML/DL Technique	Hybrid (Transformer + LSTM)	4	[3,10]
	Federated Learning	3	[14,15]
	Classical ML (RF, SVM)	8	[41,46]
Attack Type	DoS	12	[3,5,12]
	MITM	9	[4,13,46]
	Data Injection	6	[4,38]
Protocol	Modbus/TCP	16	[2,7,8]
	MQTT/IIoT	4	[42,44]
	Physical Testbed	10	[5,34]
Deployment	Software Simulation	14	[3,32]
	HIL/Cyber-Physical	7	[4,10]

Table 6. Research gaps and future directions.

Research Dimension	Current State in Reviewed Literature	Identified Gap	Recommended Future Direction
Dataset Coverage	67% of studies use synthetic or custom datasets.	Insufficient diversity of real-world attack scenarios.	Establish collaborative academic–industry dataset repositories covering energy, manufacturing, and oil and gas sectors.
Protocol Coverage	51.6% of studies focus on Modbus/TCP.	Underrepresentation of OPC UA, PROFINET, and IEC 61850.	Develop protocol-agnostic and multi-protocol IDPFs.
Threat Model Breadth	Dominant focus on DoS and MITM attacks.	Advanced Persistent Threats (APTs) remain largely unexplored.	Design MITRE ATT&CK for ICS-aligned evaluation benchmarks incorporating advanced attack scenarios.
Privacy-Preserving ML	Only 9.7% of studies employ federated learning.	Limited support for cross-organization collaborative detection.	Implement federated learning enhanced with differential privacy and secure aggregation techniques [20,21].
Response Integration	Most studies focus exclusively on intrusion detection.	Lack of integrated detect–respond–recover pipelines.	Develop SDN-based automated isolation and adaptive control mechanisms for end-to-end incident response.
Emerging Paradigms	Minimal coverage of digital twin and blockchain technologies.	Gap with the evolving security requirements of Industry 4.0.	Utilize digital twin testbeds as standardized evaluation environments [51] and explore blockchain-based collaborative IDS frameworks [21].
Evaluation Standardization	Evaluation metrics vary considerably across studies.	Prevents reliable cross-study comparison.	Establish standardized ICS security evaluation benchmarks incorporating latency, false positive rate (FPR), and computational cost metrics.

While many of the reviewed studies report impressive detection performance, translating these results into operational industrial environments remains a significant challenge. One of the primary obstacles concerns computational resource limitations. Many industrial controllers, sensors, and edge devices operate under strict processing, memory, and energy constraints, making the deployment of complex deep learning models difficult without specialized hardware acceleration or edge-computing support [11,15,44].

Another important challenge is latency sensitivity. Industrial control systems often require near real-time decision-making, particularly in critical infrastructures such as power grids, water treatment facilities, and manufacturing systems. Although advanced deep learning architectures frequently achieve high detection accuracy, their computational overhead may introduce delays that are unacceptable in safety-critical environments [4,10,34]. Consequently, practitioners must often balance detection performance against response speed and operational continuity.

Legacy system compatibility represents another major deployment barrier. Many industrial facilities continue to rely on aging protocols and equipment that were not originally designed to support modern cybersecurity mechanisms [2,3]. Integrating AI-driven security frameworks into these environments may require substantial infrastructure modifications, increasing deployment complexity and cost. Furthermore, heterogeneous industrial environments often contain multiple protocols, vendors, and communication standards, complicating system integration and interoperability.

The availability and quality of training data also remain persistent concerns. Many studies rely on synthetic datasets or controlled laboratory environments, while obtaining representative real-world industrial datasets remains difficult due to confidentiality, privacy, and operational security considerations [12,13]. This limitation raises concerns regarding model generalization and long-term effectiveness when deployed in operational settings.

From an operational perspective, explainability and trust are becoming increasingly important requirements. Security analysts and industrial operators must understand why a detection model generated a particular alert before taking corrective actions. Highly complex deep learning models may therefore face resistance in practice if their decisions cannot be adequately interpreted or audited [10]. Similarly, false positive alerts can disrupt industrial processes and reduce operator confidence in automated security systems.

Finally, maintaining AI-based security systems introduces additional lifecycle challenges. Models require periodic retraining, updating, and validation to remain effective against evolving threats. This process may impose significant operational and organizational burdens, particularly in large-scale Industrial Internet of Things environments where thousands of devices and communication flows must be continuously monitored [14,15].

These practical considerations highlight that successful industrial cybersecurity solutions must balance detection performance with scalability, interpretability, deployment feasibility, and operational reliability. Therefore, future research should evaluate proposed frameworks not only in terms of accuracy but also in terms of their readiness for real-world deployment and long-term operational sustainability.

5. Literature Gap and Future Directions

This systematic review indicates that studies on the security of industrial systems and the Industrial Internet of Things (IIoT) have gone a long way in coming up with intrusion detection and prevention models that are founded on machine learning and deep learning. Modbus/TCP is the most widely studied of all communication protocols, with IEC 60870-5-104, DNP3, MQTT, and S7Comm coming next. One trend is the increased use of hybrid datasets combining controller logs, sensor measurements and network traffic to improve model robustness and realism.

The potential of hybrid deep learning models, such as Transformer-LSTM and MODLSTM frameworks, has been demonstrated in recent research, establishing exceptionally high detection rates against DoS, MiTM, and data injection attack. No less important is the growing dependence on real-world and virtual testing conditions, and sophisticated simulators like OPAL-RT, which stress the importance of empirical validation to guarantee the practical reliability of the solutions presented. The new directions in research also

emphasize the merging of edge, fog and cloud computing layers, the application of federated learning to facilitate distributed intelligence, and the reinforcing of protocol security measures, all toward finding a balance between accuracy, latency and privacy protection.

Although there is this promising development, there are several issues that have not been addressed. The most urgent one is the lack of detailed real-world data, which involves a variety of attack categories and operational scenarios. The current literature mostly relies on synthetic or small-scale testbeds, which limit the extrapolation of findings. Future research should thus focus on the development of standardized, large-scale, and publicly available industrial data, which has been collected through a joint venture between academia and industry, especially in the energy, oil and gas, and manufacturing industries. These datasets ought to be a combination of network traffic and controller logs as well as sensor-level data to support reproducible and comparative experimentation [12,38].

A third fundamental gap is the need to develop adaptive and resilient detection models that can adapt to the threat landscape. The future directions that seem promising are attack-resistant learning, incremental learning, and reinforcement learning methods, which can help proactive and self-healing defense mechanisms [27]. Similarly, the lightweight security infrastructures to support resource-constrained IIoT devices, including older-style controllers and low-power sensors, are a relatively unexplored field that requires particular interest [11]. To instantiate these paradigms into actionable operational technology (OT) architectures, future research must pioneer an intelligent, multi-layered intrusion detection and prevention framework (IDPF) driven by deep learning mechanics. This proposed framework segments industrial defenses into specialized functional tiers: a localized boundary layer utilizing lightweight algorithms like One-Class Support Vector Machines (OC-SVMs) or Gated Recurrent Units (GRUs) deployed at edge gateways for sub-millisecond anomaly filtering; a core analytical layer hosting a central Transformer-BiLSTM network that acts as a process-aware detection engine by cross-correlating multi-protocol payload features with physical sensor telemetry; and an autonomous prevention layer powered by Deep Reinforcement Learning (DRL) agents that evaluate classified threats against strict physical control boundaries to execute risk-aware mitigation loops in real time.

Integration of detection, response, and recovery layers should also be considered in future research to foster the adoption of holistic defensive ecosystems as opposed to isolated monitoring. This can be done with the help of software-defined networking (SDN) mechanisms that isolate nodes that have been compromised on-the-fly and adaptive control mechanisms that provide continuity in operations even after cyberattacks [4,15]. Similar initiatives should be undertaken to securely redesign modern industrial protocols, including OPC UA and IEC 61850, with formal verification and cryptographic assurance to enhance protocol reliability [23].

Additionally, privacy of and trust in data are yet to be basic concerns during cross-network collaboration between the various stakeholders. Differential privacy, secure aggregation, and homomorphic encryption will play a crucial role in facilitating safe, large-scale learning and information sharing [14]. To operationalize these techniques within highly distributed and competitively sensitive industrial sectors, future work must establish a robust, multi-layer federated learning (FL) framework tailored for heterogeneous IIoT environments. This architecture should deploy decentralized local learning nodes directly onto regional industrial edge facilities, enabling them to independently train on local traffic without exposing raw operational or proprietary metadata. To mitigate the risk of model poisoning and data leakage during the global synchronization phase, the framework must integrate an advanced cryptographic layer that couples additively homomorphic encryption (such as the Paillier cryptosystem) with a localized differential privacy (LDP) protocol. Local model weight parameters are systematically masked with

calibrated Gaussian noise and encrypted before transmission to a decentralized, blockchain-backed ledger, which executes secure aggregation algorithms in the encrypted domain to guarantee absolute data privacy.

In addition to technical factors, human and organizational aspects, operator behavior, socio-technical risks, and supply chain dependencies are also to be studied by future research, as they significantly affect the resilience of industrial cybersecurity [1]. Lastly, standardized performance measures, including latency measures, false positive rate, and computational cost, will allow for comparability between studies and identify those methods that can be widely used in industry.

Beyond this, there are three additional paradigms becoming increasingly important to ICS security. For one thing, digital twin frameworks have become process-aware IDS platforms, creating a virtual copy of the physical ICS environment to detect stealthy attacks that change the condition of processes but do not generate conventional network anomaly alerts. Recent research has shown that detection F1-scores are 96.3% against FDI and DoS attacks for a digital twin-based SCADA environment with false positive rates of less than 2.5% [50]. Lin et al. then furthered this idea by implementing a hybrid deep learning model that was coupled with a SCADA digital twin to detect anomalies with low latency in real time [51].

Secondly, IDS frameworks built on blockchain technology overcome the shortcomings of centralized detection in distributed IIoT settings, providing tamper-proof logging and sharing of threat intelligence information between organizations [21]. Thirdly, by combining federated learning with 6G communication infrastructure, the security of IIoT can find very good application in the field of securing and enabling model training across geographically distributed industrial assets with small delay while preserving privacy for sensitive operational data [20].

Yet another interesting area of research is the use of large model agents within industrial cybersecurity architectures. New breakthroughs within agent-based artificial intelligence indicate that independent agents with abilities to reason and collaborate can contribute positively towards the development of intrusion detection and prevention systems. The future of ICS security architecture may include using intelligent agents, which will be able to monitor assets, correlate security incidents, manage the response to threats, and adapt to threats in real time. Additionally, concepts like embodied intelligence and the Internet of Agents may facilitate the creation of even more robust cybersecurity frameworks capable of working within heterogeneous IIoT environments. However, there are still problems associated with trustworthiness, transparency, and secure coordination of the agents, which should be considered [16–18].

6. Conclusions

From a quantitative perspective, it is found that: (1) hybrid deep learning models (Transformer + LSTM, MODLSTM) were found in 29% of the studies reviewed and consistently better than the single-model approach, with detection rates above 99% achieved; (2) Modbus/TCP was still the most widely used protocol in 51.6% of the studies, while MQTT and OPC UA are only found in 12.9% of the studies; (3) 67% of the studies used synthetic or custom datasets, reflecting the prevailing lack of real-world data; (4) only 9.7% of the studies considered federated learning to validate the models in a privacy-preserving way.

Research has strongly indicated that hybrid deep learning systems, especially hybrid systems that combine sequential-based and attention-based models, have higher detection rates of attacks like DoS, MITM, and data injection. In addition to this, the growing use of real-world testbeds and high-fidelity simulators like OPAL-RT is indicative of a growing

research culture that places importance on empirical reliability as well as algorithmic performance. Similarly, the incorporation of federated learning, edge/fog computing, and improvements to secure protocols implies an increased devotion to balancing the effectiveness of detection with the latency of the system and data confidentiality.

However, there are still major constraints, such as the lack of variety of real-world data, not enough focus on model scalability and interoperability, and excessive emphasis on detection accuracy without an equally strong emphasis on system recovery and resiliency. To solve these difficulties, it is necessary to transition to adaptive designs that combine detection and automated mitigation and response, especially for resource-constrained IIoT devices and legacy systems.

It is also recognized that IEEE Xplore was part of the search process, but this process may have excluded some important *IEEE Transactions* publications due to the stringent inclusion criteria. Future reviews should use a separate IEEE-focused search pass to guarantee a full review of the most rigorous work in this area.

Overall, the perceived literature shows that intelligent analytics, distributed architecture, and resilient operations are the future of industrial cybersecurity. Through harmonizing technological innovation and human and organizational consciousness, scientists and practitioners can get nearer to realizing safe, trustworthy, and future-proof industrial ecosystems that are able to endure the sophisticated cyber threats of the Industry 4.0 period.

Future research should focus on the following directions: (1) creation of standardized, large-scale, publicly available industrial datasets that contain data of various attack types and protocols other than Modbus/TCP; (2) design of lightweight and adaptive ML/DL models suitable for resource-constrained IIoT devices and legacy controllers; (3) extension of federated learning frameworks with differential privacy and secure aggregation to enable cross-organizational threat intelligence sharing; (4) formal verification and cryptographically strengthening of emerging protocols like OPC UA and IEC 61850; (5) construction of integrated detect–respond–recover frameworks involving SDN-based isolation and adaptive control mechanisms to ensure operational continuity during active cyber incidents.

Author Contributions: Conceptualization, K.A.-T., A.S. and S.A.-H.B.; methodology (PRISMA protocol design), K.A.-T., A.S. and S.A.-H.B.; validation (study screening and quality assessment), K.A.-T., A.S. and S.A.-H.B.; formal analysis (data extraction and synthesis), K.A.-T.; investigation (literature search), K.A.-T.; resources, K.A.-T.; data curation, K.A.-T.; writing—original draft preparation, K.A.-T.; writing—review, editing, and comprehensive proofreading, A.S. and S.A.-H.B.; supervision, A.S. and S.A.-H.B.; project administration, K.A.-T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Knowles, W.; Prince, D.; Hutchison, D.; Disso, J.F.P.; Jones, K. A survey of cyber security management in industrial control systems. *Int. J. Crit. Infrastruct. Prot.* **2015**, *9*, 52–80. [[CrossRef](#)]
2. Rahman, A.; Mustafa, G.; Khan, A.Q.; Abid, M.; Durad, M.H. Launch of denial of service attacks on the Modbus/TCP protocol and development of its protection mechanisms. *Int. J. Crit. Infrastruct. Prot.* **2022**, *39*, 100568. [[CrossRef](#)]
3. Zhang, H.; Min, Y.D.; Liu, S.; Tong, H.; Li, Y. MODLSTM: A method to recognize DoS attacks on Modbus/TCP. In *Proceedings of the IEEE International Performance Computing and Communications Conference (IPCCC)*; IEEE: New York, NY, USA, 2022. [[CrossRef](#)]

4. Leao, B.P.; Vempati, J.; Muenz, U.; Shekhar, S.; Pandey, A.; Hingos, D.; Bhela, S.; Wang, J.; Bilby, C. Machine learning-based false data injection attack detection and localization in power grids. In Proceedings of the 2022 IEEE Conference on Communications and Network Security (CNS), Austin, TX, USA, 3–5 October 2022; (No. DOE-SIEMENS-08769-CNS); Siemens Corporation, Technology.
5. Mohammed, A.S.; Saxena, N.; Rana, O. Wheels on the Modbus: Attacking Modbus/TCP communications. In *Proceedings of the ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 288–289. [\[CrossRef\]](#)
6. Gao, J.; Gan, L.; Buschendorf, F.; Zhang, L.; Liu, H.; Li, P.; Dong, X.; Lu, T. Omni SCADA intrusion detection using deep learning algorithms. *IEEE Internet Things J.* **2021**, *8*, 951–961.
7. Gueye, T.; Wang, Y.; Rehman, M.; Mushtaq, R.T.; Zahoor, S. A novel method to detect cyber-attacks in IoT/IIoT devices on the modbus protocol using deep learning. *Clust. Comput.* **2023**, *26*, 2947–2973.
8. Hossain, D.; Ochiai, H.; Arisawa, T.; Kadobayashi, Y. Smart meter modbus rs-485 spoofing attack detection by lstm deep learning approach. In *2022 9th Swiss Conference on Data Science (SDS)*; IEEE: New York, NY, USA, 2022; pp. 47–52.
9. Zhang, H.; Min, Y.; Liu, S.; Tong, H.; Li, Y.; Lv, Z. Improve the security of industrial control system: A fine-grained classification method for DoS attacks on Modbus/TCP. *Mob. Netw. Appl.* **2023**, *28*, 839–852.
10. Thiagarajan, K.; Hammad, I. Anomaly detection in air-gapped industrial control systems of nuclear power plants. In *Proceedings of the Cyber Awareness and Research Symposium (CARS)*; IEEE: New York, NY, USA, 2024; pp. 979–984. [\[CrossRef\]](#)
11. Chochtoula, D.; Ilias, A.; Stamatiou, Y.C.; Makris, C. Integrating elliptic curve cryptography with the Modbus TCP SCADA communication protocol. *Future Internet* **2022**, *14*, 232. [\[CrossRef\]](#)
12. Jadidi, Z.; Foo, E.; Hussain, M.; Fidge, C. Automated detection-in-depth in industrial control systems. *Int. J. Adv. Manuf. Technol.* **2022**, *118*, 2467–2479.
13. Dao, Q.-T.; Nguyen, L.-T.; Ha, T.-K.; Nguyen, V.-H.; Nguyen, T.-A. Investigation of secure communication of Modbus TCP/IP protocol: Siemens S7 PLC series case study. *Appl. Syst. Innov.* **2025**, *8*, 65. [\[CrossRef\]](#)
14. Purohit, S.; Govindarasu, M.; Blakely, B. FL-ADS: Federated learning anomaly detection system for distributed energy resource networks. *IET Cyber-Phys. Syst. Theory Appl.* **2025**, *10*, e70001.
15. Kotsiopoulos, T.; Radoglou-Grammatikis, P.; Lekka, Z.; Mladenov, V.; Sarigiannidis, P. Defending industrial internet of things against Modbus/TCP threats: A combined AI-based detection and SDN-based mitigation solution. *Int. J. Inf. Secur.* **2025**, *24*, 157. [\[CrossRef\]](#)
16. Wang, Y.; Pan, Y.; Su, Z.; Deng, Y.; Zhao, Q.; Du, L.; Luan, T.H.; Kang, J.; Niyato, D. Large model-based agents: State-of-the-art, cooperation paradigms, security and privacy, and future trends. *IEEE Commun. Surv. Tutor.* **2026**, *28*, 1906–1949. [\[CrossRef\]](#)
17. Wang, Y.; Niu, X.; Ba, J.; Su, Z.; Du, L.; Luan, T.H.; Kang, J.; Niyato, D. Navigating embodied intelligence: Enabling technologies, security and privacy, and emerging trends. *IEEE Internet Things J.* **2026**, *13*, 10317–10338. [\[CrossRef\]](#)
18. Wang, Y.; Guo, S.; Pan, Y.; Su, Z.; Chen, F.; Luan, T.H.; Li, P.; Kang, J.; Niyato, D. Internet of Agents: Fundamentals, Applications, and Challenges. *IEEE Trans. Cogn. Commun. Netw.* **2025**, *12*, 4476–4501. [\[CrossRef\]](#)
19. Creswell, J.W.; Creswell, J.D. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*; Sage Publications: Thousand Oaks, CA, USA, 2017.
20. Sayghe, A. Digital Twin-Driven Intrusion Detection for Industrial SCADA: A Cyber-Physical Case Study. *Sensors* **2025**, *25*, 4963. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Zhang, D.W.; Liu, G.P. Observer-based HOFA predictive cooperative control for networked multi-agent systems under time-variant communication constraints. *ISA Trans.* **2024**, *147*, 554–566. [\[PubMed\]](#)
22. Afenu, D.S.; Asiri, M.; Saxena, N. Industrial control systems security validation based on MITRE adversarial tactics, techniques, and common knowledge framework. *Electronics* **2024**, *13*, 917. [\[CrossRef\]](#)
23. Bajwa, A.; Tonoy, A.A.R.; Rana, S.; Ahmed, I. *Cybersecurity in Industrial Control Systems: A Systematic Literature Review on AI-Based threat Detection for SCADA and IOT Networks*; SSRN: Amsterdam, The Netherlands, 2025.
24. Mohammed, A.S.; Anthi, E.; Rana, O.; Saxena, N.; Burnap, P. Detection and mitigation of field flooding attacks on oil and gas critical infrastructure communication. *Comput. Secur.* **2023**, *124*, 103007. [\[CrossRef\]](#)
25. Poudel, S.; Ni, Z.; Malla, N. Real-time cyber physical system testbed for power system security and control. *Int. J. Electr. Power Energy Syst.* **2017**, *90*, 124–133. [\[CrossRef\]](#)
26. IEC 61850; Communication Networks and Systems for Power Utility Automation. International Electrotechnical Commission (IEC): Geneva, Switzerland, 2021.
27. Che, X.; Geng, Y.; Zhang, G.; Wang, M. Fuzzing technology based on information theory for industrial proprietary protocol. *Electronics* **2023**, *12*, 3041. [\[CrossRef\]](#)
28. Jhan, J.Y.; Sun, H.M. Enhancing Modbus TCP Protocol Security with eBPF Technology. *arXiv* **2023**, arXiv:2312.05665.
29. Aparcana-Tasayco, A.J.; Deng, X.; Park, J.H. A systematic review of anomaly detection in IoT security: Towards quantum machine learning approach. *EPJ Quantum Technol.* **2025**, *12*, 112. [\[CrossRef\]](#)

30. Varol, M.; İskefiyeli, M. An intrusion detection system for critical infrastructures: Modbus approach. *Eng. Appl. Artif. Intell.* **2025**, *162*, 112410. [\[CrossRef\]](#)
31. Liu, M.; Xia, Y.; Zhao, H.; Guo, L.; Shi, Z.; Zhu, H. Federated learning technologies for 6G industrial internet of things: From requirements, vision to challenges, opportunities. *J. Electron. Inf. Technol.* **2024**, *46*, 4335–4353. [\[CrossRef\]](#)
32. Slunjski, M.; Sumina, D.; Groš, S.; Erceg, I. Off-the-shelf solutions as potential cyber threats to industrial environments and simple-to-implement protection methodology. *IEEE Access* **2022**, *10*, 114735–114748.
33. Martins, T.; Oliveira, S.V.G. Enhanced Modbus/TCP security protocol. *Sensors* **2022**, *22*, 8024. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Seng, S.; Garcia-alfaro, J.; Laarouci, Y. Implementation of a stateful network protocol intrusion detection systems. In *SECRYPT 2022: 19th International Conference on Security and Cryptography*; SCITEPRESS-Science and Technology Publications: Setúbal, Portugal, 2022; pp. 398–405.
35. Sindhwad, P.; Kazi, F. Exploiting control device vulnerabilities. In *Proceedings of the Conference of Open Innovations Association (FRUCT)*; IEEE: New York, NY, USA, 2022; pp. 270–279.
36. Nguyen, H.N.; Koo, J. Enhancing SCADA Security Using Generative Adversarial Network. *J. Cybersecur. Priv.* **2025**, *5*, 73. [\[CrossRef\]](#)
37. Bahadoripour, S.; MacDonald, E.; Karimipour, H. A Deep multi-modal cyber-attack detection in industrial control systems. In *Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*; IEEE: New York, NY, USA, 2023.
38. Thulasiraman, P.; Hackett, M.; Musgrave, P.; Edmond, A.; Seville, J. Anomaly Detection in a Smart Microgrid System Using Cyber-Analytics: A Case Study. *Energies* **2023**, *16*, 7151. [\[CrossRef\]](#)
39. Ghosh, T.; Bagui, S.; Bagui, S.; Kadzis, M.; Bare, J. Anomaly detection for modbus over TCP in control systems using entropy and classification-based analysis. *J. Cybersecur. Priv.* **2023**, *3*, 895–913. [\[CrossRef\]](#)
40. Mekala, S.H.; Baig, Z.; Anwar, A.; Syed, N. DoS attacks, human factors, and evidence extraction for the industrial internet of things (IIoT) paradigm. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*; IEEE: New York, NY, USA, 2023; pp. 32–39.
41. Elrawy, M.F.; Hadjidemetriou, L.; Laoudias, C.; Michael, M.K. Detecting and classifying man-in-the-middle attacks in the private area network of smart grids. *Sustain. Energy Grids Netw.* **2023**, *36*, 101167. [\[CrossRef\]](#)
42. Banik, S.; Manicavasagam, R.; Banik, T.; Banik, S. Simulation and analysis of cyber-attack on modbus protocol for smart grids in virtual environment. In *Science and Information Conference*; Springer Nature: Cham, Switzerland, 2024; pp. 384–401.
43. Raja, D.J.S.; Hemavathi, N.; Sriranjani, R.; Arulmozhi, P. Mitigation of Man-in-the-Middle Attack in Advanced Metering Infrastructure Through Behavioral Biometrics-Based Elliptic Curve Cryptography. *IEEE Trans. Green Commun. Netw.* **2024**, *9*, 778–788.
44. Sheng, C.; Yao, Y.; Zhao, L.; Zeng, P.; Zhao, J. Scanner-hunter: An effective ICS scanning group identification system. *IEEE Trans. Inf. Forensics Secur.* **2024**, *19*, 3077–3092. [\[CrossRef\]](#)
45. Upadhyay, D.; Ghosh, S.; Ohno, H.; Zaman, M.; Sampalli, S. Securing industrial control systems: Developing a SCADA/IoT test bench and evaluating lightweight cipher performance on hardware simulator. *Int. J. Crit. Infrastruct. Prot.* **2024**, *47*, 100705.
46. Srinivasan, H.; Karimi, M. Threat-based security controls for ICS. *arXiv* **2025**, arXiv:2501.13268.
47. Mondal, S.; Prudhvi, B.; Khare, P.; Reddy, M.J.B. Real-time analysis of cyber attacks in SCADA based power system. *Aust. J. Electr. Electron. Eng.* **2025**, 1–25. [\[CrossRef\]](#)
48. Sarkar, S.; Teo, Y.M.; Chang, E.-C. A cybersecurity assessment framework for virtual operational technology in power system automation. *Simul. Model. Pract. Theory* **2022**, *117*, 102453. [\[CrossRef\]](#)
49. Sinha, A.; Patel, S.S.; Kumar, A.; Vyas, O.P. *Exploiting Vulnerabilities in the SCADA Modbus Protocol: An ICT-Reliant Perspective*; Springer: Cham, Switzerland, 2022; pp. 15–34. [\[CrossRef\]](#)
50. Ali, S.; Li, Q.; Yousafzai, A. Blockchain and federated learning-based intrusion detection approaches for edge-enabled industrial IoT networks: A survey. *Ad Hoc Netw.* **2024**, *152*, 103320.
51. Nassif, A.B.; Talib, M.A.; Nasir, Q.; Dakalbab, F.M. Machine learning for anomaly detection: A systematic review. *IEEE Access* **2021**, *9*, 78658–78700. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.