International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.4 (2025) pp. 8922-8930
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

Federated Deep Learning for Robust and Scalable Intrusion Detection in the Internet of Medical Things

Mohammed Kamel Benkaddour[1]*, Malika Abid[2], Zineb Abbazi[3], Katia Bouhnik[4]

[1] Department of Computer Science and Information Technology,Artificial Intelligence and Information Technology Laboratory (LINATI), University of Kasdi Merbah, Ouargla, 30000, Algeria.
* **Corresponding Author Email:** Benkaddour.kamel@univ-ouargla.dz- **ORCID:** 0009-0006-2471-3242

[2] Department of Computer Science and Information Technology,Artificial Intelligence and Information Technology Laboratory (LINATI), University of Kasdi Merbah, Ouargla, 30000, Algeria.
Email: Abid.malika@univ-ouargla.dz- **ORCID:** 0009-0005-7359-5177

[3] Department of Computer Science and Information Technology, University of Kasdi Merbah, Ouargla, 30000, Algeria.
Email: zineb.abbazi.2002@gmail.com- **ORCID:** 0009-0007-5393-7381

[4] Department of Computer Science and Information Technology, University of Kasdi Merbah, Ouargla, 30000, Algeria.
Email: katiabouhnik@gmail.com- **ORCID:** 0009-0007-2252-4787

Article Info:

DOI: 10.22399/ijcesen.4038
Received : 25 September 2025
Revised : 05 November 2025
Accepted : 10 November 2025

Keywords

Federated Learning,
Deep Learning,
CNN–LSTM,
Internet of Medical Things,
Intrusion Detection system

Abstract:
The Internet of Medical Things (IoMT) connects wearable devices, sensors, and healthcare systems to enable continuous patient monitoring and intelligent diagnostics. While offering significant benefits, this connectivity also exposes IoMT to cyberattacks that threaten data integrity and patient safety. Intrusion detection is therefore essential, but traditional centralized methods raise concerns of privacy leakage, high communication cost, and single points of failure. To address these issues, we propose a federated deep learning framework that employs a hybrid Convolutional Neural Network and Long Short-Term Memory CNN–LSTM architecture for intrusion detection. The federated approach allows collaborative model training across distributed clients without sharing raw medical data, preserving privacy while enhancing scalability. Experiments conducted on the CIC-IoMT2024 dataset under both IID and non-IID data distributions demonstrate that the framework achieves up to 99% accuracy in binary classification and strong robustness in multi-class scenarios. These findings confirm that federated deep learning offers a robust and scalable solution for securing IoMT networks while safeguarding sensitive medical information.

1. Introduction

Healthcare systems are implementing digital technologies to enhance diagnostic efficiency, patient monitoring, and tailored care. The Internet of Medical Things is central to this revolution, bringing together wearable sensors, medical imaging devices, and smart healthcare infrastructure. While IoMT brings significant benefits, its highly connected and distributed nature makes it vulnerable to cyberattacks, ranging from denial-of-service to data manipulation, which can compromise both data integrity and patient safety [1-2]. Traditional IDS design follows CL architectures, where all the training data are collected on a central server to be employed for model training and inference. While centralized models can ensure high detection accuracy [5], they come with major limitations like privacy concerns regarding data storage in a single point, low scalability, and vulnerability to single points of failure [6, 7]. Federated learning (FL) has emerged as a decentralized method that eliminates the aforementioned limitations [14]. FL improves data privacy and communication overhead minimization, and is competitive in terms of detection performance in decentralized environments [3,4]. This approach enhances data privacy, reduces communication overhead, and improves resilience. Despite its promise, FL in IoMT environments faces challenges such as heterogeneous data across devices, varying resource constraints, and adversarial conditions. To tackle these challenges, this paper introduces a federated deep learning framework for intrusion detection in

IoMT networks. The framework employs a CNN–LSTM model to capture both spatial features and temporal dependencies in network traffic. By distributing model training across IoMT clients and aggregating updates in a federated manner, the framework preserves privacy, scales efficiently, and maintains robustness under heterogeneous data settings. We evaluate the framework on the CIC-IoMT2024 dataset, comparing it against centralized and baseline federated methods. The results demonstrate that our approach achieves high accuracy, and robustness, making it a practical solution for next-generation medical IoT environments security. The remaining sections of this work are organized as follows. Section 2 provides a review of related research on intrusion detection and federated learning in IoMT systems. Section 3 discusses the proposed federated deep learning framework, which includes the CNN-LSTM architecture and data distribution method. Section 4 covers the experimental setup, dataset description, evaluation measures, and implementation parameters. Section 5 explains the collected results and provides a comparison to existing approaches. Finally, Section 6 summarizes the paper and provides areas for future research.

2. Related Works

Recent research in IoMT security has increasingly focused on leveraging artificial intelligence and machine learning to detect and mitigate cyber threats. Traditional centralized intrusion detection systems (IDS) have demonstrated strong detection capabilities but suffer from privacy risks and scalability issues due to centralized data aggregation. Berguiga et al. [6] introduced HIDS-IoMT, a CNN-LSTM hybrid model executed on fog nodes with Raspberry Pi, achieving 99.92% accuracy with IoTID20 and Edge-IIoT datasets. Zachos et al. [8] described a lightweight anomaly based IDS, host- and network-level monitoring combined with KNN, decision tree (DT), and random forests (RF), with over 99.6% accuracy with the TON IoT and Power trace datasets. Areia et al. [9] presented IoMT Traffic Data, traffic flow-level dataset for IoMT with emphasis on feature representation but without implementing an IDS model or taking privacy into account. Alalhareth and Hong [11] proposed ME-IDS, a meta-learning ensemble wherein stacking and dynamic voting are implemented, achieving 98% accuracy on the WUSTL dataset. Yet, the system remains centralized and energy and privacy are not considered. Thamilarasu et al. [10] presented a hierarchical IDS using mobile agents and machine learning for network and device-level defense, achieving a detection accuracy of approximately 99.6% in Castalia/OMNeT++ simulations. Hernandez-Jaimes et al. [16] used attention mechanisms and natural language processing techniques to achieve unsupervised anomaly detection using OC-SVMs. Their model attained a 95.53% F1-score on CIC- IoMT-2024 and MQTT-IoT-IDS2020, but remained centralized. Dadkhah et al. [17] introduced the CIC-IoMT- 2024 dataset with 18 attack classes suitable for benchmarking but based on centralized models without any privacy-preserving techniques. The issue stood of privacy and scalability of centralized models. That is why the FL-based ID systems had been receiving attention in recent years. Otoum et al. [7] propose a federated transfer learning system that trains deep neural networks across edge devices. Evaluation using CIC-IDS-2017 has shown that their model improved on personalization while maintaining data locality. Fahim Islam et al. [12] proposed FedIoMT, a federated framework based on KANConvNet with cluster based aggregation. It exhibited an outstanding performance profile with over 99.2% accuracy on four benchmark datasets with little computational overhead. Albahri et al. [13] presented a decision framework using fuzzy logic to choose the best classifier to be deployed for federated IDS within IoMT. Although complete, this fails to incorporate temporal architectures, such as LSTM, combined with federated learning-based defenses, thereby limiting robustness in the face of adversarial attacks.

3. Material and Methods

This section focuses on reviewing and analyzing pivotal technologies that play a key component in the development of intrusion detection systems: Centralized Learning and Federated Learning. It begins by presenting the traditional approach based on centralized data processing, highlighting its advantages and limitations. Then transitions to federated learning as an alternative paradigm that enhances data privacy and enables collaboration between multiple entities without the need to share raw data to create decentralized and secure learning environments.

3.1. Dataset Description

This study used the CIC-IoM-T2024 dataset developed by the Canadian Institute for Cybersecurity at the University of New Brunswick [18]. It is regarded as one of the top datasets in the field of IoMT security. The data was acquired in a realistic testbed environment with 40 medical equipment, both genuine and simulated. These devices communicated using several protocols. The

dataset includes 18 different forms of cyberattacks, divided into five categories: Denial of Service (DoS), Distributed Denial of Service (DDoS), reconnaissance, and MQTT.

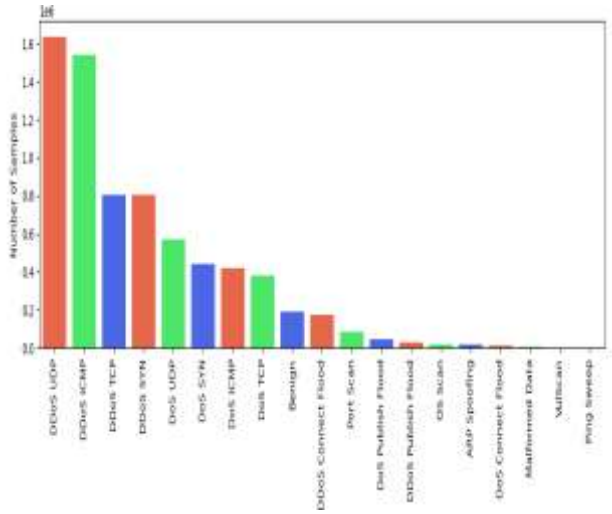


Figure 1. *Distribution of attack types and benign in the CIC-IoMT-2024 Dataset.*

3.2. Data Preprocessing

This step is important in preparing data for deep learning models since it directly affects model correctness and reliability? Noise, missing values, and class imbalances are common in raw data, and they all have a negative impact on model performance.

- **Step 1- Data Collection:** The dataset was obtained directly from the official Canadian Institute for Cybersecurity [18] repository and subsequently uploaded to Google Drive to ensure streamlined access and centralized management throughout the preprocessing and analysis pipeline.

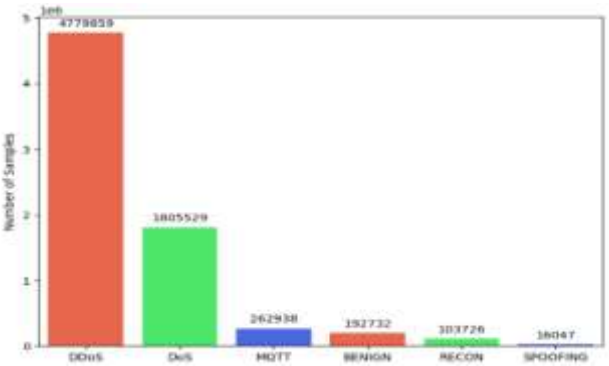
- **Step 2- Feature and Target Selection:** Columns (0-45) were selected as input features, and column 46 was used as the target variable.

- **Step 3- Removal of Non-Numeric and Constant Features:** Non-numeric features were excluded. Additionally, features with only a single unique value were removed, as they do not contribute discriminative information.

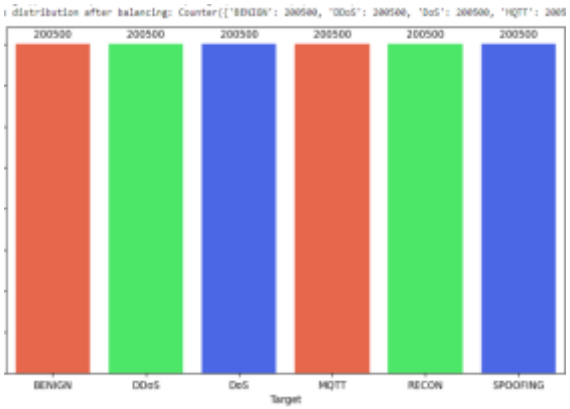
- **Step 4- Standardization:** Each feature was standardized to have zero mean and unit variance, where μ represent the mean and **σ** the standard deviation.

$$X_{standardized} = \frac{X-\mu}{\sigma} \qquad (1)$$

- **Step 5- Data Balancing:** Two complementary strategies were employed to address class imbalance in the data set: Synthetic Minority Oversampling Technique (SMOTE) and Random Undersampling. SMOTE generates synthetic instances for minority classes by interpolating between existing points and their nearest neighbors in feature space, thereby enhancing the minority class without duplication. Conversely, Random Under-sampling reduces majority classes in size by randomly removing instances, which balances the class distribution but has the possibility of losing potentially useful samples. Both techniques help generate a more balanced and representative training data set.



(a) Before balancing



(b) After balancing

Figure 2. *Data Balancing using SMOTE.*

- **Step 6 - MinMaxScaler:** The MinMaxScaler was applied to the balanced dataset in order to normalize the feature values and ensuring that all feature values fall within a specified range, typically [0, 1].

- **Step 7 - Data Splitting:** The data was split into training, testing, and validation sets according to the specified proportions. 70% of the data was allocated to the training set, while the remaining 30% was split into validation and test sets, each receiving 50%.

3.3. CL for Baseline Model Evaluation

During the CL phase, three deep learning architectures were tested to see which one could provide the intrusion detection baseline for IoMT settings. The first CNN architecture included three convolutional layers plus a fully connected layer wherein the objective was to draw spatial features from the network traffic data. The second model, an LSTM network, consisted of two LSTM layers and two fully connected layers. This model could learn long-term temporal dependencies in sequential inputs. The third architecture, namely the CNN–LSTM, consists of two convolution layers inside a single LSTM layer; the output from the LSTM units then connected to a fully output layer for jointly modeling spatial and temporal features. This allowed CNN–LSTM to learn complex patterns present in IoMT traffic data in scenarios where spatial correlations and temporal sequences were equally relevant. This model, benefiting from its architectural suitability for addressing aspects in spatiotemporal IoMT data, was tested in both binary and multiclass classification scenarios.

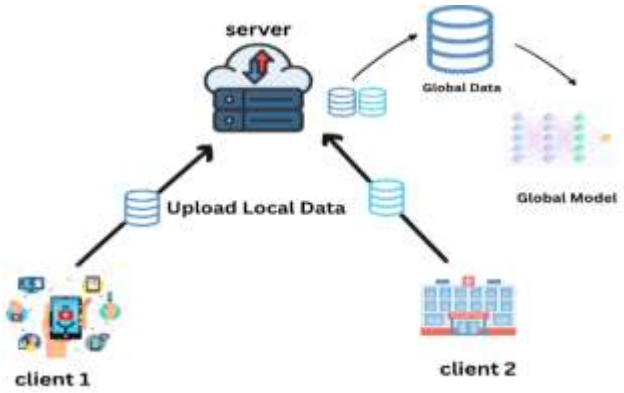


Figure 3. *Centralized Learning model.*

3.4. Federated Learning

To keep going based on what was learned during the centralized training, the CNN-LSTM model, selected for its ability to grasp spatial and temporal properties, was deployed in a federate learning setup. Federate Learning is a decentralized setup whereby many clients cooperate to train a common model without the exchange of any raw data between them, thus preserving privacy. In this phase, the CNN-LSTM model was trained across distributed clients, both under independent and identically distributed (IID) as well as Non-IID data distribution setups to better simulate realistic deployment scenarios in IoMT environments. Two different aggregation strategies were used: one is FedAvg [15], which averages client updates in a weighted fashion; the second is FedProx [20], which stabilizes the training in heterogeneous settings by imposing a proximal term. This whole setup aimed at improving privacy of data across clients, addressing statistical heterogeneity in the client system, and ultimately incrementing the performance of the global model in its intrusion-detection capacity within diverse distributed IoMT systems.

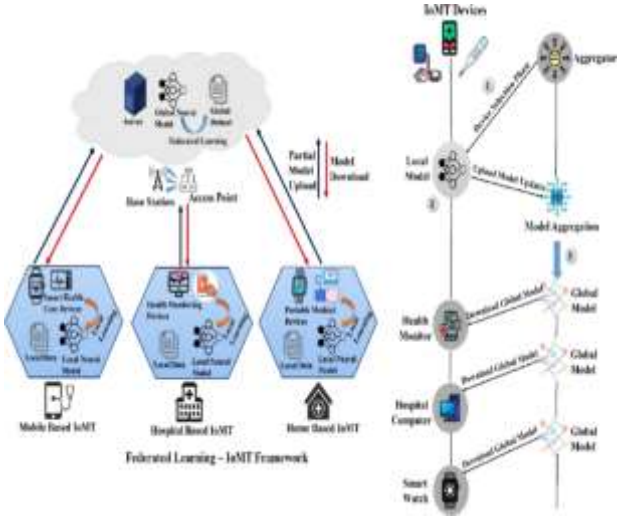
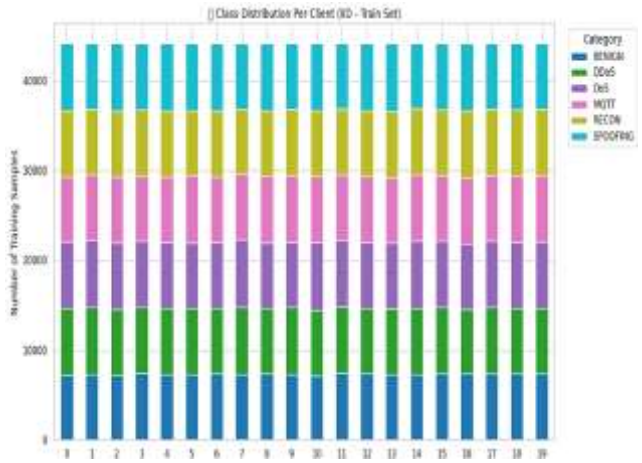


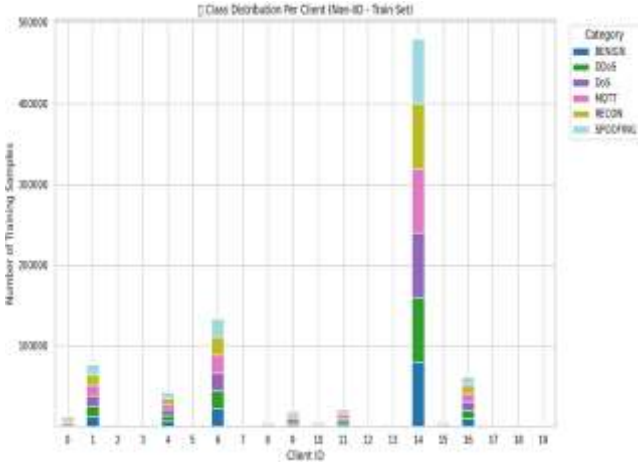
Figure 4. *Federated Learning model.*

3.5. Data Distribution

To evaluate the strength of the proposed IDS, we experimented under both IID and non-IID data settings. This enables end-to-end exploration of the impact of heterogeneity of data, common in real IoMT settings, on model performance and behaviour. In the IID scenario, data samples were distributed evenly across all contributing clients in a way that each client received a representative sample of the worldwide dataset. This ensured faster convergence and stable training dynamics. In the non-IID scenario, the simulation was closer to real-world conditions where the clients had data whose statistical features were distinct. To generate non-IID distributions, we used the Dirichlet distribution by changing the concentration parameter α, which controls the heterogeneity. The Dirichlet distribution is a simple multi-dimensional continuous probability distribution that allowed us to capture various degrees of data skewness and fragmentation. Non-IID settings represent real federated learning scenarios with heterogeneous data usage and ownership and are different from IID distributions. Combining these environments with our test platform guaranteed not just that the IDS is effective in ideal, even situations but also resilient against the decentralized, non-uniform data environments found in actual IoMT deployments.



(a) IID Data



(b) Non-IID Data

Figure 5. *Data Distributions IID and Non-IID.*

3.6. Performance Metrics

The effectiveness of the IDS classifiers within Internet of Medical Things environments was assessed using several widely performance metrics. These metrics provide a comprehensive view of the classifiers' capabilities in identifying intrusion attempts [19]. Such evaluations are critical for optimizing the classifiers to ensure high reliability and robustness of the intrusion detection systems in real-world IoMT deployments.

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

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

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

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

4. Experiments, Results and Discussions

This section outlines the outcomes of the data preprocessing stage, including dimensionality reduction and class balancing applied to the datasets. In this context, the study addressed two types of classification tasks: a multiclass classification problem consisting of six classes, and a binary classification problem. The centralized model was first evaluated as a baseline, followed by an assessment of the performance of various federated learning strategies under different experimental conditions, including variations in the number of clients, training rounds, and data partitioning methods.

4.1. Centralized Learning Result

In the Centralized Learning setting, three types of models were evaluated: CNN, LSTM, and a hybrid CNN-LSTM. These models were compared based on their performance using several evaluation metrics in order to determine the most effective architecture. In all implemented models, the Rectified Linear Unit(ReLU) activation function was utilized. The Cross-Entropy Loss function was adopted as the training criterion. Each model was trained for 20 epochs under a centralized learning setup.

In binary classification, all models demonstrated excellent performance, with the LSTM model slightly outperforming the others in terms of accuracy, loss, and F1-score, highlighting its strength in capturing temporal dependencies in the data. The CNN-LSTM model ranked second, benefiting from its combined spatial and temporal learning capabilities. Although the CNN model ranked last, it still achieved very high performance, confirming its effectiveness in feature extraction for this type of task. In the multiclass classification task, a slight decrease in performance was observed across all models, which is expected due to the increased complexity of the task. Nevertheless, the CNN-LSTM model achieved the best overall balance among the

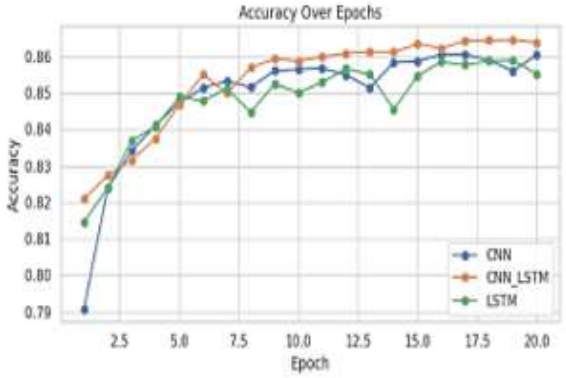
Table 1. *Binary classification parameter Model.*

Model	No. of Layers	CNN Channels	Sequence Length	No. of Classes
CNN	3 Conv + FC	64, 128, 256	45	2
LSTM	2 LSTM + 2 FC	—	45	2
CNN-LSTM	2 Conv + 1 LSTM + FC	64, 128	45	2

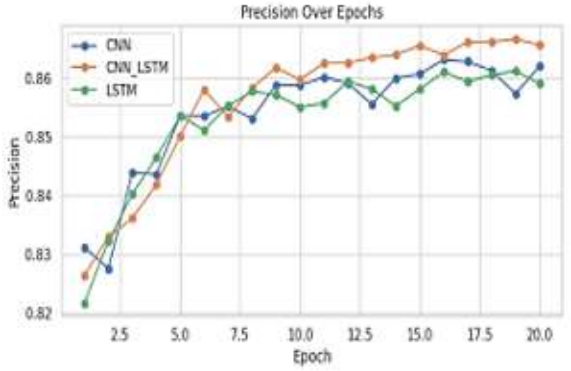
Table 2. *Binary Classification Results.*

Model	Loss	Acc	Precision	Recall	F1-score
CNN	0.0097	0.997	0.997	0.997	0.997
CNN-LSTM	0.0085	0.997	0.997	0.997	0.997
LSTM	0.0049	0.998	0.998	0.998	0.998

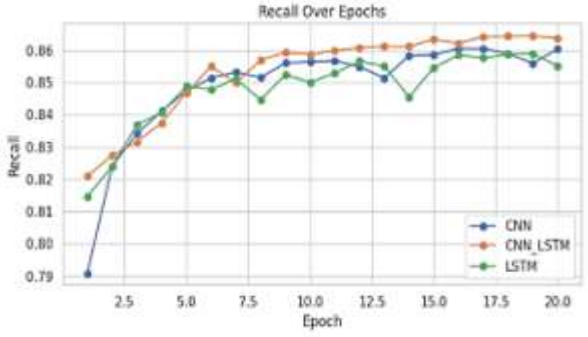
evaluation metrics, followed by the CNN model, while the LSTM model recorded the lowest performance among the three. These results indicate that combining convolutional and recurrent layers is particularly beneficial in complex multiclass tasks.



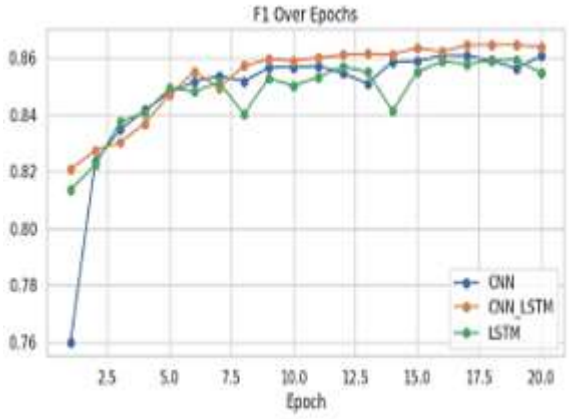
(a) Accuracy.



(b) Precision.



(c) Recall



(d) F1-score

Figure 7. *Performance of Multi-class Classification.*

4.2. Federated Learning implementation

In this study, two federated learning aggregation strategies FedAvg and FedProx were implemented and evaluated under both IID) and Non-IID data settings. The primary objective was to assess how varying the number of clients (5, 10, and 20) and communication rounds (5, 10, and 20) impacts the performance and stability of each strategy. The CNN-LSTM model, which demonstrated the highest accuracy during centralized training, was selected for the federated setup.

Table 3. *Federated Learning Accuracy (Clients = 5).*

Number of Rounds	IID		Non-IID	
	FedAvg	**FedProx**	**FedAvg**	**FedProx**
5	0.824	0.823	0.829	0.831
10	0.831	0.831	0.842	0.845
20	0.835	0.835	0.854	0.863

Table .3 presents the accuracy results when the number of clients is set to 5. Both strategies show a clear improvement in accuracy as the number of communication rounds increases. Under the IID setting, FedProx slightly outperforms FedAvg, especially at higher rounds. In contrast, under Non-IID conditions, FedAvg shows marginally better results, particularly at 20 rounds. As the number of clients increases to 10, shown in Table 4, both FedAvg and FedProx maintain stable performance under IID conditions, with nearly

Table 4. *Federated Learning Accuracy (Clients = 10).*

Number of Rounds	IID		Non-IID	
	FedAvg)	FedProx	FedAvg	FedProx
5	0.817	0.816	0.832	0.833
10	0.836	0.826	0.838	0.834
20	0.837	0.837	0.864	0.864

identical accuracy values across different communication rounds. Under the Non-IID setting, FedProx maintains a slight and consistent advantage over FedAvg, particularly as the communication rounds increase.

Table 5. *Federated Learning Accuracy (Clients = 20).*

Number of Rounds	IID		Non-IID	
	FedAvg	FedProx	FedAvg	FedProx
5	0.802	0.799	0.834	0.835
10	0.818	0.819	0.843	0.844
20	0.825	0.826	0.863	0.864

When the number of clients reaches 20, as illustrated in Table 5, a slight decline in accuracy under the IID setting is observed for both strategies. This is likely due to increased data fragmentation across more clients, which makes local training less effective. However, in the Non-IID setting, both strategies maintain high accuracy, with FedProx again demonstrating a marginal advantage, especially at higher communication rounds.Overall, both FedAvg and FedProx demonstrate strong and stable performance across various configurations. FedProx shows a slight advantage in Non-IID settings across all client numbers, whereas performance differences under IID settings are minimal. The results highlight the robustness of both strategies and the influence of communication rounds and client count on federated model accuracy. Based on the experimental results, the following observations can be made:• The results show a gradual improvement in model accuracy as the number of communication rounds increases, indicating that the model becomes more stable and effective with repeated interactions between clients and the central server.

• When comparing the FedAvg and FedProx strategies, FedProx demonstrates slightly better performance in the Non-IID data scenario, which aligns with its design objective to address data heterogeneity across clients.

• Despite the slight performance difference, the results indicate a general similarity in effectiveness between FedAvg and FedProx under the conditions of this study.

• Increasing the number of clients in a noticeable decrease in model accuracy, especially in the IID data setting. This can be attributed to the reduced amount of data available per client, which may negatively impact the quality of local model updates.

4.3 Comparison and Discussion

The proposed framework was evaluated in comparison to others techniques with different learning patterns and task categorizations. In centralized learning, our CNN-LSTM classifier showed great results in the binary task, attaining 99% accuracy, a 0.96 F1-score, and 0.98 recall. With more difficulty, our model still maintained its robustness in the multi-class task with 86.6% accuracy, a 0.92 F1-score, and 0.97 recall, outperforming several existing centralized ones. A DNN-based approach [16], for example, achieved too little accuracy (84.41%) and F1-score (91.02%) while having a good recall (98.73%), whereas a Random Forest approach [22] fared much worse with 73% accuracy and a 0.676 F1-score. Under IID data, our model gave great performance of 86% accuracy and a 0.919 F1 score; under non-IID scenarios, it remained pretty competitive, with an 85% accuracy rate and an F1 score of 0.90.Compared to the previous federated Random Forest-based binary classifier [23], which was absolutely perfect with 99% accuracy, our approach offers a better-targeted and more generalized solution that works well in multi-class scenarios, thereby ensuring better applicability and robustness to real-world intrusion detection tasks.The comprehensive experimental results of centralized, and federated learning models demonstrate a holistic view of their respective performances in intrusion detection tasks. The CNN-LSTM model performed optimally for binary and multi-class classification in the centralized learning setting, serving as a strong baseline for comparison. Shifting to the federated learning configuration, both the FedProx and FedAvg aggregation methods attained stable and progressively improved accuracy with increasing communication rounds, reflecting the utility of iterated client-server interactions for enhancing global model convergence. FedProx slightly outperformed FedAvg, particularly under non-IID conditions, confirming its robustness to client data

heterogeneity. Yet, while the number of clients increased, a loss in accuracy was observed due to the fragmentation of local datasets, which limits the usefulness of local updates and results in a scalability-accuracy trade-off.

5. Conclusions

This study presented a federated deep learning framework for intrusion detection in IoMT. The proposed system leverages the strengths of CNN–LSTM architectures to effectively capture spatiotemporal features of IoMT traffic, while federated learning enables collaborative training without sharing sensitive medical data. Experimental results on the CIC-IoMT2024 dataset confirmed that the framework achieves up to 99% accuracy in binary classification and maintains high performance under challenging non-IID data distributions. By eliminating the need for centralized data aggregation, the framework reduces privacy risks and enhances scalability, making it well-suited for real-world healthcare deployments. Moreover, the results highlight the robustness of federated deep learning against heterogeneous data and adversarial conditions, demonstrating its potential as a reliable defense mechanism for IoMT security.

Future work will focus on extending the model to more complex multi-class intrusion detection, improving efficiency for resource-constrained devices, and validating the framework in real clinical environments. Ultimately, this research contributes to building safer and more resilient digital healthcare systems where patient trust and data integrity are preserved.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Author contributions:** The article was prepared by Mohammed Kamel Benkaddour, Malika Abid, Abbazi Zineb and Bouhnik Katia. Mohammed Kamel Benkaddour and Malika Abid contributed to the literature review, conceptualization, methodology, visualization, investigation and writing original draft. Zineb Abbazi and Katia Bouhnik were responsible for the design and execution of field studies, data collection, as well as the analysis and interpretation of the data. All authors reviewed the results and approved the final version of the manuscript.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author.

References

[1] Yang, W., Zhang, J., Wang, C., & Mo, X. (2019). Situation prediction of large-scale Internet of Things network security. *EURASIP Journal on Information Security, 2019*(1), 1–12. https://doi.org/10.1186/s13635-019-0095-5

[2] Sicari, S., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks, 76*, 146–164. https://doi.org/10.1016/j.comnet.2014.11.008

[3] Ashfaq, Z., et al. (2022). A review of enabling technologies for Internet of Medical Things (IoMT) ecosystem. *Ain Shams Engineering Journal, 13*(4), 101660. https://doi.org/10.1016/j.asej.2021.10.017

[4] Khezr, S., Moniruzzaman, M., Yassine, A., & Benlamri, R. (2019). Blockchain technology in healthcare: A comprehensive review and directions for future research. *Applied Sciences, 9*(9), 1736. https://doi.org/10.3390/app9091736

[5] Hernandez-Jaimes, M. L., Martinez-Cruz, A., Ramírez-Gutiérrez, K. A., & Feregrino-Uribe, C. (2023). Artificial intelligence for IoMT security: A review of intrusion detection systems, attacks, datasets and Cloud-Fog-Edge architectures. *Internet of Things, 21*, 100887. https://doi.org/10.1016/j.iot.2023.100887

[6] Berguiga, A., Harchay, A., & Massaoudi, A. (2025). HIDS-IoMT: A deep learning-based intelligent intrusion detection system for the Internet of Medical Things. *IEEE Access.* https://doi.org/10.1109/ACCESS.2025.xxxxx

[7] Otoum, Y., Wan, Y., & Nayak, A. (2021). Federated transfer learning-based IDS for the Internet of Medical Things (IoMT). In *IEEE Globecom Workshops (GC Wkshps)* (pp. 1–6). IEEE. https://doi.org/10.1109/GCWkshps52748.2021.9682067

[8] Zachos, G., Essop, I., Mantas, G., Porfyrakis, K., Ribeiro, J. C., & Rodriguez, J. (2021). An anomaly-based intrusion detection system for Internet of Medical Things networks. *Electronics, 10*(21), 2562. https://doi.org/10.3390/electronics10212562

[9] Areia, J., Bispo, I., Santos, L., & Costa, R. L. D. C. (2024). IoMT-TrafficData: Dataset and tools for benchmarking intrusion detection in Internet of Medical Things. *IEEE Access.* https://doi.org/10.1109/ACCESS.2024.xxxxx

[10] Thamilarasu, G., Odesile, A., & Hoang, A. (2020). An intrusion detection system for Internet of Medical Things. *IEEE Access, 8*, 181560–181576. https://doi.org/10.1109/ACCESS.2020.3027983

[11] Alalhareth, M., & Hong, S. C. (2024). Enhancing the Internet of Medical Things (IoMT) security with meta-learning: A performance-driven approach for ensemble intrusion detection systems. *Sensors, 24*(11), 3519. https://doi.org/10.3390/s24113519

[12] Fahim-Ul-Islam, M., Chakrabarty, A., Alam, M. G. R., & Maidin, S. S. (2025). A resource-efficient federated learning framework for intrusion detection in IoMT networks. *IEEE Transactions on Consumer Electronics.* https://doi.org/10.1109/TCE.2025.xxxxx

[13] Albahri, O. S., et al. (2023). Rough Fermatean fuzzy decision-based approach for modelling IDS classifiers in the federated learning of IoMT applications. *Neural Computing and Applications, 35*(30), 22531–22549. https://doi.org/10.1007/s00521-023-08962-2

[14] Wen, J., Zhang, Z., Lan, Y., Cui, Z., Cai, J., & Zhang, W. (2023). A survey on federated learning: Challenges and applications. *International Journal of Machine Learning and Cybernetics*, *14*(2), 513–535. https://doi.org/10.1007/s13042-022-01621-4

[15] Li, X., Huang, K., Yang, W., Wang, S., & Zhang, Z. (2019). On the convergence of FedAvg on non-IID data. *arXiv preprint* arXiv:1907.02189. https://arxiv.org/abs/1907.02189

[16] Hernandez-Jaimes, M. L., Martinez-Cruz, A., Ramírez-Gutiérrez, K. A., & Morales-Reyes, A. (2025). Network traffic inspection to enhance anomaly detection in the Internet of Things using attention-driven deep learning. *Integration, 103*, 102398. https://doi.org/10.1016/j.vlsi.2023.102398

[17] Dadkhah, S., Neto, E. C. P., Ferreira, R., Molokwu, R. C., Sadeghi, S., & Ghorbani, A. A. (2024). CICIoMT2024: Attack vectors in healthcare devices—A multi-protocol dataset for assessing IoMT device security. *Internet of Things, 28*, 101234. https://doi.org/10.1016/j.iot.2024.101234

[18] Canadian Institute for Cybersecurity. (2024). *CICIoMT2024: A benchmark dataset for multi-protocol security assessment in IoMT*. University of New Brunswick. https://www.unb.ca/cic/datasets/iomt-dataset-2024.html

[19] Deng, Y., Eden, M. R., & Cremaschi, S. (2023). Metrics for evaluating machine learning models' prediction accuracy and uncertainty. In A. C. Kokossis, M. C. Georgiadis, & E. N. Pistikopoulos (Eds.), *Computer Aided Chemical Engineering* (Vol. 52, pp. 1325–1330). Elsevier. https://doi.org/10.1016/B978-0-443-15274-0.50211-0

[20] Li, T., Sahu, A. K., Zaheer, M., Sanjabi, M., Talwalkar, A., & Smith, V. (2020). Federated optimization in heterogeneous networks. *Proceedings of Machine Learning and Systems, 2*, 429–450.

[21] Zhu, J., Cao, J., Saxena, D., Jiang, S., & Ferradi, H. (2023). Blockchain-empowered federated learning: Challenges, solutions, and future directions. *ACM Computing Surveys, 55*(11), 1–31. https://doi.org/10.1145/3555801

[22] Misbah, A., Sebbar, A., & Hafidi, I. (2025). Innovative federated learning approach to secure Internet of Medical Things. *Innovative Technologies in Electrical Power Systems and Smart Cities Infrastructure (ICESST 2024)* (pp. 315–327). Springer. https://doi.org/10.1007/978-3-031-86705-7_21

[23] Ali, M., Saleem, Y., Hina, S., & Shah, G. A. (2025). DDoSViT: IoT DDoS attack detection for fortifying firmware over-the-air (OTA) updates using vision transformer. *Internet of Things, 30*, 101527. https://doi.org/10.1016/j.iot.2025.101527