EdgeShield: a robust and agile cybersecurity architecture for the internet of medical things

Anass Misbah¹, Anass Sebbar², Imad Hafidi¹

¹IPIM Laboratory, National School of Applied Sciences, Sultan Moulay Slimane University, Beni Mellal, Morocco ²International University of Rabat, ESIN, TICLab, Rabat, Morocco

Article Info ABSTRACT

Article history:

Received Feb 14, 2025 Revised May 13, 2025 Accepted May 26, 2025

Keywords:

Data processing Dimensionality reduction Feature selection Federated learning Internet of medical things security Intrusion detection systems Principal component analysis We present EdgeShield, a lightweight pipeline that streamlines internet of medical things (IoMT) traffic analysis by pairing aggressive dimensionality-reduction with federated model aggregation. It employs systematic preprocessing, advanced feature selection, and robust sampling to reduce computational overhead while enhancing performance. Through feature engineering techniques such as principal component analysis (PCA), targeted feature selection, and embedding methods, EdgeShield reduces dataset dimensionality by 96%, enabling near real-time detection and prevention of cyber attacks on resource-constrained edge devices. To harden the IoMT perimeter, EdgeShield trains ten lightweight edge models in just 54s and merges their parameters into a single global classifier with negligible extra delay. This method requires no additional training or predictions, thus accelerating deployment. Additionally, by using a compact dataset with five top-performing features and PCA with two components, EdgeShield consistently achieves accuracy levels exceeding 99.2% for individual edge models and the consolidated global model. With a built-in continuous improvement loop, EdgeShield dynamically adapts to emerging data patterns and operational conditions, driving substantial advancements in IoMT ecosystem management. This approach delivers both rapid machine learning model deployment and robust cyber attack detection, illustrating its potential to revolutionize IoMT security and elevate healthcare data integrity.

This is an open access article under the <u>CC BY-SA</u> license.



1128

Corresponding Author:

Anass Misbah IPIM Laboratory, National School of Applied Sciences, Sultan Moulay Slimane University Beni-Mellal 23000, Morocco Email: anassmisbah@gmail.com

1. INTRODUCTION

Internet of medical things (IoMT) devices now enable clinicians to monitor patients continuously and in real time, opening new doors for personalised care [1], [2]. Despite offering considerable benefits such as improved diagnostic accuracy and cost-effective care, the highly connected IoMT infrastructure also introduces major challenges. Chief among these are data security threats, limited computational resources on edge devices, and increasing volumes of medical information that must be processed promptly [3], [4]. Cyberattacks on IoMT systems can jeopardize patient privacy, disrupt vital services, and even endanger lives [5], [6].

In response to these challenges, we propose EdgeShield, a comprehensive framework designed to streamline IoMT data processing while enhancing cyber resilience. This proposed approach is applied to CICIoMT2024 dataset [7]. EdgeShield integrates systematic preprocessing, advanced feature selection (FS),

controlled sampling, and a novel model aggregation technique to bolster performance and efficiency. Central to its design is an aggressive dimensionality reduction pipeline: targeted FS methods and principal component analysis (PCA) [8]-[10] together reduce the dataset's dimensionality by approximately 96%, thereby lowering the computational burden and enabling real-time analysis on resource-constrained edge devices.

To further strengthen cyber defense, EdgeShield trains 10 edge models in a mere 54 seconds, consolidating them into a unified global model with near-zero additional latency. By leveraging a compact dataset of five top-performing features and two PCA components, both individual edge models and the aggregated global model consistently achieve accuracy rates exceeding 99.2%. This synergy of rapid dimensionality reduction, real-time analytics, and robust aggregation paves the way for more secure, adaptive, and efficient IoMT ecosystems.

The remainder of this paper is structured as follows. Section 2 provides an overview of related work in IoMT security and dimensionality reduction approaches. Section 3 outlines the EdgeShield methodology and its key components, while section 4 concludes with a discussion of the results, open challenges, and directions for future research.

2. METHOD

2.1. Related work

This section highlights the role of PCA in diverse fields, the advantages of combining PCA with FS, and related frameworks that address security and efficiency in IoMT systems.

2.1.1. PCA in various domains

Across disciplines from power-grid monitoring to image recognition, researchers rely on PCA to compress high-dimensional data while preserving most of the signal. For example, the work of [8] combined PCA with support vector machines (SVM)-assisted neural networks and cut state-of-charge prediction error for LiFePO4 batteries to a very low value, while another study used PCA alongside gradient tree boosting to predict biocrude oil yields with near-perfect correlation [9]. Similarly, PCA-driven dimensionality reduction enhanced freshness detection of fruits and vegetables [10], optimized data aggregation in IoT networks [11], and improved feature learning in a deep kernel PCA framework [12]. Its broad applicability—spanning coal outburst prediction [13], lithium exploration [14], and hotel occupancy forecasting [15] underscores PCA's versatility in boosting model accuracy and interpretability.

2.1.2. Combining FS with PCA

FS removes less informative attributes from high-dimensional datasets, thereby reducing noise and improving model performance [16]. When combined with PCA, FS ensures that only the most relevant features contribute to the principal components, yielding a more efficient representation of the data. For instance, a mutual-information-based technique has been proposed for better intrusion detection in IoMT systems [17], whereas distance correlation has been applied to random forest (RF) models to handle high-dimensional, non-linear data [18]. Empirical evidence further supports the synergy of PCA and FS for improved accuracy in classification tasks [19]-[21], especially in resource-constrained environments like IoMT. Moreover, carefully orchestrating FS before PCA has been shown to enhance clustering and reduce computational overhead [22].

2.1.3. Similar approaches in IoMT security

Numerous frameworks address IoMT security from different angles. For example, a privacy-aware system integrates Federated Learning and eXplainable AI for intrusion detection [23], while another focuses on machine learning algorithms to safeguard IoT-based healthcare data in real time [24]. IoMT-SAF introduces an ontological method to systematically evaluate device vulnerabilities and recommend tailored defense strategies [25]. Although these solutions mitigate various security risks, most still face challenges in handling large, heterogeneous datasets and ensuring low-latency performance—critical factors for IoMT environments. The proposed EdgeShield framework (detailed in subsequent sections) aims to fill these gaps by leveraging an efficient dimensionality reduction pipeline, model aggregation with minimal overhead, and high-accuracy threat detection.

2.2. Main approach

EdgeShield provides a streamlined yet effective framework for processing internet of medical things (IoMT) data, balancing high accuracy with low computational overhead on resource-constrained edge devices.

1130

This section details each step of the pipeline—ranging from optional preprocessing to final model aggregation—and explains how the framework's components fit together.

2.2.1. Overall framework

Figure 1 offers an overview of EdgeShield's workflow. Data first enters a (potentially optional) preprocessing phase. From there, the system applies FS techniques, performs controlled sampling, and executes PCA to reduce dimensionality. Multiple lightweight models (simulating IoMT edge devices) are trained on these reduced datasets, and their trained parameters are aggregated in a fog or cloud environment to form a robust global model. Finally, the aggregated model returns to edge devices, along with updated security rules, closing the continuous improvement loop.



Figure 1. EdgeShield global architecture: from optional data preprocessing through FS, sampling, PCA-based reduction, local model training at the edge, and aggregation into a global model

2.2.2. Data preprocessing

Although CICIoMT2024 is generally delivered in a well-structured format, data preprocessing may still be relevant when working with raw or diverse IoMT data sources. The key steps are:

- Cleaning: removes noise, outliers, or inconsistent entries.
- Scaling: standardizes feature magnitudes (e.g., via z-score normalization).
- Normalization: maps features onto a uniform scale, often [0, 1].
- Harmonization: aligns disparate data from multiple sources, ensuring coherent structure.

In many IoMT applications, these steps significantly reduce data heterogeneity, thereby improving subsequent modeling. For CICIoMT2024, these tasks are minimal because the dataset arrives precleaned.

2.2.3. Feature selection

Next, we identify and retain only the most influential features—a crucial step to reduce computational cost without harming accuracy. Figure 2 compares four primary methods:

- RF feature importance: evaluates how often each feature contributes to reducing impurity across decision trees as shown in Figure 2(a).
- Mutual information (MI): quantifies the shared information between a given feature and the target label as shown in Figure 2(b).
- Gradient boosting (GB): uses feature importance metrics derived from boosting iterations as shown in Figure 2(c).
- PCA loadings: determines contributions of original features to principal components (though PCA is typically an unsupervised method) as shown in Figure 2(d).



Figure 2. Feature importance rankings from four selection approaches: (a) RF, (b) MI, (c) GB, and (d) PCA loadings (down-right) consistently highlight inter-arrival time (IAT) and related network behavior features as top contributors to attack detection

Table 1 further compares their accuracy and computational cost on CICIoMT2024. Although mutual information can yield slightly higher accuracy, it demands significantly more runtime. RF strikes a practical balance, consistently achieving near-optimal results. We ultimately keep the top five features from whichever method performed best under time/accuracy constraints (in many experiments, this was RF feature importance).

Table 1. Comparison of FS methods on CICIoMT2024									
FS method	Top 10 acc.	Time (Top 10)	Top 5 acc.	Time (Top 5)					
RF importance	0.9972	23m 9.9s	0.9960	10m 30s					
PCA loadings	0.8920	66m 61s	0.8326	39m 33s					
Gradient boosting (10 k)	0.9793	12m 39s	0.9789	26.8s					
Mutual info	0.9979	1010m 22.9s	0.9976	96m 29.2s					

EdgeShield: a robust and agile cybersecurity architecture for the internet of medical things (Anass Misbah)

2.2.4. Data sampling and PCA

Random sampling, since CICIoMT2024 is fully labeled and already representative, we sample 100–1000 rows to manage computational overhead. This approach balances data diversity with speed, especially on edge devices where memory and processing power are scarce. Active learning [26] though beneficial for unlabeled or partially labeled datasets is not mandatory here.

Dimensionality reduction over PCA, PCA rotates the original axes so that the resulting components are orthogonal and capture descending amounts of variance. By retaining just the top components (two in our experiments), we drastically lower dimensionality while preserving key variance. Figures 3(a) and (b), Figures 4(a) and (b) show how PCA consistently enhances accuracy across different sample sizes, particularly when combined with a carefully chosen subset of features (top five).



Figure 3. Sampling size vs accuracy of (a) all features and (b) all features + PCA. PCA stabilizes performance, especially as the dataset size varies from 100 to 1000



Figure 4. Sampling size vs accuracy using: (a) top five features and (b) top five + PCA. This further reduces noise and can approach near-perfect detection rates

2.2.5. Edge model training and performance

To simulate ten IoMT edge devices, we split the dataset into ten balanced subsets, each retaining approximately 1% of labeled instances for each attack type and benign traffic. This design addresses real-world IoMT scenarios where each device sees only a fraction of overall traffic. We then train lightweight RF (or XGBoost) classifiers on these subsets. Figure 5 shows that despite minimal training time (under a minute in total), each local model achieves strong accuracy.



Figure 5. Accuracies of the ten local edge models (each subset covers 1% of dataset labels). The uniform performance underscores the representativity of each sub-dataset

2.2.6. Global model construction and aggregation

Once trained, each local model's parameters are transferred to a fog or cloud node for further evaluation and consolidation into a single RF ensemble. Table 2 outlines four sub-model experiments on a 5,000-row sample. Combining PCA with the top five features (Exp. 4) yields a high accuracy of 0.9984, slightly outperforming the all-features approach.

Table 2.	Comparison	of Sub-model	experiments	(5.000-row	subset.	train/test)
14010 2.	companioon	or buo mouer	experimento	(5,000 10	540500,	traing (CSC)

Exp.	Features	PCA	Model	Sub-models	Meta-acc.
1	All feats	No	RF	10	0.9664
2	All feats	Yes	RF	10	0.9978
3	All feats	Yes	XGB	10	0.9915
4	Top 5	Yes	RF	10	0.9984

Aggregation mechanics, EdgeShield assembles the global model by merging the decision trees learned on each device into one RF ensemble. Because no retraining is needed, the step adds virtually zero latency while raising accuracy, as the consolidated classifier benefits from the diversity of local edge models. Privacy is maintained, as raw data remains on each device, only the learned parameters (tree estimators) are shared.

2.2.7. Global model validation and adjustment

After aggregation, the meta-model is subjected to various data range tests (e.g., 100 to 1000 samples) to confirm stability and generalization. Figures 6(a) and (b), Figures 7(a) and (b) depict how PCA usage, FS strategy, and chosen model type (RF vs XGB) affect final accuracy. We observe that PCA + top-five FS can surpass 99% accuracy under multiple configurations. A summary is given in Table 3.



Figure 6. Aggregated RF meta-model: (a) with PCA vs (b) without PCA using all features. PCA consistently yields higher and more stable accuracy

EdgeShield: a robust and agile cybersecurity architecture for the internet of medical things (Anass Misbah)



Figure 7. The RF-based approach can slightly outperform XGB in certain scenarios: (a) XGBoost with PCA on all features and (b) RF with PCA on top-5 features

				0
Exp.	Acc. Range	(200-400 samples)	(400-600 samples)	(+600 samples)
1	0.92-0.98	>0.95	>0.95	>0.97
2	0.95-0.99	>0.98	>0.98	Up to 0.999
3	0.95-0.99	>0.98	>0.98	Up to 0.998
4	0.9825-0.9990	>0.9925	>0.9925	>0.9990

Table 3. Global meta-model accuracy results for various data ranges

2.2.8. Continuous improvement

EdgeShield incorporates a self-learning loop to sustain reliability in evolving IoMT environments. As new data accumulates, each edge device updates its local model, whose parameters are then merged into an upgraded global classifier. This iterative flow aids in:

- Adaptation: coping with emerging attack patterns or changing traffic behaviors.

- Scalability: supporting more devices or expanded datasets as IoMT networks grow.
- Efficiency: preserving privacy by never sharing raw data and only exchanging tree estimators.

3. RESULTS AND DISCUSSION

EdgeShield compresses IoMT traffic into just five discriminative features and two principal components, allowing each of ten edge devices to finish training in under a minute yet still exceed 99% local accuracy; when these lightweight RF models are federated, meta-accuracy climbs to 0.998 while privacy is preserved because raw data never leave the device. The combination of aggressive feature pruning and PCA, rather than the specific base learner, drives the gain, yielding performances that remain above 95% even on 200-row subsets and scale to near perfect detection as sample size grows. Compared with recent federated IoMT intrusion-detection systems that top out at 95–98% accuracy [23], [24], EdgeShield improves the state-of-the-art by 1–4 percentage points and compresses end-to-end processing time from hours to minutes, providing a practical, real-time and privacy-aware defence for resource-constrained medical devices.

4. CONCLUSION

By uniting fast FS, two-component PCA and a privacy-preserving aggregation of ten edge classifiers, EdgeShield cuts compute load yet still surpasses 99% detection accuracy. Its modular design allows for continuous updates, making it well-suited to adapt as new threats emerge or IoMT device behaviors shift. Looking ahead, future work could explore integrating explainable AI (XAI) techniques to increase transparency in decision-making, applying more advanced sampling or active learning strategies to reduce labeling costs, and extending model aggregation approaches beyond RF for improved scalability. Such developments stand to further strengthen EdgeShield's role as a secure, efficient, and adaptable framework in next-generation healthcare systems.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	С	Μ	So	Va	Fo	Ι	R	D	0	Ε	Vi	Su	Р	Fu
Anass Misbah					\checkmark					\checkmark	\checkmark	\checkmark		
Anass Sebbar														
Imad Hafidi										\checkmark		\checkmark		
 C : Conceptualization M : Methodology So : Software Va : Validation Fo : Formal Analysis 		I R D O E	: Investigation : Resources : Data Curation : Writing - Original Draft : Writing - Paviaw & Editing						Vi Su P Fu	: Visu : Sup : Proje : Fun	ualizatio ervision ect Adr ding Ad	on 1 ninistra cquisiti	ation ion	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [A Misbah], upon reasonable request.

REFERENCES

- M. Sigala, A. Beer, L. Hodgson, and A. O'Connor, "Big data for measuring the impact of tourism economic development programmes: A process and quality criteria framework for using big data," *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications*, pp. 57–73, 2019, doi: 10.1007/978-981-13-6339-9.4.
- [2] G. Nguyen et al., "Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey," Artificial Intelligence Review, vol. 52, no. 1, pp. 77–124, Jun. 2019, doi: 10.1007/s10462-018-09679-z.
- [3] R. Dwivedi, D. Mehrotra, and S. Chandra, "Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review," *Journal of Oral Biology and Craniofacial Research*, vol. 12, no. 2, pp. 302–318, 2022, doi: 10.1016/j.jobcr.2021.11.010.
- [4] J. P. A. Yaacoub, H. N. Noura, O. Salman, and A. Chehab, "Ethical hacking for IoT: Security issues, challenges, solutions and recommendations," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 280–308, 2023, doi: 10.1016/j.iotcps.2023.04.002.
- [5] Gerodimos, L. Maglaras, M. A. Ferrag, N. Ayres, and I. Kantzavelou, "IoT: Communication protocols and security threats," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 1–13, 2023, doi: 10.1016/j.iotcps.2022.12.003.
- [6] M. L. Hernandez-Jaimes, A. Martinez-Cruz, K. A. Ramírez-Gutiérrez, and C. Feregrino-Uribe, "Artificial intelligence for IoMT security: A review of intrusion detection systems, attacks, datasets and Cloud–Fog–Edge architectures," *Internet of Things (Netherlands)*, vol. 23, 2023, doi: 10.1016/j.iot.2023.100887.
- [7] S. Dadkhah, E. C. P. Neto, R. Ferreira, R. C. Molokwu, S. Sadeghi, and A. Ghorbani, "CICIOMT2024: Attack Vectors in Healthcare devices-A Multi-Protocol Dataset for Assessing IoMT Device Security," *Journal of Computer Communications*, 2024.
- [8] C. Mehta, A. V. Sant, and P. Sharma, "SVM-assisted ANN model with principal component analysis based dimensionality reduction for enhancing state-of-charge estimation in LiFePO4 batteries," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 8, 2024, doi: 10.1016/j.prime.2024.100596.
- [9] T. Katongtung, T. Onsree, K. Y. Tippayawong, and N. Tippayawong, "Prediction of biocrude oil yields from hydrothermal liquefaction using a gradient tree boosting machine approach with principal component analysis," *Energy Reports*, vol. 9, pp. 215–222, 2023, doi: 10.1016/j.egyr.2023.08.079.
- [10] Y. Yuan and X. Chen, "Vegetable and fruit freshness detection based on deep features and principal component analysis; Vegetable and fruit freshness detection," *Current Research in Food Science*, vol. 8, 2024, doi: 10.1016/j.crfs.2023.100656.
- [11] Bajpai, H. Verma, and A. Yadav, "Optimizing data aggregation and clustering in Internet of things networks using principal component analysis and Q-learning," *Data Science and Management*, vol. 7, no. 3, pp. 189–196, 2024, doi: 10.1016/j.dsm.2024.02.001.
- [12] F. Tonin, Q. Tao, P. Patrinos, and J. A. K. Suykens, "Deep Kernel Principal Component Analysis for multi-level feature learning," *Neural Networks*, vol. 170, pp. 578–595, 2024, doi: 10.1016/j.neunet.2023.11.045.
- [13] C. Fan, X. Lai, H. Wen, and L. Yang, "Coal and gas outburst prediction model based on principal component analysis and improved support vector machine," *Geohazard Mechanics*, vol. 1, no. 4, pp. 319–324, 2023, doi: 10.1016/j.ghm.2023.11.003.
- [14] S. Esmaeiloghli, A. Lima, and B. Sadeghi, "Lithium exploration targeting through robust variable selection and deep anomaly detection: An integrated application of sparse principal component analysis and stacked autoencoders," *Geochemistry*, 2024, doi: 10.1016/j.chemer.2024.126111.
- [15] D. Contessi, L. Viverit, L. N. Pereira, and C. Y. Heo, "Decoding the future: Proposing an interpretable machine learning model for

EdgeShield: a robust and agile cybersecurity architecture for the internet of medical things (Anass Misbah)

hotel occupancy forecasting using principal component analysis," International Journal of Hospitality Management, vol. 121, 2024, doi: 10.1016/j.ijhm.2024.103802.

- [16] Guyon I and Elissef A, "An Introduction to Variable and Feature Selection," *Journal of machine learning research*, vol. 3, pp. 1157–1182, 2003.
- [17] M. Alalhareth and S. C. Hong, "An Improved Mutual Information Feature Selection Technique for Intrusion Detection Systems in the Internet of Medical Things," *Sensors*, vol. 23, no. 10, 2023, doi: 10.3390/s23104971.
- [18] S. Ratnasingam and J. Muñoz-Lopez, "Distance Correlation-Based Feature Selection in Random Forest," *Entropy*, vol. 25, no. 9, 2023, doi: 10.3390/e25091250.
- [19] V. Rupapara, F. Rustam, A. Ishaq, E. Lee, and I. Ashraf, "Chi-Square and PCA Based Feature Selection for Diabetes Detection with Ensemble Classifier," *Intelligent Automation & Soft Computing*, vol. 36, no. 2, pp. 1931–1949, 2023, doi: 10.32604/iasc.2023.028257.
- [20] Tripathi and P. Rani, "PCA-Based Feature Selection and Hybrid Classification Model for Speech Emotion Recognition," *Lecture Notes in Networks and Systems*, vol. 703 LNNS, pp. 347–353, 2023, doi: 10.1007/978-981-99-3315-0_26.
- [21] P. A. Vysakh and P. Mayank, "Solar Flare Prediction and Feature Selection Using a Light-Gradient-Boosting Machine Algorithm," *Solar Physics*, vol. 298, no. 11, 2023, doi: 10.1007/s11207-023-02223-5.
- [22] J.-S. Dessureault and D. Massicotte, "Feature selection or extraction decision process for clustering using PCA and FRSD," *Machine Learning*, 2021.
- [23] A. Si-ahmed, M. A. Al-Garadi, and N. Boustia, "Explainable Machine Learning-Based Security and Privacy Protection Framework for Internet of Medical Things Systems," *Cryptography and Security*, Mar. 2024.
- [24] S. Pirbhulal, N. Pombo, V. Felizardo, N. Garcia, A. H. Sodhro, and S. C. Mukhopadhyay, "Towards machine learning enabled security framework for iot-based healthcare," *Proceedings of the International Conference on Sensing Technology*, ICST, vol. 2019–December, 2019, doi: 10.1109/ICST46873.2019.9047745.
- [25] F. Alsubaei, A. Abuhussein, V. Shandilya, and S. Shiva, "IoMT-SAF: Internet of Medical Things Security Assessment Framework," *Internet of Things (Netherlands)*, vol. 8, 2019, doi: 10.1016/j.iot.2019.100123.
- [26] Z. Zhao, Y. Shang, J. Wu, and Y. Yan, "Dataset Quantization with Active Learning Based Adaptive Sampling," in *Lecture Notes in Computer Science*, 2025, pp. 346–362. doi: 10.1007/978-3-031-73027-6_20.

BIOGRAPHIES OF AUTHORS





Anass Misbah D R ecceived the Engineering degree from the École Nationale Supérieure des Ingénieurs de Bourges, France, in 2008. He has been working in software engineering ever since and is currently pursuing the Ph.D. degree at the IPIM Laboratory, National School of Applied Sciences, Sultan Moulay Slimane University, Morocco. His research interests include cybersecurity, Internet of Things (IoT), machine learning, deep learning, and federated learning. His contributions in the current study encompass the design of advanced machine learning algorithms for IoT network security, the development of secure data transmission protocols for resource-constrained IoT devices, and the application of deep/federated learning techniques to enhance intrusion detection and threat mitigation. He can be contacted at email: anassmisbah@gmail.com.



Imad Hafidi (D) (S) ((S) (S) (S) ((S) (S) ((S) (S) ((S) ((S)