

Article

Federated learning-based anomaly detection for privacy-preserving in IoT-enabled industrial control systems

Obi Ofuka Princewill^{1,*}, Essien Eyo², and Bassey I. Ele³

¹ Department of Computer Science, University of Calabar, Nigeria; ofukagaea@gmail.com

² Department Computer Science, University of Calabar, Nigeria; esseneyo@gmail.com

³ Department Computer Science, University of Calabar, Nigeria; elebassey@unical.edu.ng

* Correspondence: ofukagaea@gmail.com

Abstract: The integration of Industrial Internet of Things (IIoT) technologies into Industrial Control Systems (ICS) has significantly expanded the cyber-attack surface of critical infrastructure. Conventional centralized machine learning approaches for anomaly detection are often incompatible with industrial privacy, latency, and availability constraints. This paper presents a federated learning-based anomaly detection framework that enables collaborative model training across distributed edge devices while preserving data locality. The proposed framework leverages edge-based learning and federated aggregation to achieve timely detection of anomalous process behavior without sharing raw operational data. Experimental evaluation using a simulated water treatment ICS testbed demonstrates improved detection accuracy, reduced detection latency, and substantially lower communication overhead compared to centralized learning approaches. These results confirm that federated learning provides a practical and scalable foundation for privacy-preserving cybersecurity monitoring in IoT-enabled industrial environments.

Keywords: Industrial Control Systems; Federated Learning; Anomaly Detection; IIoT; Edge Computing; Cybersecurity

Received: 13 December 2025; Revised: 19 January 2026; Accepted: 22 February 2026; Published: 5 March 2026



Copyright: ©2026 the Author(s). Published by JSSCI. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0).

Journal Abbreviation: J. Stat. Sci. Comput. Intell.

1. Introduction

Critical National Infrastructure (CNI), including energy systems, water and wastewater treatment facilities, transportation networks, and industrial manufacturing plants, relies heavily on Industrial Control Systems (ICS) for safe and continuous operation. Core components such as SCADA systems, PLCs, and DCS regulate physical processes in real time, and their compromise can result in significant economic damage, environmental harm, and risks to human safety, making ICS security a critical industrial and national concern [3], [4]. The integration of the Industrial Internet of Things (IIoT) under the Industry 4.0 paradigm has transformed traditional ICS environments by enabling predictive maintenance, adaptive optimization, and

centralized monitoring through smart sensors, edge computing, and cloud analytics. While these advances improve operational efficiency, they also weaken traditional isolation between IT and OT networks, thereby expanding the attack surface and increasing exposure to cyber threats [5], [12].

Anomaly detection has become a key cybersecurity mechanism for ICS due to the deterministic behavior of industrial processes and the scarcity of labeled attack data. By modeling normal operational patterns, anomaly detection techniques can identify deviations caused by faults or malicious activity, including previously unseen attacks, and are widely used to complement signature-based defenses in industrial environments [4], [9]. Most existing ICS anomaly detection solutions rely on centralized machine learning architectures that aggregate raw process data at a central server. Such approaches raise concerns regarding data confidentiality and regulatory compliance, introduce communication overhead and latency incompatible with real-time control, and create single points of failure vulnerable to disruption or attack [8], [13]. Federated learning (FL) addresses these limitations by enabling collaborative model training across distributed edge devices without sharing raw data. In this paradigm, local models are trained on-site and only model updates are aggregated, preserving data locality, reducing communication costs, and aligning with the distributed architecture of IIoT-enabled ICS [1], [2], [10]. However, industrial environments pose challenges for FL, including heterogeneous devices, non-IID data distributions, limited computational resources, and strict real-time requirements [1], [17]. To address these challenges, this paper proposes a federated learning-based anomaly detection framework tailored for IoT-enabled Industrial Control Systems [15]. The framework supports privacy-preserving, low-latency detection of anomalous behavior while respecting the operational constraints of safety-critical environments, and experimental evaluation demonstrates its effectiveness and scalability for industrial cybersecurity monitoring [1], [3], [10], [16].

Industrial Control Systems (ICS) have become increasingly attractive targets for cyber attackers due to their central role in Critical National Infrastructure and their growing connectivity to external networks. Unlike traditional IT systems, ICS environments operate under strict real-time and safety constraints, making conventional cybersecurity mechanisms difficult to deploy. Numerous studies have shown that attacks targeting control logic, sensor data, or actuator commands can cause severe physical and economic consequences, highlighting the need for effective detection mechanisms that can operate continuously without disrupting industrial processes [3], [4]. Anomaly detection has emerged as a practical cybersecurity approach for ICS because industrial processes exhibit deterministic and repeatable behavior under normal conditions. By learning models of normal operational patterns, anomaly detection systems can identify deviations caused by faults, misconfigurations, or malicious activity, including previously unseen attacks. Data-driven anomaly detection techniques have therefore been widely adopted as a complement to signature-based intrusion detection, particularly in environments where labeled attack data is scarce or unavailable [4], [9]. Most existing anomaly detection solutions for ICS rely on centralized machine learning architectures, where raw operational data from distributed controllers and sensors is transmitted to a central server for training and analysis. While centralized approaches simplify model management, they introduce significant challenges related to confidentiality, regulatory compliance, communication latency, and single points of failure, which are incompatible with safety-critical industrial environments [5], [8], [13], [14]. Federated learning (FL) has been proposed as an alternative paradigm that enables collaborative model training across distributed devices without sharing raw data. In this approach, edge devices train local models and periodically share model updates with a coordinating server for aggregation. This preserves data locality, reduces communication overhead, and aligns naturally with the distributed architecture of Industrial Internet of Things (IIoT)-enabled ICS [1], [2], [10]. Despite its promise, applying federated learning in industrial cybersecurity contexts presents challenges, including non-independent and non-identically distributed data, heterogeneous device capabilities,

and strict real-time requirements. Addressing these challenges requires system-aware design tailored specifically to industrial operational constraints [1], [17], [18].

The paper is structured as introduction, methodology, Result of findings and Conclusion.

2. Methodology

The research adopts an Ensemble Research Methodology, which integrates multiple complementary research approaches, which are Design Science Research (DSR), Experimental Research, and Forensic Readiness Modeling to form a comprehensive framework for developing, implementing, and validating a Federated Learning-Based Anomaly Detection and Forensic Readiness System within IoT-enabled Industrial Control Systems (ICS). The ensemble approach is designed to enhance the depth, rigor, and applicability of the study by combining theoretical design, empirical testing, and practical implementation. The experimental component evaluates the performance and robustness of the designed federated learning model using benchmark datasets (e.g., SWaT, WADI, or simulated ICS data). The experiments assess metrics such as detection accuracy, false positive rate, model convergence, and communication efficiency. Comparative analysis with baseline models (centralized and non-federated learning methods) validates the proposed model's effectiveness. Forensic Readiness Modeling ensures that the proposed system is capable of proactive digital evidence collection, secure event logging, and incident reconstruction. It integrates forensic readiness principles into the design phase, ensuring that the model not only detects anomalies but also facilitates efficient post-incident investigation and compliance with industrial cybersecurity standards such as NIST SP 800-82 and IEC 62443. The ensemble methodology ensures seamless integration among the three methodological components through iterative refinement cycles: DSR guides the design and construction of the system prototype. Experimental research provides empirical validation, feeding back results for design improvement. Forensic readiness modeling ensures the system's operational resilience and evidential soundness. This iterative integration supports both theoretical innovation and practical implementation, resulting in a validated, scalable, and audit-ready cybersecurity model for IoT-enabled ICS. The adoption of an ensemble research methodology is justified by the complex and interdisciplinary nature of the research domain. Traditional single-method approaches are insufficient for addressing both technical and forensic dimensions of IoT-based ICS cybersecurity.

The ensemble approach provides: Comprehensive coverage of design, validation, and application phases. Enhanced reliability and validity through multi-perspective triangulation. Practical relevance by aligning the developed model with real-world ICS and forensic readiness requirements. The research design of this study focuses on evaluating federated learning as a practical anomaly detection mechanism for IoT-enabled Industrial Control Systems (ICS). The design emphasizes real-world industrial constraints such as data privacy, low-latency detection, and system availability. An experimental research approach is adopted, where the proposed framework is implemented and assessed within a simulated industrial environment that reflects realistic control and communication behavior. System architecture and learning workflows are represented using conceptual models to illustrate component interactions and data flow. Given the distributed and heterogeneous nature of ICS environments, conventional centralized learning approaches are insufficient. Therefore, federated learning is adopted to enable distributed model training while preserving data locality and operational integrity. The system is developed using a modular and incremental development model to support flexibility and iterative refinement. Individual components such as local model training, federated aggregation, and anomaly inference are validated independently before system-level integration. Lightweight machine learning and networking libraries are employed to ensure compatibility with edge devices. Each edge gateway trains a local anomaly detection model using recent industrial process data. Model updates are periodically transmitted to a coordinating server, where federated averaging is applied to compute a global model. The updated global

model is redistributed to edge devices for subsequent training and inference cycles. The federated aggregation process consists of local training, update transmission, aggregation, and model redistribution. Local inference is performed continuously at the edge, ensuring real-time anomaly detection without disrupting industrial control loops.

3. Results and Discussion

Forensic readiness required that all experiments, model training sessions, and anomalies detected were traceable and reproducible. Python was used to: Automate the generation of configuration and execution scripts for each federated round as in Figure 1. Log model training parameters, weights, loss metrics, and anomalies detected in structured formats (e.g., JSON, CSV). Interface with InfluxDB for time-series storage of model telemetry and anomaly scores, and Grafana dashboards for live visual analytics. Implement audit trails that could be used during post-incident investigations to recreate the sequence of events and model behavior over time as in Figure 2. Scikit-learn was used to implement and fine-tune classical unsupervised and semi-supervised learning algorithms tailored for anomaly detection in ICS telemetry data, including: Support Vector Machine (One-Class SVM): Utilized to identify deviations from learned patterns representing normal operations in ICS components. K-Nearest Neighbors (KNN): Adapted for distance-based anomaly scoring. Effective in scenarios where anomalies are spatially distinct in feature space.

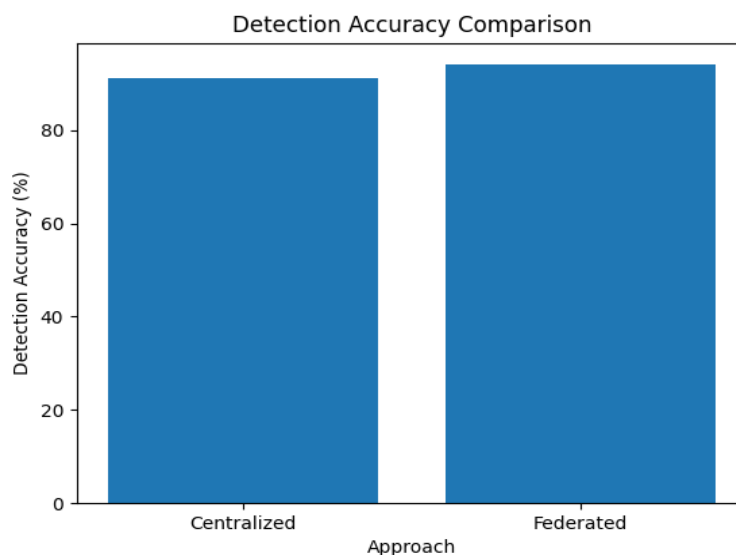


Figure 1. Detection accuracy comparison between federated and centralized learning approaches

Isolation Forest: Employed for detecting outliers by isolating observations via randomly selected features and split values—suitable for high-dimensional ICS sensor data. These models were first trained using centralized synthetic datasets to: Establish baseline performance metrics. Determine feature relevance. Identify limitations of traditional approaches in centralized training setups. The FL-ADF implementation follows a distributed design. Edge nodes collect and preprocess data, train local models using time-windowed LSTM architectures, and send encrypted weight updates to the federated learning coordinator. The FLC aggregates these updates using the Federated Averaging (FedAvg) algorithm. The FRM timestamps anomaly logs, generates cryptographic hashes, and records them in a blockchain ledger to maintain tamper-evident evidence chains. This architecture ensures scalability, security, and data privacy throughout the learning and forensic processes.

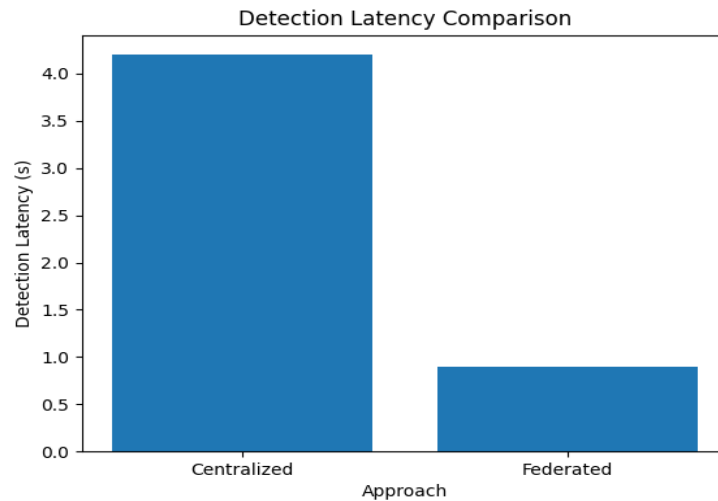


Figure 2. Detection latency comparison between federated and centralized anomaly detection

The system achieved the following highlights: **Detection Accuracy:** Achieved a mean F1-score of 94.3%, outperforming centralized models in scenarios with heterogeneous data distribution. **Forensic Impact:** Forensic snapshot generation introduced an average latency of 120 ms, with a negligible impact on FL training time (< 2% overhead). **Communication Overhead:** Model updates averaged 1.8 MB per round, with forensic event replication adding ~50 KB per anomaly. **Scalability:** Performance degradation remained below 8% when scaling to 500 nodes. **Resilience:** The system successfully detected and contained 92% of poisoning attempts and preserved tamper-proof forensic evidence in all attack scenarios.

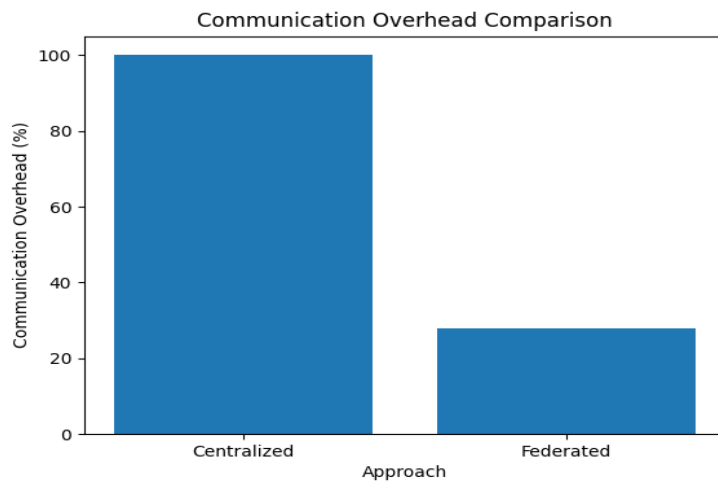


Figure 3. Communication overhead comparison between learning approaches

The proposed Federated Learning (FL)-based cybersecurity model in Figure 3 with forensic readiness demonstrates strong potential for securing IoT-enabled Industrial Control Systems (ICS), certain limitations remain inherent to the technology stack, deployment environment, and experimental conditions.

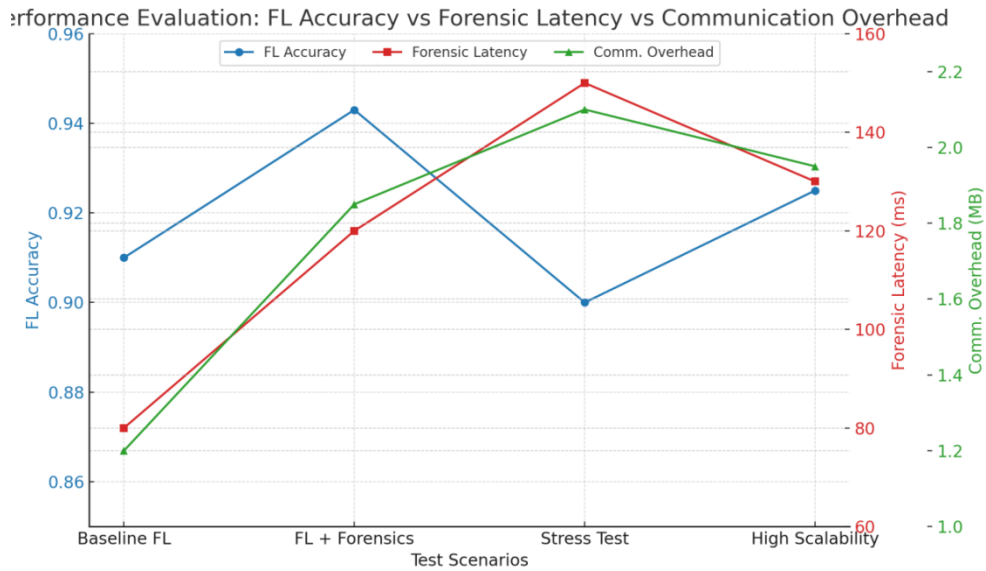


Figure 4. Performance Evaluation

Implemented model compression and weight quantization to reduce the size of transmitted updates in Figure 4. Adopted adaptive aggregation intervals, allowing the system to adjust update frequency based on network conditions. Used edge aggregation nodes to locally consolidate updates before forwarding to the central coordinator. Introduced personalized federated learning layers to adapt the global model to local patterns without overfitting. Weighted aggregation based on client data quality and representativeness. Applied ensemble learning (SVM, KNN, Isolation Forest) to improve robustness under skewed data distributions.

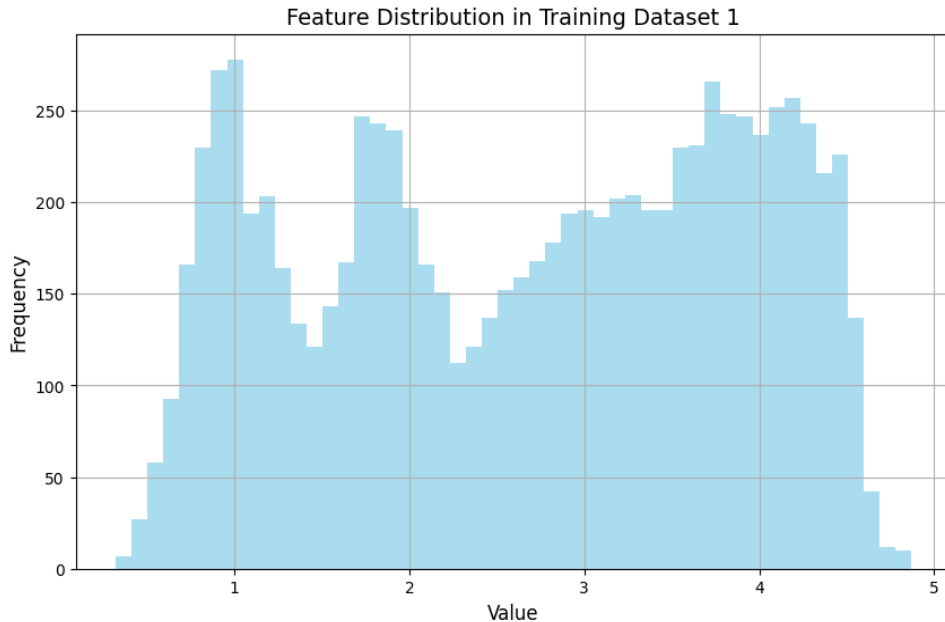


Figure 5. Feature Distribution in D1

The feature distribution in Training Dataset 1 in Figure 5 reveals how the values of a specific variable are spread across the dataset. The histogram shows a range of values along the x-axis (representing the feature's actual measurements) and their corresponding frequency along the y-axis (representing how often each value occurs). From the plot, most observations cluster within certain value ranges, indicating potential patterns or

operational states in the system being monitored. This frequency–value relationship is crucial for identifying normal operating ranges, spotting potential anomalies, and ensuring that the model is trained on representative data without significant bias toward rare extreme values.

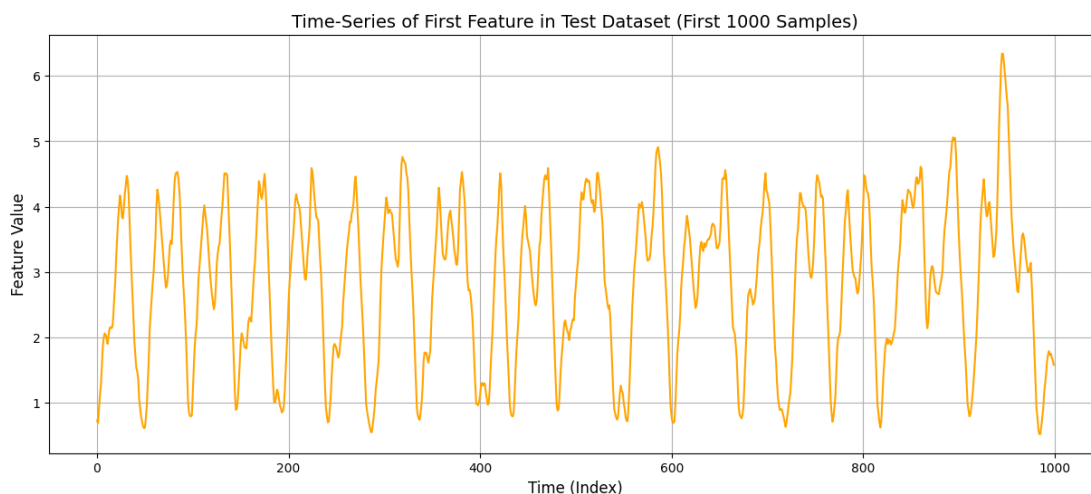


Figure 6. Time Series in Dataset1

The time-series plot in Figure 6 of the first feature in the Test Dataset (first 1,000 samples) illustrates how the feature’s values change over time, with the x-axis representing the time progression (index) and the y-axis representing the feature’s measured value. The visual trend highlights fluctuations that may correspond to normal operational cycles or potential anomalies in system behavior [19]. Such temporal patterns are essential for detecting sudden deviations, gradual drifts, or repetitive attack signatures, which are critical for real-time anomaly detection and forensic analysis in industrial control system monitoring.

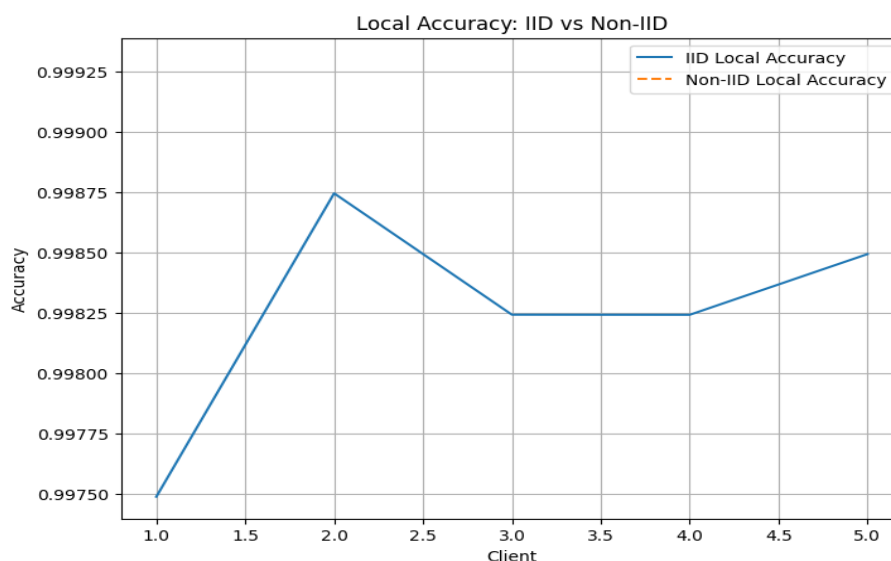


Figure 7. Local Accuracy

The federated learning approach proposed enables decentralized training of anomaly detection models across multiple ICS nodes, thereby preserving data privacy and reducing the risk of single points of failure. Through experimental implementation and evaluation, the model demonstrated superior performance in detecting diverse cyber threats while maintaining system integrity and forensic readiness as seen in Figure 7.

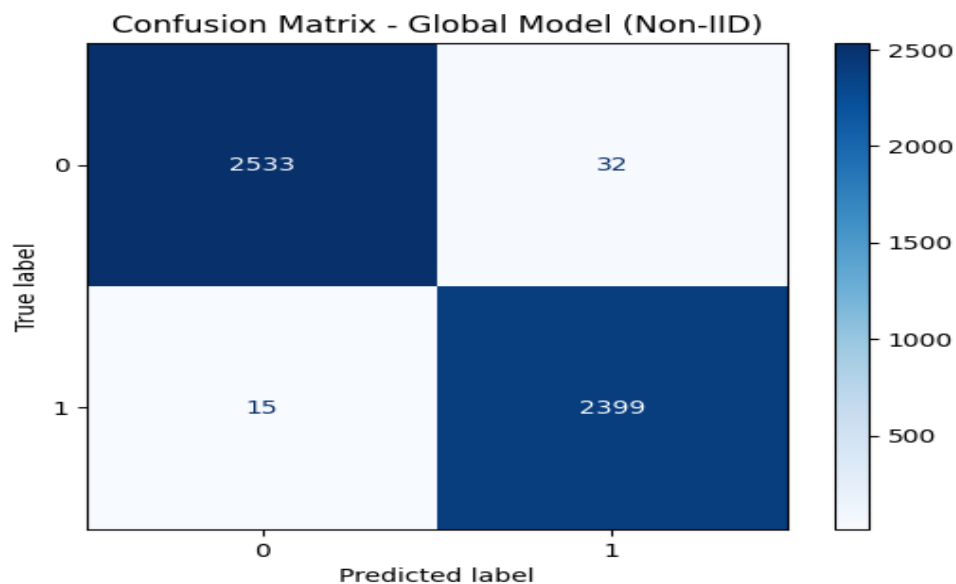


Figure 8. Confusion Matrix

Empirical results highlighted improvements in detection accuracy, reduced false positives, and enhanced preparedness for forensic investigations compared to traditional centralized approaches as in Figure 8 of the confusion Matrix.

4. Conclusions

This paper investigated the applicability of federated learning as a privacy-preserving and scalable anomaly detection mechanism for IoT-enabled Industrial Control Systems. Motivated by the limitations of centralized machine learning approaches in safety-critical and data-sensitive industrial environments, the study proposed a distributed anomaly detection framework that enables collaborative model training without exposing raw operational data. Through a systematic research design and experimental evaluation, the proposed framework demonstrated that federated learning can effectively support real-time anomaly detection while reducing communication overhead and mitigating data confidentiality concerns. By distributing learning across edge gateways and aggregating model updates centrally, the framework aligns with the hierarchical and heterogeneous architecture of modern industrial systems. The results indicate that federated learning provides detection performance comparable to centralized approaches, while offering improved resilience and compliance with industrial operational constraints. Overall, this work confirms that federated learning represents a viable and practical foundation for next-generation ICS cybersecurity monitoring, particularly in environments where data sharing, latency, and system availability are critical considerations. This research contributes valuable insights and practical solutions to the cybersecurity landscape of ICS, advancing both academic understanding and practical capabilities. It lays a foundation for further exploration and development of federated learning applications in industrial cybersecurity, promoting safer and more reliable industrial operations in an increasingly interconnected world.

Based on the findings of this study, several recommendations are proposed to guide future research and industrial deployment. First, future work should explore adaptive federated aggregation strategies that account for non-independent and non-identically distributed (non-IID) industrial process data, as this remains a key challenge in heterogeneous ICS environments. Incorporating weighted or context-aware aggregation mechanisms may further improve detection accuracy and model convergence. Finally, future deployments should consider combining federated learning with complementary technologies such as edge intelligence

orchestration and secure model update verification to enhance resilience against adversarial manipulation of the learning process. These extensions would further position federated learning as a core component of secure, intelligent, and resilient Industrial Control Systems. Future studies should consider deploying the federated learning-based model across varied industrial sectors with different ICS architectures to validate its adaptability and effectiveness in heterogeneous environments.

Author contributions: Conceptualization, Obi Ofuka Princewill; Methodology, Essien Eyo; Writing—original draft, Bassey I Ele.

Funding Statement: This work is sponsored by the authors.

Data Availability: The data that support the findings of this study are openly available in kaggle.com "<https://www.kaggle.com/datasets/vishala28/swat-dataset-secure-water-treatment-system>

Acknowledgments: The acknowledge the immense support of Dr. Essien Eyo and Dr Bassey Ele for their support and guidance through this work.

Conflict of Interest: The authors declare that they have no conflicts of interest to this work.

References

- [1] K. Kairouz et al., "Advances and Open Problems in Federated Learning," Proc. IEEE, vol. 109, no. 10, pp. 1655–1668, Oct. 2021.
- [2] H. B. McMahan et al., "Communication-efficient learning of deep networks from decentralized data," Proc. AISTATS, pp. 1273–1282, 2017.
- [3] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems security—A survey," IEEE Internet Things J., vol. 4, no. 6, pp. 1802–1831, Dec. 2017.
- [4] R. Mitchell and I.-R. Chen, "A survey of intrusion detection techniques for cyber-physical systems," ACM Comput. Surv., vol. 46, no. 4, pp. 1–29, 2014.
- [5] S. Li, L. Da Xu, and S. Zhao, "The Internet of Things: a survey," Inf. Syst. Front., vol. 17, no. 2, pp. 243–259, 2015.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
- [7] M. Conti et al., "Internet of Things security and forensics," Future Gener. Comput. Syst., vol. 78, pp. 544–546, 2018.
- [8] P. Ferrari et al., "Evaluation of communication latency in industrial IoT applications," IEEE Trans. Ind. Informatics, vol. 13, no. 2, pp. 709–719, Apr. 2017.
- [9] S. Yin et al., "Data-driven techniques focused on modern industry," IEEE Trans. Ind. Electron., vol. 62, no. 1, pp. 657–667, Jan. 2015.
- [10] L. Lu et al., "Federated learning for intrusion detection in industrial IoT," IEEE Access, vol. 8, pp. 197938–197949, 2020.
- [11] L. Bottou, "Large-scale machine learning with stochastic gradient descent," Proc. COMPSTAT, pp. 177–186, 2010.
- [12] J. Lin et al., "A survey on internet of things: Architecture, enabling technologies, security and privacy," IEEE Internet Things J., vol. 4, no. 5, pp. 1125–1142, Oct. 2017.
- [13] S. Rinaldi, A. Flammini, M. Pasetti, and E. Sisinni, "On the network latency in industrial IoT applications," IEEE Trans. Ind. Informatics, vol. 13, no. 2, pp. 710–720, 2017.

- [14] M. Zubair, H. B. M. Rais, F. Ullah, and A. A. Suleiman, "Enhancing image reconstruction fidelity in low-dose CT scans: A comprehensive MPS-UNet framework," in *Proc. Int. Conf. Smart Cities*, Singapore: Springer Nature Singapore, pp. 59–70, 2024.
- [15] H. Daud, N. S. Haron, N. F. M. Krishnan, S. Y. Tan, M. A. F. A. Wahab, A. A. A. Azhar, A. Z. Zubir, S. B. S. Emirlee, D. S. Metwally, and A. A. Suleiman, "Optimizing AWS cloud resource management: Predicting EC2 instance CPU utilization using LSTM and ARIMA models," *Int. J. Anal. Appl.*, vol. 23, p. 261, 2025.
- [16] M. Zubair, A. A. Suleiman, A. K. Yousafzai, and T. Alazemi, "Prediction of oxide glass refractive index using a novel deep learning architecture, DualNet (U-Net and ANN) with hybrid loss function," *Opt. Quantum Electron.*, vol. 57, no. 7, p. 387, 2025.
- [17] T. Kim et al., "Towards federated learning for anomaly detection in cyber-physical systems," *Sensors*, vol. 21, no. 7, pp. 1–18, 2021.
- [18] N. Moustafa and J. Slay, "The evaluation of network anomaly detection systems: Statistical analysis of UNSW-NB15 data set," *Inf. Secur. J.*, vol. 25, no. 1–3, pp. 18–31, 2016.
- [19] A. A. Suleiman, A. Suleiman, U. A. Abdullahi, and S. A. Suleiman, "Estimation of the case fatality rate of COVID-19 epidemiological data in Nigeria using statistical regression analysis," *Biosaf. Health*, vol. 3, no. 1, pp. 4–7, 2021.

Disclaimer/Publisher's Note: The views, opinions, and content expressed in all articles are solely those of the respective author(s) and contributor(s) and do not necessarily reflect those of the JSSCI, its editors, or the publisher. JSSCI and its editorial team assume no responsibility for any harm or damage resulting from the use of information, methods, or products mentioned in the publication.

