

Federated Learning–Enabled Privacy Protection and Cybersecurity in Distributed IoMT Environments

Tahar Zioual*, Bouazza Abdelhammid†, Mohammed Hoceini*, Abdelhafidh Kacher*, Silarbi Ibrahim‡, Horri Mourad‡

* Department of Computer Science
University of Tiaret, Tiaret, Algeria

Email: zioualt@gmail.com, mohammed.hoceini@univ-tiaret.dz, abdelhafidh.kacher@univ-tiaret.dz

† Department of Computer Science
University of Msila, Msila, Algeria
Email: a.hamid223@gmail.com

‡ Department of Computer Science
University of Tiaret, Tiaret, Algeria

Email: silarbi.ibrahim.48@gmail.com, mouradhorri10@gmail.com

Abstract—The proliferation of the Internet of Medical Things (IoMT) has fundamentally transformed contemporary healthcare delivery by facilitating continuous patient monitoring, intelligent clinical decision support, and seamless integration of interconnected medical devices. Nevertheless, the inherently sensitive nature of medical data, coupled with the architectural heterogeneity of IoMT environments, introduces substantial challenges for data privacy, cybersecurity, and system reliability. Conventional centralised machine learning methodologies necessitate the aggregation of patient data within a unified repository, thereby elevating vulnerability to data breaches, contravening stringent medical privacy regulations, and constraining resilience against evolving cyber threats.

Federated Learning (FL) has emerged as a compelling paradigm to address these challenges by enabling multiple healthcare institutions and IoMT devices to collaboratively train machine learning models without exchanging raw patient data. This paper investigates an FL-based cybersecurity framework for distributed IoMT environments, with particular emphasis on privacy-preserving intrusion detection. Each participating client conducts local training on its proprietary network traffic data and transmits exclusively model updates to a coordinating server, thereby ensuring that sensitive medical information remains within institutional boundaries.

The proposed framework is evaluated utilising the WUSTL-EHMS-2020 IoMT dataset and a deep Multilayer Perceptron (MLP) classifier trained via Federated Averaging (FedAvg). Experimental results obtained under controlled conditions demonstrate that federated intrusion detection achieves robust detection performance whilst preserving data privacy.

Index Terms—Federated Learning, Internet of Medical Things, Data Privacy, Cybersecurity, Distributed Learning, Intrusion Detection, Privacy-Preserving Artificial Intelligence.

I. INTRODUCTION

The accelerating digital transformation of healthcare has established the Internet of Medical Things (IoMT) as a foundational infrastructure layer of modern clinical practice, interconnecting wearable sensors, implantable devices, bedside monitors, and clinical gateway systems with edge–cloud computing architectures [1], [2]. This interconnected ecosystem enables continuous physiological monitoring, real-time clinical

decision support, and personalised telemedicine services. At the same time, it generates large volumes of sensitive, high-frequency data whose compromise can directly affect patient safety, privacy, and institutional operational resilience [3], [6].

IoMT deployments are inherently heterogeneous, comprising devices that vary widely in computational capability, communication protocols, and operational context, and that are often deployed in resource-constrained and latency-sensitive environments [3]. Traditional perimeter defence mechanisms and centrally trained intrusion detection systems (IDS) face significant limitations in such settings, as they assume homogeneous network conditions, stable connectivity, and the feasibility of aggregating raw traffic data in a centralised repository. These assumptions conflict with prevailing data-protection regulations, such as GDPR and HIPAA, and amplify the potential impact of data breaches [4], [5]. Moreover, inter-institutional differences in clinical workflows, patient populations, and infrastructure configurations result in non-independent and identically distributed (non-IID) traffic patterns, leading to label imbalance, feature drift, and temporal variability that can substantially degrade the generalisation performance of centrally trained models.

Federated Learning (FL) has emerged as a promising paradigm for addressing these privacy and heterogeneity challenges by enabling collaborative model training without requiring the exchange of raw data [7], [8]. In the FL setting, participating clients train local models on-site and share only model parameters or updates with a coordinating server. Recent IoMT-focused studies indicate that FL can maintain effective performance even under heterogeneous data distributions and diverse client characteristics, particularly when combined with appropriate aggregation strategies and preprocessing techniques [9], [12], [13].

This paper investigates a federated intrusion detection framework based on a deep multilayer perceptron (MLP) trained on IoMT network traffic from the WUSTL-EHMS-2020 dataset. The primary objective of this work is to explore

federated learning as a privacy-preserving and secure machine learning paradigm for intrusion detection in distributed IoMT systems.

II. RELATED WORK

A. Intrusion Detection in IoMT

Centralised machine-learning-based intrusion detection systems have been extensively investigated for IoMT and healthcare networks due to their interpretability and relatively modest computational requirements. However, such approaches require aggregating sensitive network traffic data, which raises substantial privacy, security, and regulatory compliance concerns [4], [6]. Benchmark studies evaluating tree-based classifiers, Support Vector Machines (SVM), and multilayer perceptron (MLP) architectures on medical and Internet of Things (IoT) datasets report competitive detection accuracy. Nevertheless, these approaches rely on centralised data collection and often assume balanced or only moderately imbalanced class distributions [6], [10].

More recent IoMT-focused research emphasises the importance of domain-specific datasets and tailored preprocessing methodologies for reliable IDS evaluation. Dome`nech *et al.* [9] demonstrate that IDS models trained on generic IoT datasets can misrepresent performance characteristics when deployed in IoMT environments, and show that careful temporal windowing and dataset-specific handling substantially improve both detection accuracy and robustness. Related studies propose feature selection and adaptive anomaly detection techniques that balance detection performance against resource utilisation, but continue to operate within a centralised training paradigm [11]. Collectively, these contributions motivate the adoption of federated approaches that preserve the benefits of sophisticated feature engineering while eliminating the need for centralised raw-data aggregation.

B. Federated Learning for IoMT Security

FL enables distributed clients to collaboratively learn a shared model by periodically exchanging model parameters or gradients rather than raw data, thereby reducing privacy risks and regulatory friction [7], [8]. Within the IoMT domain, FL has been explored in hierarchical, energy-aware, and reinforcement-based IDS architectures, where hospital gateways, edge devices, and fog nodes coordinate model training while respecting data locality constraints and device limitations [12], [13].

Recent federated IDS frameworks tailored to IoMT environments identify several critical design considerations. Non-independent and identically distributed (non-IID) data arising from heterogeneous institutions and diverse device roles can significantly affect convergence behaviour and final model performance, necessitating techniques such as adaptive client selection, label-aware partitioning, or hybrid federated-transfer learning strategies to stabilise training dynamics [9], [12]. Fault tolerance—encompassing client dropout handling, network partition management, and resilience to Byzantine

behaviour—is increasingly recognised as essential in safety-critical medical environments. In parallel, temporal models based on long short-term memory (LSTM) or related architectures have demonstrated improved capability to capture sequential patterns in medical traffic compared to purely static models, particularly when combined with temporal aggregation strategies to reduce communication and computational overhead [9].

In contrast to these more complex federated LSTM-based pipelines, the present work focuses on a deep MLP architecture that is straightforward to deploy, interpret, and analyse. This design choice prioritises implementation simplicity and transparency while remaining aligned with the broader research direction of privacy-preserving, federated intrusion detection for IoMT systems.

III. METHODOLOGY

This section describes the proposed privacy-preserving cybersecurity framework for IoMT environments. The methodology is presented through a formal description of the hierarchical system architecture, the assumed threat model, the data preprocessing pipeline derived from the experimental implementation, and the FL formulation adopted for intrusion detection.

A. System Architecture and Operational Workflow

The proposed framework targets distributed IoMT environments characterised by heterogeneous endpoints and strict privacy constraints. As illustrated in Figure 1, the system architecture is organised into four hierarchical and interdependent layers: the physical IoMT sensor layer, the edge computing and local training layer, the federated coordination layer, and the global deployment backend.

1) *Physical IoMT Sensor Layer*: The foundational layer comprises a heterogeneous collection of medical devices interconnected within healthcare institution networks. These endpoints include wearable vital-sign monitors, implantable sensors, and bedside patient monitoring systems. Each device continuously generates biomedical telemetry and control traffic, resulting in network patterns that differ significantly from those observed in generic IoT environments.

2) *Edge Computing and Local Training Layer*: At the network edge, hospital gateway nodes operate as federated learning clients. A strict **privacy boundary** is enforced at this layer, as indicated by the padlock symbols in Figure 1. Raw network traffic features X_k , and their corresponding labels y_k are processed exclusively within local institutional environments. Only model parameters, such as weights or gradients, are transmitted to the central aggregation server. This design ensures that raw patient data and internal network configurations never leave institutional boundaries, thereby preserving privacy and regulatory compliance.

3) *Federated Aggregation and Global Distribution*: The central federated aggregation server coordinates the collaborative learning process. During each communication round t , the server broadcasts the current global model parameters $w^{(t)}$

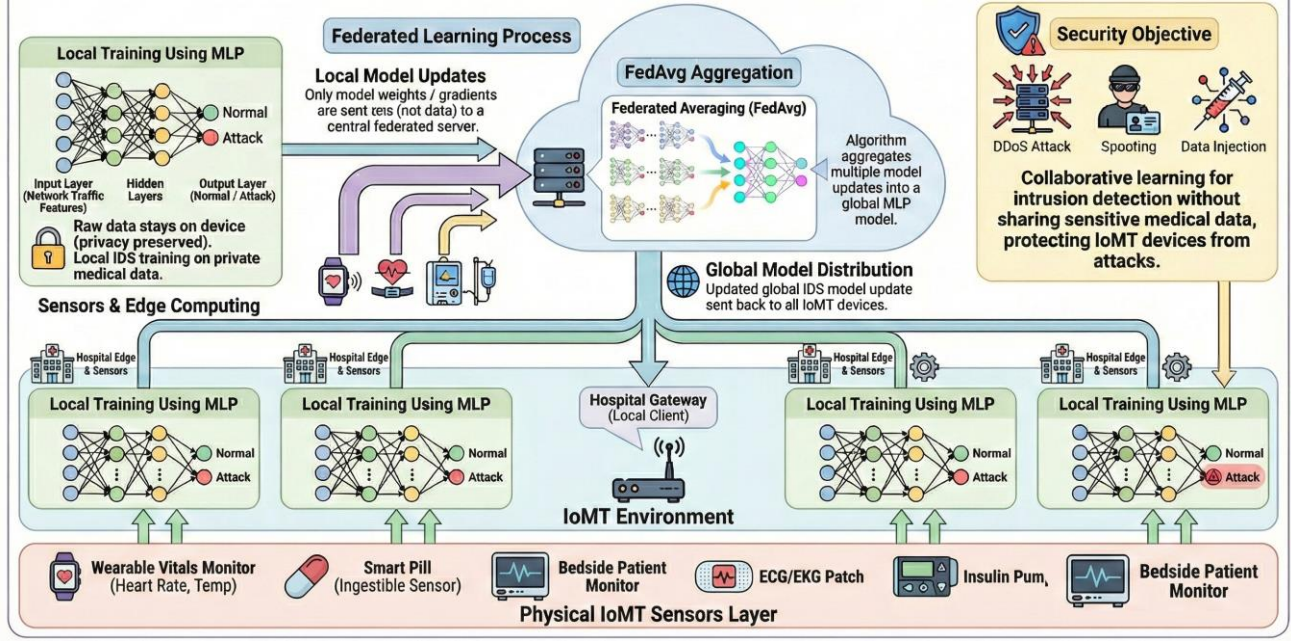


Fig. 1. Privacy-preserving federated learning architecture for IoMT intrusion detection: system overview and security objectives.

to all participating clients and aggregates the locally updated parameters $w_k^{(t+1)}$ using the Federated Averaging (FedAvg) algorithm. Following aggregation, the updated global model is redistributed to all gateway nodes, enabling real-time detection of cyber threats such as denial-of-service attacks and device impersonation attempts.

B. Threat Model

The threat model assumes an adversary capable of generating malicious network traffic targeting IoMT devices and gateways. The primary security objectives, illustrated in the right panel of Figure 1, include the detection of the following attack categories:

- **Denial-of-Service (DoS) Attacks:** Volumetric or resource-exhaustion attacks targeting gateway computational and communication resources.
- **Spoofing and Masquerading Attacks:** Attempts to impersonate legitimate medical devices or gateways within the IoMT network.
- **Malicious Data Injection:** Injection of falsified telemetry data or unauthorised control messages intended to disrupt clinical operations.

The system objective is to maximise the detection of anomalous network flows ($y = 1$) while maintaining high precision for benign clinical traffic ($y = 0$).

C. Dataset Characterisation and Preprocessing Pipeline

The proposed framework is evaluated using the **WUSTL-EHMS-2020** dataset, a domain-specific benchmark that captures real-world cyber-biomedical telemetry from an elec-

tronic health monitoring system. To ensure reliable model training and evaluation, the raw dataset undergoes a rigorous preprocessing pipeline, as described in the following subsections.

1) *Feature Selection and Sanitisation:* To prevent model overfitting to site-specific or non-generalizable identifiers, observational attributes such as Flow ID, Source IP, Source Port, Destination IP, Destination Port, Protocol, and Timestamp are systematically removed. The retained feature set, therefore, focuses exclusively on flow-level statistical characteristics. In addition, the dataset is sanitised by removing instances containing missing or infinite values.

2) *Label Binarisation:* The original multiclass labels are mapped to a binary intrusion detection formulation. All attack-related classes are assigned to the *abnormal* class ($y = 1$), while legitimate Electronic Health Monitoring System (EHMS) traffic is assigned to the *normal* class ($y = 0$). This binarisation reflects practical operational requirements in clinical environments, where rapid distinction between benign and malicious traffic is critical.

3) *Feature Normalisation:* To stabilise gradient-based optimisation during neural network training, Z-score standardisation is applied to the feature matrix X . Each feature is transformed to have zero mean and unit variance using statistics computed exclusively from the training set, thereby preventing information leakage into the test data.

4) *Class Imbalance Mitigation:* Given the inherent class imbalance typical of intrusion detection datasets, random over-sampling is applied to the minority (attack) class in the training

set until the classes are balanced ($N'_0 = N'_1$). This strategy mitigates classifier bias toward the majority benign class and improves sensitivity to rare but critical attack patterns.

D. Federated Learning Formulation and Experimental Configuration

1) *Federated Averaging Protocol*: A federated learning environment is simulated with $K = 3$ clients. The balanced training dataset is randomly partitioned into K equal-sized subsets, resulting in an independent and identically distributed (IID) experimental setting. The central server aggregates local model updates using the standard Federated Averaging (FedAvg) rule:

$$w^{(t+1)} = \frac{1}{K} \sum_{k=1}^K w_k^{(t+1)}, \quad (1)$$

where $w_k^{(t+1)}$ denotes the updated model parameters returned by client k at communication round t .

2) *Deep Neural Network Architecture*: The local intrusion detection model is implemented as a deep multilayer perceptron (MLP) comprising an input layer, four fully connected hidden layers with 256, 128, 64, and 32 neurons, respectively, and a single output neuron with sigmoid activation for binary classification. Rectified Linear Unit (ReLU) activation functions are employed in all hidden layers. To enhance generalisation capability and reduce overfitting, Dropout regularisation with probability $p = 0.3$ is applied after each hidden layer.

3) *Training Configuration*: The federated training protocol is configured for $T = 30$ communication rounds. During each round, participating clients perform $\tau = 5$ local training epochs using a batch size of $B = 32$. Model optimisation is carried out using the Adam optimiser with the binary cross-entropy loss function.

IV. EXPERIMENTAL RESULTS

A. Evaluation Metrics

The global federated model is evaluated using standard IDS performance metrics [10], [11]. Let TP, TN, FP, and FN denote true positives, true negatives, false positives, and false negatives, respectively. The evaluation metrics are defined as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (2)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

In addition, the Receiver Operating Characteristic Area Under the Curve (ROC-AUC) metric is computed from predicted class probabilities using the Receiver Operating Characteristic (ROC) curve, in accordance with established methodological practice in IDS research [9].

B. Federated Data Distribution Analysis

Before analysing classification outcomes, it is necessary to examine the data distribution across participating federated clients to verify adherence to the IID experimental configuration. Figure 2 illustrates the sample allocation among the three federated clients following the preprocessing and class-balancing procedures described in Section III-C.

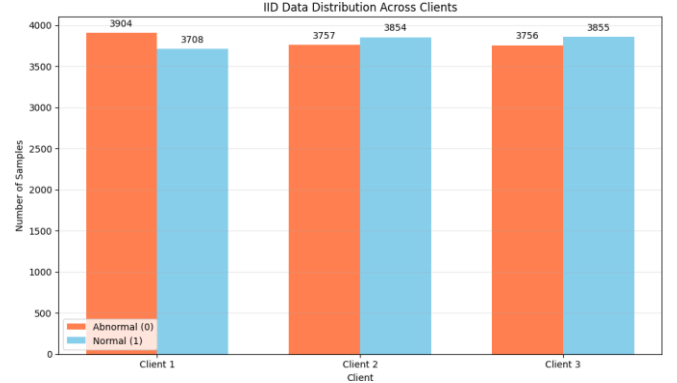


Fig. 2. Balanced sample distribution across federated clients under IID partitioning.

As shown in Figure 2, the stochastic partitioning strategy results in a near-uniform sample distribution across all participating clients, with each institutional node receiving approximately 7,611 samples. Furthermore, intra-client class balance is preserved, with normal and abnormal traffic instances represented in approximately equal proportions within each local partition. This configuration ensures that local gradient estimates remain statistically unbiased during federated optimisation, thereby facilitating stable convergence of the Federated Averaging (FedAvg) protocol and mitigating the risk of client drift.

C. Federated Training Convergence Behaviour

Figure 3 illustrates the temporal evolution of global model loss and classification accuracy over 30 federated communication rounds, providing insight into the convergence dynamics of the distributed training process.

Fig. 3. Temporal evolution of global model performance across federated communication rounds.

The loss trajectory depicted in the left panel of Figure 3 demonstrates consistent convergence behaviour, with a steady reduction in global loss across successive communication rounds. Minor oscillations are observed, characteristic of federated optimisation due to stochastic local updates and inter-client variability, and are subsequently attenuated by averaging. The accuracy curve shown in the right panel indicates rapid convergence to near-optimal performance within the first few communication rounds, followed by sustained stability. Minor fluctuations observed at specific rounds (e.g., rounds 7,

11, 17, and 25) temporally align with loss oscillations and reflect the inherent stochasticity of distributed stochastic gradient descent.

D. Classification Performance Results

Under the evaluated experimental conditions, the federated MLP-based IDS achieves strong classification performance. Table I summarises the quantitative evaluation results.

TABLE I
QUANTITATIVE CLASSIFICATION PERFORMANCE OF THE FEDERATED
MLP-BASED INTRUSION DETECTION SYSTEM

Performance Metric	Achieved Value
Accuracy	1.0000
Precision	1.0000
Recall (Sensitivity)	1.0000
F1-score	1.0000
ROC-AUC	1.0000

Figure 4 presents the corresponding confusion matrix, offering a detailed view of the classification outcomes for each traffic class.

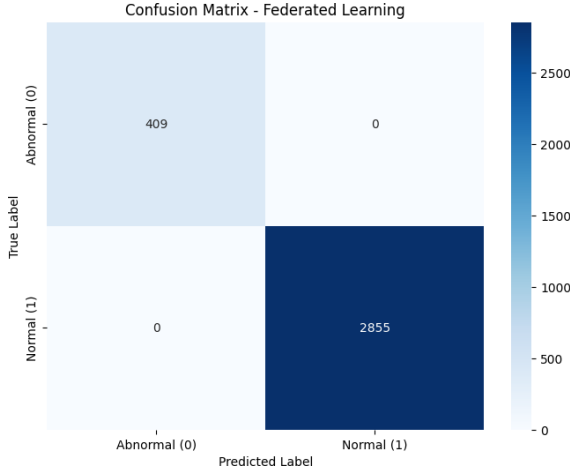


Fig. 4. Confusion matrix of the federated MLP-based intrusion detection system.

As illustrated in Figure 4, the federated model correctly classifies all abnormal traffic instances (409 attack samples) and all normal traffic instances (2,855 benign samples), resulting in zero false positives and zero false negatives. The absence of false positives indicates that no legitimate clinical traffic is erroneously flagged as malicious, a critical operational requirement in healthcare environments where excessive false alarms can disrupt clinical workflows. Similarly, the absence of false negatives ensures comprehensive detection of malicious activity, preventing undetected attacks from compromising system integrity.

The Receiver Operating Characteristic curve shown in Figure 5 provides further confirmation of the discriminative capability of the trained federated model.

The ROC curve in Figure 5 indicates that the federated model achieves a true positive rate of 1.0 at a false positive

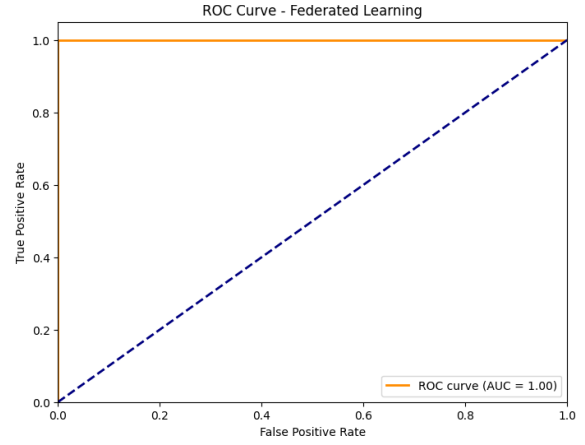


Fig. 5. Receiver Operating Characteristic (ROC) curve demonstrating optimal classifier discrimination.

rate of 0.0, corresponding to an AUC of 1.00. This result suggests that, under the evaluated conditions, the learned decision boundary effectively separates normal and abnormal traffic patterns within the feature space.

E. Discussion

The empirical results presented in this study indicate that, under balanced and IID experimental conditions, a federated MLP architecture can achieve strong intrusion detection performance on IoMT network traffic while strictly preserving data locality. This observation is consistent with, and extends, prior findings demonstrating that deep learning models trained on domain-specific IoMT datasets can attain very high classification accuracy when preprocessing procedures and feature engineering strategies are carefully aligned with the statistical properties of the underlying traffic.

The convergence analysis shown in Figure 3 further demonstrates that the FedAvg protocol achieves stable optimisation within 30 communication rounds. This behaviour suggests favourable computational efficiency and communication characteristics, which are important considerations for deployment in resource-constrained healthcare network infrastructures. Nonetheless, it is important to note that these results are obtained under controlled IID conditions and balanced class distributions, which may not fully reflect the heterogeneity and operational complexity encountered in real-world IoMT deployments.

V. APPLICATION TO IoMT SYSTEMS

IoMT-based healthcare and remote monitoring systems generate high-frequency sensor and control data streams that are both safety-critical and privacy-sensitive [3]. Within such environments, the proposed federated learning framework offers several operational advantages:

- Each IoMT endpoint or subsystem can operate as a federated learning client, locally training the intrusion detection model on control and network logs while sharing only model parameters with a coordinating server.

- Anomalous commands, timing irregularities, and atypical network flows can be mapped into a unified feature space and identified by the federated intrusion detection system.
- The intrusion detection system can be integrated with runtime safety and risk-management frameworks (e.g., ISO 14971) to trigger safe-state transitions or human-in-the-loop interventions upon detection of anomalous behaviour [14].

These application scenarios underscore the need for lightweight yet robust federated intrusion detection solutions that can operate within the computational, safety, and regulatory constraints characteristic of IoMT healthcare environments.

VI. CONCLUSION

This paper presented a federated learning-based intrusion detection framework for distributed IoMT healthcare environments. Leveraging the WUSTL-EHMS-2020 dataset and a deep multilayer perceptron classifier trained using the Federated Averaging protocol across three clients, the proposed framework demonstrates strong intrusion detection performance on a held-out test set under controlled, IID conditions. These results indicate that architecturally simple deep learning models can provide effective intrusion detection capabilities while preserving data locality and avoiding the centralisation of sensitive network traffic. Future work will focus on developing non-IID-aware training strategies, including label-aware client partitioning and adaptive aggregation mechanisms, to improve robustness under realistic cross-site heterogeneity.

ARTIFACT AVAILABILITY

To support transparency, reproducibility, and future research, the complete source code, data preprocessing scripts, and federated learning implementation developed and used in this study are publicly available at the following repository:

<https://github.com/imadim/Federated-Learning-Secure-IoMT>

REFERENCES

- [1] C. Huang, J. Wang, S. Wang, and Y. Zhang, "Internet of medical things: A systematic review," *Neurocomputing*, vol. 557, p. 126719, 2023. [Online]. Available: <https://doi.org/10.1016/j.neucom.2023.126719>
- [2] A. H. Dalloul, F. Miramirkhani, and L. Kouhalvandi, "A review of recent innovations in remote health monitoring," *Micromachines*, vol. 14, no. 12, p. 2157, 2023. [Online]. Available: <https://doi.org/10.3390/mi14122157>
- [3] G. D. Gallo and D. Micucci, "Internet of medical things systems review: Insights into non-functional factors," *Sensors*, vol. 25, no. 9, p. 2795, 2025. [Online]. Available: <https://doi.org/10.3390/s25092795>
- [4] U.S. HHS, "Summary of the HIPAA Security Rule," 2024. [Online]. Available: <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>
- [5] European Union, "Regulation (EU) 2016/679 (General Data Protection Regulation)," 2016. [Online]. Available: <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
- [6] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, "Intrusion detection system for healthcare systems using medical and network data: A comparison study," *IEEE Access*, vol. 8, pp. 106576–106584, 2020. [Online]. Available: <https://doi.org/10.1109/ACCESS.2020.3000421>
- [7] Y. Jin, H. Zhu, J. Xu, and Y. Chen, *Federated Learning Fundamentals and Advances*. Springer, 2022.
- [8] X. Gu, F. Sabrina, Z. Fan, and S. Sohail, "A review of privacy enhancement methods for federated learning in healthcare systems," *Int. J. Environ. Res. Public Health*, vol. 20, no. 15, p. 6539, 2023. [Online]. Available: <https://doi.org/10.3390/ijerph20156539>
- [9] J. Dome'nech, O. Leo'n, M. S. Siddiqui, and J. Pegueroles, "Evaluating and enhancing intrusion detection systems in IoMT: The importance of domain-specific datasets," *Internet of Things*, vol. 32, p. 101631, 2025. [Online]. Available: <https://doi.org/10.1016/j.iot.2025.101631>
- [10] A. Binbusayyis, H. Alaskar, T. Vaiyapuri, and M. Dinesh, "An investigation and comparison of machine learning approaches for intrusion detection in IoMT network," *J. Supercomput.*, vol. 78, no. 15, pp. 17403–17422, 2022. [Online]. Available: <https://doi.org/10.1007/s11227-022-04568-3>
- [11] D. Alsalman, "A comparative study of anomaly detection techniques for IoT security using adaptive machine learning for IoT threats," *IEEE Access*, vol. 12, pp. 14719–14730, 2024. [Online]. Available: <https://doi.org/10.1109/ACCESS.2024.3359033>
- [12] I. Ioannou *et al.*, "GEMLIDS-MIOT: A green effective machine learning intrusion detection system based on federated learning for medical IoT network security hardening," *Comput. Commun.*, vol. 218, pp. 209–239, 2024. [Online]. Available: <https://doi.org/10.1016/j.comcom.2024.02.023>
- [13] I. A. Khan *et al.*, "Fed-Inforce-Fusion: A federated reinforcement-based fusion model for security and privacy protection of IoMT networks against cyber-attacks," *Inf. Fusion*, vol. 101, p. 102002, 2024. [Online]. Available: <https://doi.org/10.1016/j.inffus.2023.102002>
- [14] ISO, "ISO 14971:2019 — Medical devices — Application of risk management to medical devices," 2019. [Online]. Available: <https://www.iso.org/standard/72704.html>