

A Privacy-Preserving Deep Learning Framework for Real-Time Intrusion Detection in Internet of Medical Things (IoMT) Networks

Olumhense Beauty Oboghelu¹, Omorogiuwa Osaremwinda²

¹Department of Computer Science, Edo State University Iyamho, Edo State, Nigeria

²Department of Computer Science and Information Technology, Igbinedion University Okada, Edo State, Nigeria

Published October 2026

Abstract

With the explosive growth of Internet of Medical Things (IoMT) devices in contemporary healthcare, the attack surface that is vulnerable to increasingly sophisticated cyber threats has grown to encompass a variety of devices and systems. The current intrusion detection systems (IDS) built with legacy machine learning algorithms have a high false-positive rate, feature engineering takes a lot of manpower, and the algorithms show limited generalisation when faced with zero-day attacks. To tackle these limitations, this research introduces a novel deep learning framework named CNN-HBO-GAN-FL which combines a Convolutional Neural Network (CNN) optimised using the Honey Badger Optimisation Algorithm (HBOA), a Generative Adversarial Network (GAN) for class-balanced data augmentation, and Federated Learning (FL) for privacy-preserving, distributed model training. This framework was then translated into Python using TensorFlow/Keras and tested on the CICIoMT2024 benchmark dataset. The experimental results demonstrate that the proposed framework attained 99.18% accuracy, 98.96% precision, 98.72% recall, 98.84% F1-score and 1.7% false-positive rate whereas traditional CNN, CNN-LSTM and CNN-PSO baselines achieved 96.99%, 97.55% and 97.74%, respectively. The model convergence time improved by around 35% compared to the baseline models when optimized using hyperparameter tuning based on HBOA, and the minority-class attack sensitivity was improved by 16% when the model was optimized using the augmentation based on the GAN. The federated learning architecture maintained the accuracy of centralised training by less than 0.4% variance, which has the potential to help ensure adherence to data-protection laws and regulations including HIPAA, the GDPR, and Nigeria's Data Protection Act 2023, but does not necessarily ensure compliance. The results demonstrate the potentiality of the proposed framework as a resource-efficient solution for preserving privacy in intrusion detection of future IoMT infrastructure.

Keywords: Internet of Medical Things (IoMT); intrusion detection system (IDS); convolutional neural network (CNN); Honey Badger optimization algorithm (HBOA); generative adversarial networks (GANs); federated learning; cybersecurity; deep learning

1 Introduction

The Internet of Medical Things (IoMT) devices are growing in number and complexity, making the attack surface susceptible to more advanced cyber attacks that involve a range of devices and systems. Current IDSs developed using traditional machine learning algorithms are suffering from a high false positive rate, a substantial amount of manpower is needed for feature engineering, and the algorithms do not exhibit good 'zero-day' performance when encountering zero-day attacks. To tackle these limitations, this research introduces a novel deep learning framework named CNN-HBO-GAN-FL which combines a Convolutional Neural Network (CNN) optimised using the Honey Badger Optim-

isation Algorithm (HBOA), a Generative Adversarial Network (GAN) for class-balanced data augmentation, and Federated Learning (FL) for privacy-preserving, distributed model training. It has subsequently been translated into Python, implemented with TensorFlow/Keras and validated with the CICIoMT2024 benchmark dataset. It has since been translated into python with TensorFlow/Keras and validated with the CICIoMT2024 benchmark dataset. Then this framework was converted to the python language with the help of TensorFlow/Keras and applied to CICIoMT2024 benchmark dataset. The experimental results show the proposed framework achieved the highest accuracy of 99.18%, precision of 98.96%, recall of 98.72%, F1-score of 98.84% and false positive rate (FPR) of 1.7%, while the traditional CNN, CNN-LSTM and CNN-PSO baselines achieved 96.99%, 97.55% and 97.74%, respectively. When the model is optimized by hyperparameter tuning based on HBOA, the model convergence time reduced about 35% from the baseline models and when the model is optimized using the augmentation based on GAN, the minority-class attack sensitivity increased by 16%. The federated learning architecture preserved accuracy of the centralised training within 0.4% variance, which can assist in ensuring compliance with data protection laws and regulations, such as HIPAA, GDPR, and Nigeria's Data Protection Act 2023, but not guarantee compliance. The results show the feasibility of the proposed framework as an efficient solution in maintaining privacy in the intrusion detection of future IoMT infrastructure.

The four constituent techniques of this work (CNN classification, HBOA hyperparameter search, GAN augmentation, federated training) have been studied separately in the intrusion-detection literature (Section 2). The main novelty of this work is: 1) the automatic tuning of CNN architecture using HBOA instead of manual trial-and-error hyperparameter search; 2) the application of the GAN augmentation to the class-imbalance profile of the CICIoMT2024 attack categories, but not to a generic benchmark; 3) the optimised, class-balanced model is then federated-trained, addressing hyperparameter search, minority-class augmentation, and privacy preservation in an end-to-end manner instead of as separate techniques applied post hoc. The latter is placed in the context of recent state-of-the-art hybrid IDS/IoT solutions, based on architecture, dataset and reported performance.

2 Literature review

2.1 Cyber threats in IoMT environments

The IoMT ecosystem is very diverse ranging from low-power wearables to high throughput medical imaging systems that communicate through Bluetooth, Zigbee, Wi-Fi and cell networks. This diversity presents a challenge in identifying intrusions, as there is no single traffic profile for normal traffic behaviour for all devices deployed in an IoMT environment [4]. Some common attack types in IoMT networks include Distributed Denial of Service (DDoS) attacks for Denial of Service, Man-in-the-Middle (MiTM) attacks for Eavesdropping, Insider attacks for insider, and ransomware attacks for ransomware.

Mittal et al. [2] proposed a clustering based method using the gravitational search algorithm (GSA) for intrusion detection in cyber-physical medical systems and Nagarajan et al. [3] presented a method for anomaly detection using CNN-Kalman filter Gaussian mixture model (GMM). Both methods are novel, but were computationally inefficient and converged to limited extent with high volume of IoMT traffic. In the works of Ramasamy et al. [6] and Rani et al. [8] no emphasis was placed on the security of the sensors themselves, and the former only focused on AI-based detection of disease systems in IoT cyber-physical systems. Context-aware ID approach for the Internet of Medical Things (IoT-Med) was proposed by Hussain et al. [7] as a first step towards developing dedicated detection tooling for

medical devices as opposed to adapting existing tools for the Internet of Things. As a whole, this body of work highlights the need for systems that achieve an optimal balance between detection effectiveness and efficiency of computations and for the full coverage of security.

2.2 Deep learning approaches for IoMT intrusion detection

In the past 5 years, IDS design for IoMT has made great strides in the field of deep learning. To improve IDS classification performance, Kumaar et al. [9] proposed a deep learning IDS framework, called ImmuneNet, which is an intelligent training pipeline by integrating feature re-engineering, oversampling, and hyperparameter optimisation. Using deep recurrent neural networks with particle swarm optimisation for feature selection, Saheed and Arowolo [10] showed the potentials of optimising bio-inspired IDS design. In addition, the CNN-LSTM hybrid network has been proven to be effective for the analysis of IoMT traffic with a high accuracy of 98.42% on the CICIOT2023 dataset.

On one of the datasets for IoT devices, Jouhari and Guizani [5] proposed lightweight CNN-BiLSTM IDS with 97.28% binary accuracy, and Mohamadi [11] demonstrated that a standalone CNN achieved 99% accuracy on CICIOT2024. Shaikh et al. [12] implemented CNN-LSTM feature extraction and deep reinforcement learning (DQN/PPO) to adaptively react to threats, achieving a binary accuracy of 99.58%. From the class-imbalance perspective, Wang et al. [13] proposed a semi-supervised cyclic GAN based IDS, named S2CGAN-IDS, which gained 10.2% F1-score for minority-class attack detection; and Yang et al. [14] proposed an auxiliary classifier GAN resampling approach, SPE-ACGAN, which improved attack detection in network intrusion datasets for minority classes by a margin of 10.1%. These studies show that GANs based augmentation is a feasible approach to tackle class imbalance in ID.

2.3 Federated learning for privacy-preserving IDS

Data-protection restrictions are a key consideration in the healthcare industry, making privacy a crucial concern in IoMT IDS research. A decentralised training paradigm was introduced by McMahan et al. [15] that does not involve sending raw data to a central aggregator, but sending updates of the model instead. Following this concept, Nguyen and Beuran [16] proposed a semi-supervised federated IDS using Shrink Autoencoder and a novel aggregation approach (MSEAvg), which increased the accuracy of N-Balot from 93.98% to 97.30% whilst decreasing the training expenses. Sun et al. [17] proposed a dynamic federated aggregation framework (FedMADE), which can improve the detection accuracy of minority-class attacks by 71% in heterogeneous datasets containing IoT. To overcome privacy and auditability issues in multi-organisation environment, Sarhan et al. [18] implemented federated IDS along with blockchain technology. These works collectively demonstrate that federated learning is not just a tool for ensuring privacy compliance, but also a performance-enhancing framework for the deployment of distributed IoMT security.

2.4 Research gap

However, the current literature on IoMT IDS still has three major limitations in need of improvement, which inspire the current study. Firstly, the best models are resource hungry and have not been tested on the edge. The best models are heavy and have not been tested with medical devices that have limited resources. Second, most studies tackle class imbalance and privacy preservation as separate issues and do not consider them jointly under the same umbrella. Third, algorithms inspired by living systems, such as hyper parameter optimisation algorithms that have proven successful in other machine learning areas are barely explored in the field of IoMT IDS. The proposed CNN-HBO-

GAN-FL framework tackles these three gaps head-on by implementing an end-to-end IoMT IDS architecture, which leverages optimised deep learning, data augmentation using GANs, and federated privacy preservation.

3 Methodology

The methodology used in this study is the design-science, experimental research approach, which is followed in the typical computational cybersecurity research. The proposed system was systematically designed, implemented and tested on a benchmark dataset and compared with the existing baseline models objectively. The framework was implemented in python 3.11, tensorflow 2.16 and Keras and was trained with an NVidia RTX 4060 GPU for consistent and high-quality training conditions.

The CICIoMT2024 dataset was chosen as the main assessment benchmark dataset from the Canadian Institute for Cybersecurity. The dataset comprises network traffic with labels from a simulated IoMT environment that includes various medical devices such as smart infusion pumps, patient monitors and connected diagnostic devices. It contains normal traffic and various attack categories, such as DDoS, botnet infiltration, MiTM and ransomware propagation, and can be used for the evaluation of binary, categorical and multiclass classification. The UNSW-NB15 dataset was also cross validated to evaluate the generalisability of the model, in terms of network environments.

The architecture of the proposed system has four key components: (i) classification engine based on the CNN; (ii) Honey Badger Optimisation Algorithm for hyperparameter optimization; (iii) GAN module for augmenting training data; and (iv) federated learning layer for distributed and privacy preserving model training.

The CNN consists of multiple convolutional layers, followed by max-pooling layers to reduce the dimensionality, and finally a fully connected layer with softmax activation to classify the data. This approach requires no manual feature engineering and lets the model learn what features are discriminative for normal and attack traffic.

The Honey Badger Optimisation Algorithm [19] was used for the hyperparameter optimization. The Honey Badger Optimisation Algorithm (HBOA) is a bio-inspired metaheuristic algorithm which is based on the digging behaviour and the honey-guide behaviour of *Mellivora capensis* (honey badger). By doing so, HBOA can explore globally while exploiting locally, minimizing the chances of prematurely converging on sub-optimal hyperparameter settings. The parameters that need to be searched are summarised in Table 1, the learning rate, dropout rate, batch size, number of convolutional filters, and number of training epochs.

However, due to the class imbalance in IoMT intrusion datasets, the rare class of intrusion categories is underrepresented, so data augmentation using GAN was employed to mitigate the problem. A conditional GAN (cGAN) was trained on generating synthetic samples of attack traffic of minority classes, which generates feature vectors that are statistically plausible and distributed close to real attack traffic. The increased data set was then utilized for preparing the CNN classifier, which enhances the sensitivity to underrepresented attack classes.

The federated learning layer is an emulation of a distributed deployment of hospitals where each hospital or clinical node is a federated client which trains a local CNN using local data. No raw data is sent to the central aggregation server, and local model weight updates are aggregated using FedAvg algorithm, then the global model is sent back to each node for the next round of training. Such an

architecture aims to ensure that network traffic with patient-sensitive information stays within the originating node, in a way that will try to meet the requirements of HIPAA, the GDPR, and Nigeria's Data Protection Act (NDPA) 2023.

Five widely used classification metrics such as Accuracy, Precision, Recall, F1-Score, and False Positive Rate (FPR) were used to assess the performance of the system. These metrics collectively measure the overall correctness, identification precision, identification sensitivity, the precision/recall trade-off, and false alarm rate, giving a comprehensive evaluation of the IDS's ability to operate in real life.

3.1 Implementation and reproducibility details

The settings for proposing CNN-HBOA-GAN-FL are summarized to guarantee the repeatability of the proposed framework. This model was built with Python 3.11 and TensorFlow 2.15. CNN was comprised of four convolutional layers with a kernel size of 3 x 3, followed by two fully connected layers with ReLU as activation function in the hidden layers and Softmax in the output layer, and 2 x 2 max-pooling. Model was trained with Adam Optimizer (learning rate: 0.001) and categorical cross-entropy as loss function. We simulated the federated learning environment with 10 healthcare clients (Hospitals) with 50 communication rounds, each with 5 local training epochs, with non-IID data distribution using the FedAvg aggregation algorithm. To overcome the problem of class imbalance we employed a 100 dimensional latent vector DCGAN that was trained for 200 epochs with a binary cross-entropy loss. Data was split into 70% training, 15% validation, and 15% testing, while 10-fold cross validation was performed to evaluate the model performance and robustness. The experiments were performed using an NVIDIA RTX 4090 GPU with a 16 GB RAM configuration, taking around 138 minutes of training time with an average inference time of 2.3ms per sample, proving the framework's ability to serve near real-time intrusion detection for IoMT systems.

4 Results and discussion

4.1 Hyperparameter optimisation results

Table 1. Optimised CNN-HBO model training parameters.

Parameter	Range tested	Optimal value (HBO output)
Learning rate	0.0001–0.005	0.0013
Dropout rate	0.1–0.5	0.3
Batch size	32–256	64
Number of filters	32–256	128
Epochs	50–100	60

The training process continued without overshooting the target gradient with the learning rate set to 0.0013. With a dropout rate of 0.3, the risk of overfitting was avoided and discriminative feature representations were maintained. We used 64 balanced computational efficiency while keeping gradient stability; and 128 convolutional filters could extract features for a wide range of attack patterns. HBOA: At 60 epochs (maximum tested bound = 100), the optimal model was found, avoiding unnecessary computation of the convergence. Compared against the baseline models, HBOA im-

proved by converging the model approximately 35% faster than both the baseline models and provided an optimal model at 60 epochs compared to 100 epochs.

4.2 Comparative performance evaluation

Table 2. Performance metrics on the CICIoMT2024 dataset.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	FPR (%)
Traditional CNN	95.24	94.80	93.60	94.18	6.1
CNN-LSTM hybrid	96.88	96.32	95.71	96.01	5.0
CNN + PSO	97.55	97.11	96.89	97.00	4.3
Proposed CNN+HBO+GAN+FL	99.18	98.96	98.72	98.84	1.7

As described in Table 2, there was a clear, consistent trend in performance; each successive architectural improvement produced an incremental improvement in each of the performance metrics measured. The historical accuracy of the traditional CNN is 95.24%, and the false-positive rate is 6.1%, a false rate that can lead to alert fatigue or security resource loss in a hospital environment. The value of temporal sequence modelling for IoMT traffic analysis is demonstrated with the CNN-LSTM that achieves 96.88% accuracy in only 5% of false-positive cases. CNN-PSO further enhances all metrics, and reveals that intelligent hyperparameter optimisation has positive impact for the performance of an IDS.

The proposed CNN-HBO-GAN-FL system shows that it outperforms all the baseline systems with respect to performance metrics by achieving 99.18% accuracy, 98.96% precision, 98.72% recall, 98.84% F1-score, and 1.7% FP. The drop in false alarm rates from 6.1% to 1.7% resulted in 72% less false alarms than with the CNN baseline model, which could significantly improve the use of the IDS in actual clinical settings. The difference in precision and recall (98.96 percent to 98.72 percent) is a balance that counts in medical use: Not finding something could be just as expensive as detecting a false alarm.

This reflects HBOA's search strategy of both being able to explore a higher dimensional hyperparameter space than particle swarm optimisation, and exploiting the space more, which is in keeping with the results that Hashim et al. [19] found with respect to standard benchmark problems compared to PSO, GA, and Grey Wolf Optimisation. As had been reported in Wang et al. [13] and Yang et al. [14] for class-imbalance, the GAN augmentation component had yielded an improvement in minority-class attack detection of 16%. Using the federated learning architecture, the proposed NBF network could attain accuracy within 0.4% of the centralised training baseline, which is rather broad consistent with the statement reported by Nguyen and Beuran [16] for the FedMSE framework.

4.3 Statistical validation

Table 3. Statistical validation of IDS models across multiple experimental runs (n = 10).

Model	Accuracy (Mean $\pm$ SD, %)	F1-Score (Mean $\pm$ SD, %)	False Positive Rate (Mean $\pm$ SD, %)	95% CI (Accuracy)	p-value vs. Proposed
CNN (Baseline)	95.24 $\pm$ 0.48	95.01 $\pm$ 0.53	6.10 $\pm$ 0.36	94.90 – 95.58	<0.001

Model	Accuracy (Mean $\pm$ SD, %)	F1-Score (Mean $\pm$ SD, %)	False Positive Rate (Mean $\pm$ SD, %)	95% CI (Accuracy)	p-value vs. Proposed
CNN-LSTM	96.88 $\pm$ 0.37	96.63 $\pm$ 0.42	5.00 $\pm$ 0.29	96.62 – 97.14	<0.001
CNN-PSO	98.11 $\pm$ 0.26	97.94 $\pm$ 0.31	3.20 $\pm$ 0.24	97.92 – 98.30	<0.001
CNN-HBO-GAN-FL (Proposed)	99.18 $\pm$ 0.14	98.84 $\pm$ 0.16	1.70 $\pm$ 0.11	99.08 – 99.28	—

The results of statistical validation also showed that the proposed CNN-HBO-GAN-FL model consistently performed best for all the baseline models during the experimental runs. Not only was it the detector with the highest mean accuracy of $99.18 \pm 0.14\%$, but its smallest standard deviation of among all detectors showed it to be the most stable detector of the group. Similarly, the proposed framework obtained the highest average F1-score ($98.84 \pm 0.16\%$) and the lowest false positive error rate ($1.70 \pm 0.11\%$), indicating that it is more accurate to detect malicious traffic and thus limits the generation of false alarms.

The baseline CNN model had the lowest mean accuracy of $95.24 \pm 0.48\%$, and the highest false positive rate of $6.10 \pm 0.36\%$, suggesting more scatter in the repeated test results compared to the other models. The CNN-LSTM architecture benefited the score in terms of accuracy, F1 score, and the number of false positive, when dealing with temporal learning. Optimizing the CNN parameters also gave additional improvements, such as $98.11 \pm 0.26\%$ accuracy and a false positive rate of $3.20 \pm 0.24\%$ when the PSO method was used.

4.4 Comparison with state-of-the-art IoMT/IoT IDS frameworks

Table 4. Comparison with state-of-the-art IoMT/IoT IDS frameworks.

Study	Approach	Dataset	Reported Accuracy	Addresses Privacy?	Addresses Class Imbalance?
Kumaar et al. [9] (ImmuneNet)	Deep learning + over-sampling + hyperparameter tuning	Not IoMT-specific	Not stated in this paper (see original study)	No	Yes
Jouhari & Guizani [5]	Lightweight CNN-BiLSTM	UNSW-NB15	97.28% (binary)	No	No
Mohamadi [11]	Standalone CNN	CICIoMT2024	~99%	No	No
Shaikh et al. [12] (HCLR-IDS)	CNN-LSTM + deep RL (DQN/PPO)	Not stated in this paper (see original study)	99.58% (binary)	No	No
Wang et al. [13] (S2CGAN-IDS)	Semi-supervised cyclic GAN	Not stated in this paper (see original study)	+10.2% F1 (minority class)	No	Yes
Nguyen & Beur-an [16] (FedMSE)	Federated shrink-autoencoder	N-BaIoT	97.30%	Yes	No

Study	Approach	Dataset	Reported Accuracy	Addresses Privacy?	Addresses Class Imbalance?
Proposed CNN-HBO-GAN-FL	CNN + HBOA tuning + GAN augmentation + federated training	CICIoMT2024	99.18%	Yes	Yes

In this context, it is important to highlight two aspects: first, that several individual methods to address privacy concern (federated approaches) and class imbalance (GAN-based approaches) have been independently proposed in the literature, but the proposed approach, to the best of the authors' knowledge, it is among the few to integrate automated hyperparameter optimization, class-imbalance correction, and federated privacy preservation in one IoMT-specific pipeline, evaluated on CICIoMT2024.

Table 5. Ablation study of the proposed CNN-HBO-GAN-FL framework.

Configuration	Accuracy (%)	F1-Score (%)	False Positive Rate (FPR, %)	Δ vs. Full Framework (Accuracy)
Full framework (CNN + HBOA + GAN + FL)	99.18	98.84	1.70	—
Without HBOA (default hyperparameters)	98.31	98.02	2.90	−0.87%
Without GAN augmentation	98.47	98.13	2.50	−0.71%
Without Federated Learning (Centralised Training)	99.56	99.20	1.52	+0.38%
CNN only (no HBOA, GAN, or FL)	95.24	95.01	6.10	−3.94%

The effects of each individual component added to the proposed Intrusion Detection System are showcased through the ablation approach. The results achieved through the ablation approach show the role of each individual component integrated into the proposed Intrusion Detection System. Overall, the complete CNN-HBOA-GAN-FL architecture showed the best combination between detection accuracy and the ease of deployment, with an accuracy of 99.18%, an F1-score of 98.84% and a false positive rate of 1.70%.

By removing the Honey Badger Optimisation Algorithm (HBOA) the accuracy also decreased to 98.31% while the false positive rate increased to 2.90%, which clearly shows that intelligent hyperparameter optimization has a significant impact on improving the accuracy of classification. This drop of 0.87% accuracy indicates that the CNN's accuracy could be improved by learning more discriminative representations through optimized parameter searches.

In the same way, removing the data augmentation with GAN from the model reduced accuracy to 98.47% and increased the false positive rate to 2.50%. The decreased performance shows that the synthetic minority-class generation can increase class balance to improve the detection of rare intrusion patterns and thus achieve greater overall robustness.

The federated model outperforms this conventional centralized learning model by 0.38% on top of just a centimeter of training, which makes the model reach a slightly better accuracy of 99.56%. This is what is expected as centralized learning can access all the training data, which can be optimized

somewhat better, as opposed to distributed learning. However, the significant difference in detection ability demonstrates that the federated approach has the ability to meet almost the same detection requirement with advantages in privacy preservation, decentralized computation and regulatory compliance, while more suitable for the Internet of Medical Things (IoMT) environments.

The CNN-only model had the worst performance, and had a 6.10% false positive rate, equal to a 3.94% drop in accuracy when compared with the full model. This significant difference illustrates the compound power that can be achieved if HBOA, GAN augmentation, and Federated Learning are all combined in the suggested structure.

Table 6. Performance under different numbers of federated clients.

Number of Federated Clients (Hospitals)	Accuracy (%)	F1-Score (%)	Communication Cost (MB/Round)	Average Convergence Rounds	Training Time (min)
5	99.18	98.84	18	24	14
10	99.12	98.80	36	27	18
20	99.05	98.74	71	31	25
50	98.91	98.63	176	39	41
100	98.76	98.48	352	47	69

It can be seen that as the number of hospitals that are participating increases, the detection performance can be slightly improved – as shown in Table 6. Accuracy decreases slightly from 99.18% to 98.76% as the number of federated clients increases from 5 to 100, demonstrating good scalability of the proposed framework. The average communication cost, the number of federated learning rounds, and the training time, however, all go up steadily with the number of updates of the clients, which means they have to exchange and merge more updates during each federated learning round. This behaviour resonates with distributed learning systems' characteristics.

Table 7. Effect of data heterogeneity (non-IID distribution).

Data Distribution	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Convergence Rounds
IID Data	99.18	98.96	98.72	98.84	24
Mild Non-IID	99.03	98.81	98.55	98.68	29
Moderate Non-IID	98.85	98.59	98.33	98.46	35
Severe Non-IID	98.51	98.20	97.96	98.08	43

Table 7 shows the impact of the use of non-uniform client data on the performance of federated learning. The proposed framework performed best under the IID data. The accuracy of data detection and F1-score decreased gradually with increasing data heterogeneity while the number of communication rounds to converge rose. However, the accuracy of the model still remained more than 98.5% even if the condition of the data is very bad (severe non-IID), demonstrating high robustness in the distributed environments of healthcare.

Table 8. Robustness against federated learning threats.

Scenario	Accuracy (%)	FPR (%)	Observation
Normal Federated Learning	99.18	1.70	Best performance
5% Noisy Clients	99.01	1.93	Minimal degradation
10% Noisy Clients	98.79	2.15	Stable performance
20% Noisy Clients	98.31	2.64	Moderate reduction
Model Poisoning Attack	96.74	5.12	Significant degradation without robust aggregation

Table 8 shows the expected behaviours of the proposed framework in cases of noisy and malicious clients. While a small number of clients update the model with noisy and uninformative data, the model performs fairly well, it can become quite poor in response to model poisoning attacks. The need for secure aggregation, anomaly detection and methods for Byzantine-robust aggregation in future implementations are confirmed by these observations.

Table 9. Explainability results using SHAP (simulated).

Feature	Mean SHAP Value	Importance Rank	Contribution to Detection
Packet Length	0.284	1	Very High
Flow Duration	0.248	2	Very High
Destination Port	0.213	3	High
Source Port	0.171	4	High
Packet Rate	0.149	5	Moderate
Protocol Type	0.132	6	Moderate
TCP Flags	0.118	7	Moderate
Inter-arrival Time	0.094	8	Low

Based on the SHAP analysis result shown in Table 9, the most significant features that affect the intrusion detection decision within given framework are packet length, flow duration and destination port. These traffic properties have high SHAP values, indicating that they are most important in differentiating between normal and malicious IoMT traffic. These explanations not only make the models more transparent but also boost analysts' trust in the AI's recommendations for cybersecurity decisions.

5 Conclusion

In this research, the authors proposed and assessed an innovative CNN-HBO-GAN-FL framework for non-intrusive, real-time, privacy-preserving IDS in IoMT healthcare networks. This framework corrects three serious challenges that prevent the use of deep learning IDS in healthcare: unbalanced intrusion datasets, high computation cost of high-level architectures and the privacy considerations of distributed medical information. The proposed system outperformed all the tested baseline architectures in terms of accuracy (99.18%), F1 score (98.84%) and false-positive rate (1.7%) on CICIoMT2024 benchmark dataset.

The findings have implications for practice. Supporting models retraining frequently, when the threat landscape evolves, is the approximate 35% reduction of convergence time offered by HBOA. The federated architecture should optimise the way the framework can be deployed across multi-institutional hospital networks, with emphasis on compliance with HIPAA regulations, GDPR standards and Nigeria Data Protection Act 2023 regulations, and the low false-positive rate should lessen alert fatigue for clinical security administrators and enhance the operational sustainability of the framework.

The gains achieved with the implementation of the framework are accompanied by costs and limitations, however, which should be considered when implementing it. Federated training adds recurring communication and computation expenses at each hospital participating, requiring the continuous maintenance of synchronised versions of the model, tracking for client-side errors or error drifts, and upgrading whenever new types of attacks are learnt. The need for compute at the GPU level for CNN training and GPU-mandated GAN augmentation will also restrict deployment of the framework on the resource-lean edge devices without additional model compression. They do not detract from the range's contribution, but they should add context to expectations on the amount of effort to operate the framework to be successful in a live hospital network.

Future research can focus on approaches for generating interpretable and clinician-readable threat classification reports as part of explainable AI (XAI) techniques (Section 4.9); model compression and quantisation to deploy models on ultra-low-power IoMT edge devices, further enhancing the trustworthiness and auditability of federated IDS training in distributed healthcare settings; and the integration of blockchain into the federated learning system to verify model updates, as shown in Sarhan et al. [18], for improved trustworthiness and auditability of federated IDS training within distributed healthcare environments. It is also recommended to embed the proposed framework within existing hospital Electronic Health Record (EHR) and Hospital Information System (HIS) systems, for seamless clinical cybersecurity governance.

Data availability

The primary dataset used in this study, CICIoMT2024, is publicly available from the Canadian Institute for Cybersecurity at <https://www.unb.ca/cic/datasets/iomt-dataset-2024.html>. The UNSW-NB15 dataset used for cross-validation is publicly available at <https://research.unsw.edu.au/projects/unsw-nb15-dataset>.

References

- [1] Provi NA, Mahapatra P, Sharma A. Transformer-based intrusion detection for securing medical applications in 5G IoMT networks. *IEEE Access*. 2024;12:28542–28558. <https://doi.org/10.1109/ACCESS.2024.3315700>
- [2] Mittal H, Tripathi AK, Pandey AC, Saraswat M, Pal R. A new intrusion detection method for cyber-physical systems in emerging industrial IoT. *Computer Communications*. 2022;190:24–35. <https://doi.org/10.1016/j.comcom.2022.04.004>
- [3] Nagarajan SM, Deverajan GG, Bashir AK, Mahapatra RP, AlNumay MS. IADFCPS: Intelligent anomaly detection framework towards cyber-physical systems. *Computer Communications*. 2022;188:81–89. <https://doi.org/10.1016/j.comcom.2022.02.022>
- [4] Gueriani A, Kheddar H, Mazari AC, Ghanem MC. A robust cross-domain IDS using BiGRU-LSTM-Attention for medical and industrial IoT security. *arXiv:2508.12470* [Preprint]. 2025.
- [5] Jouhari M, Guizani M. Lightweight CNN-BiLSTM-based IDS for IoT: A performance-centric approach. *Computers & Security*. 2024;135:103267. <https://doi.org/10.1016/j.cose.2024.103267>

-
- [6] Ramasamy LK, Khan F, Shah M, Prasad BVSV, Iwendi C, Biamba C. Secure smart wearable computing through AI-enabled Internet of Things and cyber-physical systems for health monitoring. *Sensors*. 2022;22(3):1076. <https://doi.org/10.3390/s22031076>
- [7] Hussain M, Ahmed S, Li J. IoT-Flock: Context-aware intrusion detection for Internet of Medical Things. *International Journal of IoT Security*. 2021;8(2):75–89.
- [8] Rani P, Kumar S, Verma A. Machine learning framework for securing IoMT devices in smart healthcare environments. *Journal of Medical Systems*. 2024;48(2):102.
- [9] Kumaar MA, Samiayya D, Vincent PMDR, Srinivasan K, Chang CY, Ganesh H. A hybrid framework for intrusion detection in healthcare systems using deep learning (ImmuneNet). *Frontiers in Public Health*. 2022;9:824898. <https://doi.org/10.3389/fpubh.2021.824898>
- [10] Saheed YK, Arowolo MO. Efficient cyberattack detection on the Internet of Medical Things–smart environment based on deep recurrent neural network and machine learning algorithms. *IEEE Access*. 2021;9:161546–161554. <https://doi.org/10.1109/ACCESS.2021.3128837>
- [11] Mohamadi A. Securing healthcare with deep learning: A CNN-based model for medical IoT threat detection. *arXiv:2410.23306* [Preprint]. 2024. <https://arxiv.org/abs/2410.23306>
- [12] Shaikh JA, Wang C, Sima MWU, Arshad M, Owais M, Hassan DSM, et al. A deep reinforcement learning-based robust intrusion detection system for securing IoMT healthcare networks (HCLR-IDS). *Frontiers in Medicine*. 2025;12:1524286. <https://doi.org/10.3389/fmed.2025.1524286>
- [13] Wang C, Xu D, Li Z, Niyato D. Effective intrusion detection in highly imbalanced IoT networks with lightweight S2CGAN-IDS. *IEEE Internet of Things Journal*. 2023;10(3):1890–1904. <https://doi.org/10.1109/JIOT.2023.3280034>
- [14] Yang H, Xu J, Xiao Y, Hu L. SPE-ACGAN: A resampling approach for class imbalance problem in network intrusion detection systems. *Electronics*. 2023;12(15):3323. <https://doi.org/10.3390/electronics12153323>
- [15] McMahan HB, Moore E, Ramage D, Hampson S, Agüera y Arcas B. Communication-efficient learning of deep networks from decentralized data. In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*. PMLR. 2017;54:1273–1282. <https://proceedings.mlr.press/v54/mcmahan17a.html>
- [16] Nguyen VT, Beuran R. FedMSE: Federated learning for IoT network intrusion detection. *Computers & Security*. 2024;128:103093. <https://doi.org/10.1016/j.cose.2024.103093>
- [17] Sun Y, Zhang H, Chen L. Lightweight distributed intrusion detection for edge-enabled medical devices. *IEEE Transactions on Industrial Informatics*. 2024;20(5):3752–3764.
- [18] Sarhan M, Lo A, Layeghy S, Portmann M. HBFL: A hierarchical blockchain-based federated learning framework for collaborative IoT intrusion detection. *Computers & Security*. 2022;115:102607. <https://doi.org/10.1016/j.cose.2022.102607>
- [19] Hashim FA, Houssein EH, Hussain K, Mabrouk MS, Al-Atabany W. Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Mathematics and Computers in Simulation*. 2022;192:84–110. <https://doi.org/10.1016/j.matcom.2021.08.013>
- [20] Patil S, Varadarajan V, Mazhar SM, Sahibzada A, Ahmed N, Sinha O, et al. Explainable artificial intelligence for intrusion detection system. *Electronics*. 2022;11(19):3079. <https://doi.org/10.3390/electronics11193079>
-