Journal of Information and Organizational Sciences
Volume 49, Number 2 (2025)
Journal homepage: jios.foi.hr

JIOS

DOI: 10.31341/jios.49.2.11
Open Access

UDC 004.056:004.8:519.6:61
Original Scientific Paper

A Hybrid IG-PCA and Machine Learning Approach for Accurate Intrusion Detection in IoMT with Imbalanced Data

Willy Riyadi[1*], Kurniabudi[1], Jasmir[1], Yudi Novianto[1], Desi Kisbianty[1] and Xaverius Sika[1]

[1]Faculty of Computer Science, Universitas Dinamika Bangsa, Jambi, Indonesia
*Correspondence: willyriyadi@unama.ac.id

PAPER INFO

Paper history:
Received 22 August 2025
Accepted 02 December 2025

Citation:
Riyadi, W., Kurniabudi, Jasmir, Novianto, Y., Kisbianty, D. & Sika, X. (2025). A Hybrid IG-PCA and Machine Learning Approach for Accurate Intrusion Detection in IoMT with Imbalanced Data. In Journal of Information and Organizational Sciences, vol. 49, no. 2, pp. 345-359

ABSTRACT

The rapid growth of the Internet of Medical Things (IoMT) has introduced critical cybersecurity challenges, highlighting the need for robust and accurate intrusion detection systems (IDS). This study presents a hybrid machine learning (ML) framework to strengthen intrusion detection in IoMT networks using the CIC-IoMT2024 dataset. The framework combines Information Gain (IG) and Principal Component Analysis (PCA) for feature selection and dimensionality reduction, while SMOTEENN and SMOTETomek are applied to address severe class imbalance. The processed data are classified using Random Forest (RF), K-Nearest Neighbors (KNN), XGBoost (XGB), Multi-Layer Perceptron (MLPC), and Logistic Regression (LR), with hyperparameters optimized through Bayesian Optimization. Performance is evaluated using Accuracy, Precision, Recall, F1-Score, and AUC. Experimental results reveal that the optimized XGB classifier with SMOTEENN achieves a peak accuracy of 99.811%. This top-tier performance surpasses several existing benchmarks, validating the effectiveness of integrating IG-PCA with advanced resampling and optimization strategies. This work contributes a lightweight, scalable, and highly accurate IDS, offering a practical and efficient solution for enhancing security in resource-constrained, next-generation medical IoT systems.

Keywords: Internet of Medical Things, Intrusion Detection Accuracy, IG-PCA, Machine Learning, Bayesian Optimization

1. Introduction

The proliferation of the Internet of Medical Things (IoMT) has catalyzed a paradigm shift in healthcare, enabling innovations such as remote patient monitoring, automated diagnostics, and interconnected smart health services (Alturki et al., 2025; Razdan & Sharma, 2022; Wang et al., 2021). This digital transformation holds immense potential for improving patient outcomes and operational efficiency. However, the increasing connectivity and inherent heterogeneity of IoMT devices have also introduced significant security vulnerabilities (Ahmed et al., 2024; Binbusayyis et al., 2022; Mathkor et al., 2024). Operating in resource-constrained environments and often lacking standardized security protocols, these devices are prime targets for cyberattacks, making data privacy, secure communication, and real-time threat detection paramount concerns (Papaioannou et al., 2022; Zachos et al., 2021).

To mitigate these threats, the development of robust Intrusion Detection Systems (IDS) specifically designed for IoMT ecosystems has become a critical area of research (Alalhareth & Hong, 2024; Ibrahim & Al-Wadi, 2024). m (ML)-based IDS have demonstrated considerable promise by learning network behavior to

intelligently classify legitimate and malicious traffic (Berguiga et al., 2025; Chaganti et al., 2022). Nonetheless, their practical implementation is hindered by several persistent challenges, including the 'curse of dimensionality' from high-dimensional data, severe class imbalance where attack data is sparse, and the stringent requirement for lightweight models that can operate in real-time on devices with limited computational power (Husain et al., 2025).

In response, the literature presents several strategies. To manage data complexity, hybrid methods combining Information Gain (IG) for feature selection and Principal Component Analysis (PCA) for dimensionality reduction have been explored to enhance classification accuracy while minimizing computational load (Kumar et al., 2022; Nasir et al., 2022; Odhiambo Omuya et al., 2021; R.M. et al., 2020). To counteract class imbalance, over-sampling and hybrid sampling techniques like SMOTEENN and SMOTETomek are widely used to improve model sensitivity towards minority attack classes (Alsharaiah et al., 2025; Bouke et al., 2024; Sarkar et al., 2024; Talukder et al., 2024). Recent advanced models have achieved high accuracy; for instance, systems integrating deep learning with blockchain have reported accuracies of 99.7% across 18 attack types (Abdiwi, 2024). while federated learning approaches have reached 97.31%. Similarly, combining SMOTE with classical ML models has yielded accuracies up to 99.2% (Mohsin & Jony, 2024).

Beyond the technical vulnerabilities, security breaches in IoMT ecosystems pose direct threats to organizational stability and patient safety. A successful cyberattack could lead to the compromise of sensitive patient health information (PHI), disruption of critical clinical workflows, or even manipulation of life-sustaining medical devices, causing irreparable harm to patients and catastrophic reputational damage to healthcare providers (Chuma & Ngoepe, 2022). Despite these advancements, a significant research gap persists. Many existing solutions focus on maximizing a single metric like accuracy, often at the cost of computational efficiency, model interpretability, or scalability—qualities that are non-negotiable for real-world IoMT deployment. A holistic framework that systematically integrates and optimizes preprocessing, feature engineering, and model tuning to create a balanced, lightweight, and effective IDS is still needed. This study addresses this gap by proposing a comprehensive ML-based framework to optimize intrusion detection, validated on the contemporary CIC-IoMT 2024 dataset (Dadkhah et al., 2024). Our approach prioritizes a pragmatic balance between high detection performance and computational feasibility, ensuring its suitability for resource-constrained IoMT infrastructures.

The major contributions of this research are as follows:

1. Development of an Optimized IDS Framework: We propose and evaluate an end-to-end intrusion detection framework that synergistically integrates data preprocessing, feature engineering (using IG and PCA), and a comparative analysis of advanced data balancing techniques (SMOTEENN and SMOTETomek). This creates a robust and computationally efficient system specifically tailored for IoMT environments.
2. Systematic Hyperparameter Optimization: We employ Bayesian Optimization (BO) for automated and efficient hyperparameter tuning of multiple ML classifiers. This approach systematically enhances model performance and reduces the manual effort and computational cost associated with traditional grid-search or random-search methods.
3. Comprehensive Performance Validation**:** The proposed framework is rigorously validated on the recent and highly relevant CIC-IoMT 2024 dataset. This provides a robust and up-to-date benchmark of its effectiveness against modern cyber threats, assessing performance across multiple metrics, including accuracy, precision, recall, F1-score, and Area Under Curve (AUC).

We present a comprehensive performance evaluation of our proposed method against baseline models and other optimization approaches from the literature. This demonstrates that our framework achieves an enhanced balance between high detection accuracy and practical computational efficiency.

2. Related work

The growing complexity and interconnectivity of IoMT systems have prompted extensive research into developing secure, efficient, and intelligent IDS. A significant body of literature has focused on leveraging ML to identify malicious behavior in healthcare-related IoT environments. Traditional signature-based IDS approaches are often ineffective in IoMT contexts due to their inability to detect novel or evolving threats. Consequently, ML-based IDS have gained traction due to their adaptive learning capabilities and ability to generalize from data. Studies such as those by (Chaganti et al., 2022) and (Ibrahim & Al-Wadi, 2024) have demonstrated the feasibility of ML models like Decision Trees, Random Forests, and Support Vector Machines in IoMT-specific datasets. These models achieved high detection accuracy but often required further optimization for real-time deployment due to their computational overhead.

Efficient feature engineering plays a pivotal role in enhancing IDS performance, particularly when handling high-dimensional IoMT data. To improve detection speed and simplify model complexity, several studies have adopted hybrid feature reduction techniques. For instance, (Sayed et al., 2023) and (Kumari & Jain, 2024), combined Information Gain (IG) and Principal Component Analysis (PCA), achieving higher classification accuracy and reduced training time. Similarly, (Odhiambo Omuya et al., 2021) developed a hybrid PCA-IG model to address the challenges of high-dimensional datasets. By minimizing redundant and irrelevant attributes, the PCA-IG model significantly improved classification accuracy and computational efficiency. This approach achieved superior performance in accuracy, precision, recall, and training time compared to conventional feature selection methods and classifiers such as Naïve Bayes. In addition to feature optimization, researchers have explored various data resampling techniques to mitigate class imbalance. Methods such as SMOTE (Synthetic Minority Over-sampling Technique), SMOTEENN, and SMOTETomek have been widely employed to enhance model sensitivity toward minority classes. (Husain et al., 2025) and (Alsharaiah et al., 2025) reported substantial improvements in minority class detection using these balancing strategies, while (Mohsin & Jony, 2024) integrated SMOTE-based balancing with classical classifiers, achieving accuracy gains exceeding 99% across multiple attack categories.

Furthermore, (Doménech et al., 2025) demonstrated the significant impact of preprocessing and balancing techniques, such as Random Undersampling and SMOTE, on the CIC-IoMT2024 dataset, improving accuracy by 26.35% and the F1-score by 29.40% over baseline models. However, they also observed that SMOTE occasionally reduced the detection rate for minority classes, such as "ARP Spoofing," due to the generation of ambiguous synthetic samples. Complementarily, (Torre et al., 2025) employed Federated Learning integrated with Differential Privacy and Homomorphic Encryption to ensure decentralized training with strong privacy protection, albeit at the expense of higher computational complexity. Overall, while the existing literature presents numerous innovations in IDS for IoMT, several limitations persist. Many models are designed for general IoT scenarios and do not account for the stringent latency and power requirements in healthcare settings. Additionally, model interpretability, essential for healthcare providers, is often overlooked. Few studies offer scalable solutions that combine lightweight processing with high detection accuracy and balanced class performance.

3. Proposed methodology

The methodology of this study follows a structured ML pipeline designed to construct and evaluate an effective IDS for IoMT networks. The entire workflow, from data preparation to performance evaluation, is illustrated in Figure 1 to validate their effectiveness and suitability for deployment in resource-constrained IoMT environments.

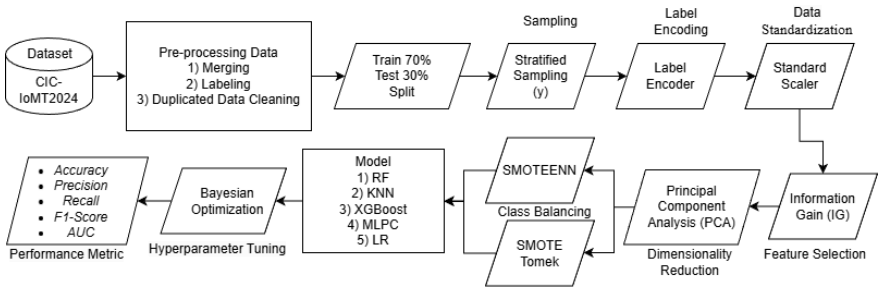

Figure 1. Proposed Method

3.1. Datasets Description

This research is grounded in the CIC-IoMT 2024 dataset (Dadkhah et al., 2024), provided by the Canadian Institute for Cybersecurity. The dataset encompasses diverse and representative network traffic patterns commonly observed in IoMT environments, including both normal operations and various forms of malicious activity such as DDoS, infiltration, and other significant cyber threats.

3.2. Data Preprocessing and Sampling

The initial stage of this study involved a systematic preprocessing pipeline applied to the CIC-IoMT2024 dataset to ensure high data quality and efficient computational handling. Initially, all raw data files were merged into a unified dataframe. A deduplication procedure was then performed to eliminate redundancy, resulting in the removal of 5,119 duplicate entries. This yielded a clean dataset consisting of 7,155,712 unique records, each described by 46 attributes. Each record in the dataset was categorized using the 'Label' feature as either 'Benign' or one of 18 distinct malicious attack types, as illustrated in Figure 2. To prepare the dataset for ML, all categorical variables, including the target 'Label', were converted into numerical form using Label Encoding.

The dataset was subsequently divided into training and testing subsets using a 70/30 split ratio, 5,008,998 samples for training (X_train, y_train) and 2,146,714 samples for testing (X_test, y_test). Given the large dataset size, training models directly on the full data would lead to substantial computational overhead. To address this, stratified random sampling was employed to maintain the proportional representation of all classes within both subsets, ensuring consistency with the original class distribution and minimizing sampling bias. A maximum threshold of 15,000 samples per class was established for downsampling the majority classes. Classes with counts exceeding this threshold (e.g., TCP_IP-DDoS-UDP, MQTT-DDoS-Connect_Flood) were randomly reduced to 15,000 samples, while minority classes with fewer instances (e.g., Recon-Ping_Sweep, Recon-VulScan) were retained entirely. This hybrid downsampling strategy effectively reduced the size of the training data while preserving the diversity across all 19 classes. The resulting sampled training dataset comprised 232,087 records, with class distribution summarized as follows:
Benign (15,000), TCP_IP-DoS-UDP (15,000), TCP_IP-DDoS-TCP (15,000), TCP_IP-DDoS-SYN (15,000), TCP_IP-DDoS-ICMP (15,000), MQTT-DoS-Publish_Flood (15,000), MQTT-DDoS-Connect_Flood (15,000), MQTT-DDoS-Publish_Flood (15,000), TCP_IP-DoS-ICMP (15,000), TCP_IP-DoS-SYN (15,000), Recon-Port_Scan (15,000), TCP_IP-DoS-TCP (15,000), TCP_IP-DDoS-UDP (15,000), Recon-OS_Scan (11,314), ARP_Spoofing (11,233), MQTT-DoS-Connect_Flood (8,941), MQTT-Malformed_Data (3,591), Recon-VulScan (1,490), and Recon-Ping_Sweep (518).

To enhance the performance and convergence of the learning algorithms, feature normalization was applied using StandardScaler, which transformed each feature to have a mean of zero ($\mu = 0$) and a standard deviation of one ($\sigma = 1$), as expressed in Equation (1).

$$z = (x - \mu) / \sigma \tag{1}$$

Standardization is a critical step to ensure that all features are on a comparable scale, preventing features with larger magnitudes from disproportionately influencing the model.

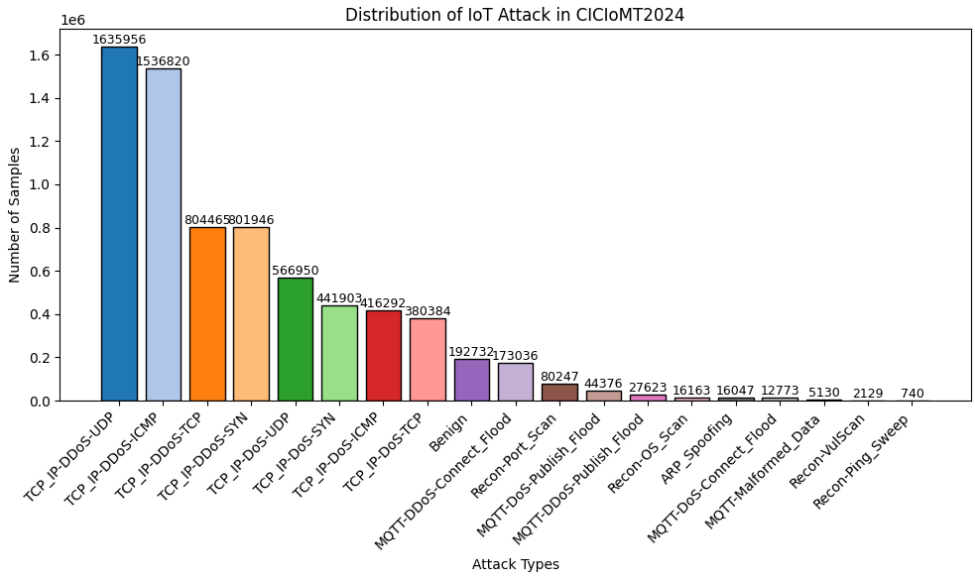

Figure 2. Distribution of cyberattacks

3.3. Feature Selection

IG is employed to select features with the highest predictive power by measuring their mutual information with the target class. It is calculated as the reduction in entropy after splitting the dataset based on an attribute. Formally, IG for an attribute A relative to a dataset S as shown in Equation (2):

$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{v \in \text{Values}(A)} \frac{|S_v|}{|S|} \text{Entropy}(S_v) \quad (2)$$

Where, $Entropy(S)$ is the entropy of the entire dataset, Values A denotes the set of possible values for attribute A, S_v is the subset of S where attribute A has value v, $\frac{|S_v|}{|S|}$ represents the proportion of the dataset falling into subset S_v. A higher IG value indicates a stronger dependency between the feature and the target variable, suggesting greater predictive relevance. The IG scores were computed for all 42 standardized features using the training dataset. Based on the distribution of IG values, an empirical threshold of IG > 0.001 was established to differentiate high-impact from low-impact features. Consequently, 40 features with IG scores above this threshold were retained for further model training and evaluation. The ranked list of features is presented in Table 1.

Rank	Feature Name	IG Value
1	IAT	2.644043
2	Tot sum	1.742246
3	Tot size	1.711671
4	Min	1.648812
5	Max	1.637822
6	AVG	1.609903
7	Magnitue	1.603614
8	Header_Length	1.537972
9	Srate	1.422746
10	Rate	1.422336
…	…	…
31	ARP	0.028318
32	IPv	0.027613
33	LLC	0.025614
34	DNS	0.009515
35	HTTP	0.007317
36	IRC	0.006482
37	cwr_flag_number	0.004747
38	Telnet	0.003675
39	SMTP	0.001834
40	ece_flag_number	0.001489
41	SSH	0.000744

Table 1. IG result

The remaining three features, namely Drate, DHCP, and IGMP, exhibited IG scores of 0.000000, indicating negligible contribution to class discrimination. Therefore, these features were excluded from further analysis to reduce noise and enhance model efficiency.

3.4. Dimensionality Reduction

To further refine the feature space and mitigate potential multicollinearity among the 42 features previously selected by Information Gain (IG), Principal Component Analysis (PCA) was applied. PCA is an unsupervised linear transformation technique that projects the data into a new subspace defined by orthogonal principal components (PCs) that capture the maximum variance within the original feature set. Considering the high-dimensional and multi-class characteristics of the CIC-IoMT2024 dataset, the number of principal components

was empirically determined to be 31. This configuration was selected to ensure that both dominant and subtle variance patterns were retained—critical for distinguishing among the 19 different traffic classes, including those with closely related characteristics. The transformation process is grounded in the covariance relationship between features, mathematically expressed in Equation (3):

$$cov_{x,y} = \frac{\sum(x_i - \underline{x})(y_i - \underline{y})}{N-1} \quad (3)$$

Where x_i and y_i are individual sample values, $\underline{x}$ and $\underline{y}$ are the mean values of variables x and y, and N denotes the total number of observations. The PCA transformation demonstrated that 31 components collectively explained 99.99% of the total variance, confirming that the reduced feature space preserved nearly all relevant information from the IG-selected features while significantly minimizing redundancy. This dimensionality reduction enhanced computational efficiency and reduced the potential risk of overfitting during model training. The transformed training and testing datasets exhibited shapes of (232,087, 31) and (2,146,714, 31), respectively, indicating successful feature compression while maintaining discriminative power. Table 2 summarizes the proportion of variance explained by each principal component. The first few components—PC1 (21.20%), PC2 (9.67%), PC3 (7.58%), PC4 (7.44%), and PC5 (6.57%)—collectively accounted for over 52% of the total variance, demonstrating that most of the variability in the data was captured by a relatively small number of components.

Principal Component (PC)	Individual Variance Explained (%)	Cumulative Variance Explained (%)
PC1	21.20	21.20
PC2	9.67	30.87
PC3	7.58	38.45
PC4	7.44	45.90
PC5	6.57	52.47
PC6	5.20	57.67
PC7	4.43	62.10
PC8	4.25	66.35
PC9	4.15	70.50
PC10	4.01	74.51
…	…	…
PC26	0.33	99.21
PC27	0.28	99.50
PC28	0.23	99.73
PC29	0.13	99.85
PC30	0.08	99.94
PC31	0.05	99.99

Table 2. PCA result

3.5. Class Balancing

A persistent challenge in most cybersecurity datasets, including the CIC-IoMT2024 dataset used in this study, is class imbalance. This issue arises when certain attack types or benign traffic dominate the dataset, causing machine learning (ML) models to become biased toward majority classes and perform poorly in detecting minority or rare attacks. To address this problem, two robust hybrid resampling techniques were applied to the PCA-transformed training data (comprising 31 components): SMOTEENN and SMOTETomek. Both methods combine synthetic over-sampling of minority classes with under-sampling or cleaning mechanisms to achieve a more balanced and representative dataset.

1. **SMOTEENN:** The Synthetic Minority Over-sampling Technique combined with Edited Nearest Neighbors (SMOTEENN) first generates synthetic samples for underrepresented attack classes to increase their presence in the dataset. Subsequently, the ENN cleaning process removes ambiguous or noisy samples from both majority and minority classes to refine the decision boundaries. Applying this approach produced a training dataset consisting of 227,025 samples across all 19 classes. The post-resampling distribution showed improved balance, with class counts ranging approximately between 5,692 and 14,909 instances per class.

2. **SMOTETomek:** The SMOTE combined with Tomek Links (SMOTETomek) method similarly begins by synthesizing minority samples using SMOTE and then identifies Tomek links—pairs of samples from different classes that are nearest neighbors. The majority-class instances involved in these links are removed to better delineate class boundaries. This method resulted in a slightly larger, more uniform dataset containing 271,764 samples, with class counts ranging approximately between 12,013 and 14,983 instances per class.

Both methods successfully mitigated the class imbalance observed in the original dataset (where class counts ranged from 518 to 15,000). While SMOTEENN introduced a moderately balanced structure emphasizing data cleanliness, SMOTETomek achieved a near-uniform class distribution and clearer decision margins. These balanced datasets were subsequently used to independently train and evaluate the machine learning models to assess their robustness and generalization across balanced conditions..

3.6. Model

Four ML models were chosen due to their established effectiveness in intrusion detection tasks:

1) Random Forest (RF) builds a large collection of decision trees and then aggregates their individual predictions. This aggregation process boosts the model's overall accuracy and reduces the risk of overfitting to the training data (Al-Abadi et al. (2023)).
2) K-Nearest Neighbors (KNN) is a non-parametric algorithm used for classification by labeling a data point according to the most common class among its k nearest neighbors within the feature space (Sun & Chen (2021)).
3) XGBoost (XGB) is an advanced gradient boosting algorithm. It develops a predictive model by creating a series of decision trees one after another, with each new tree aiming to rectify the inaccuracies of the preceding models in the sequence (Salehpour et al. (2024)).
4) Logistic Regression (LR) is a fundamental statistical algorithm used for binary classification tasks. Despite its name, it is a classification model that calculates the probability of a specific outcome. It employs the logistic (sigmoid) function to transform the output of a linear equation into a probability score between 0 and 1, making it highly interpretable and computationally efficient (Chalichalamala et al. (2023)).
5) Multi-Layer Perceptron Classifier (MLPC) is a supervised artificial neural network model consisting of interconnected layers of nodes: input, hidden, and output. It applies nonlinear activation functions to capture complex patterns in data and employs backpropagation for weight optimization, enabling robust classification performance across diverse domains, particularly in high-dimensional and nonlinear problem spaces (Zhao et al. (2025)).

3.7. Hyperparameter Optimization

To fine-tune each model and maximize its predictive performance, BO was employed. BO efficiently explores the hyperparameter search space by building a probabilistic surrogate model that estimates the objective function based on prior evaluations. In this study, the objective function $f(x)$ represents the mean 5-fold cross-validation accuracy for a given set of hyperparameters x. The optimization iteratively updates its belief about $f(x)$ using a Gaussian Process (GP) and selects the next candidate point x^* by maximizing an acquisition function, typically the Expected Improvement (EI). Mathematically, the BO process can be expressed as (4).

$$x^* = \arg\max_{x \in X} \alpha(x \mid D_t), \quad \text{where} \quad \alpha(x \mid D_t) = \mathbb{E}[max(f(x) - f(x^+), 0)] \tag{4}$$

Here, $D_t = \{(x_i, f(x_i))\}_{i=1}^{t}$ denotes the set of observed evaluations, $f(x^+)$ is the best performance obtained so far, and $\alpha(x \mid D_t)$ represents the acquisition function guiding the exploration–exploitation trade-off. This automated and probabilistically guided tuning process ensures that each classifier—trained on both SMOTEENN and SMOTETomek balanced datasets—is optimized toward its most effective configuration before final evaluation, thereby enhancing the robustness, generalizability, and reliability of our findings.

3.8. Evaluation

To validate the dependability and stability of the proposed models, we evaluated their performance on intrusion detection within the IoMT environment using a diverse range of evaluation metrics. These metrics include:

1) $Accuracy = \frac{TN+TP}{TN+FP+TP+FN}$, measuring overall correctness
2) $Precision = \frac{TP}{TP+FP}$, measures the fraction of true positive cases among all instances classified as positive, indicating how accurately the model identifies positive outcomes.
3) $Recall = \frac{TP}{TP+FN}$, reflecting the model's sensitivity to actual positives.
4) $F1\,Score = 2 * \frac{Precision * Recall}{Precision + Recall}$, balancing precision and recall.
5) $AUC = TPR * \frac{TP}{TP + FN}$, evaluates a model's capability to differentiate between classes in classification tasks, providing insight into its overall discriminative power.

This evaluation framework facilitates the effective development of a smart IDS tailored for the intricate IoMT landscape.

4. Result and Analysis

This research presents a thorough evaluation of the ML models developed for intrusion detection within IoMT environments. A systematic performance analysis is conducted for each model using datasets balanced through SMOTEENN and SMOTETomek techniques. The results are critically interpreted and positioned within the context of existing scholarly works to underscore the novel contributions of this study. The implementation was carried out on Google Colab Pro, leveraging high-performance CPUs and substantial RAM resources to ensure efficient execution. This computationally robust configuration facilitates the handling of memory-intensive operations and accelerates processing, thereby enhancing the overall experimental workflow.

4.1. Experimental Design

The discussion section interprets these results, compares them with existing literature, and highlights the contributions of this research. The evaluation was conducted using five state-of-the-art ML classifiers: RF, KNN, XGB, MLPC and LR. Each model was trained and tested on two distinct datasets that were processed using the IG-PCA feature engineering pipeline and balanced with either SMOTEENN or SMOTETomek. Model performance was assessed using standard evaluation metrics: Accuracy, Precision, Recall, F1-Score and AUC.

To ensure each model achieved its maximum potential, all classifiers were fine-tuned using the BO process. A predefined search space, outlining the range of possible values for each hyperparameter, was established to guide the optimization. The specific search spaces for each classifier are detailed in Tables 3 through 7.

Hyperparameter	Type	Search Space
'n_neighbors'	Integer	[100, 1000]
'max_depth'	Integer	[5, 100]
'min_samples_split'	Integer	[2, 20]
'min_samples_leaf'	Integer	[1, 10]
'max_features'	Categorical	[sqrt, log2, none]
'bootstrap'	Categorical	[True, False]
'criterion'	Categorical	['gini', 'entropy']

Table 3. RF Search Space Perimeter

Hyperparameter	Type	Search Space
'n_neighbors'	Integer	[1, 90]
'weights'	Categorical	['uniform', 'distance']
'metric'	Categorical	['euclidean', 'manhattan', 'minkowski']
'algorithm'	Categorical	['auto', 'ball_tree', 'kd_tree', 'brute']
'leaf_size'	Integer	[10, 900]
'p'	Integer	[1, 2]

Table 4. KNN Search Space Perimeter

Perimeter name	Type	Search Space
'n_estimators'	Integer	[50, 500]
'max_depth'	Integer	[3, 90]
'learning_rate'	Float	[0.01, 0.3]
'subsample'	Float	[0.5, 1.0]
'colsample_bytree'	Float	[0.5, 1.0]
'gamma'	Float	[0, 5]
'min_child_weight'	Integer	[1, 10]
'reg_alpha'	Float	[0, 5]
'reg_lambda'	Float	[0, 5]
'scale_pos_weight'	Float	[0.5, 5]

Table 5. XGB Search Space Perimeter

Hyperparameter	Type	Search Space
'Penalty'	Categorical	['l1', 'l2', 'elasticnet', 'none']
'C'	Real	[1e-3, 1e3, prior = 'log-uniform']
'solver'	Categorical	['lbfgs', 'liblinear', 'saga', 'newton-cg']
'max_iter'	Integer	[100, 2000]
'l1_ratio'	Real	[0, 1, prior = 'uniform']

Table 6. LR Search Space Perimeter

Hyperparameter	Type	Search Space
'n_layers'	Categorical	[1, 2, 3]
'hl1'	Integer	32 – 512
'hl2'	Integer	16 – 512
'hl3'	Integer	8 – 512
'activation'	Categorical	['relu', 'tanh', 'logistic']
'solver'	Categorical	['adam', 'sgd', 'lbfgs']
'alpha'	Real	1e-5 – 1e-1 (log-uniform)
'learning_rate'	Categorical	['constant', 'invscaling', 'adaptive']
'learning_rate_init'	Real	1e-4 – 1e-1 (log-uniform)
'max_iter'	Integer	200 – 2000
'batch_size'	Integer	32 – 512
'early_stopping'	Categorical	[True, False]

Table 7. MLPC Search Space Perimeter

4.2. Performance on the SMOTEENN-Balanced Dataset

Models trained on the dataset balanced with the SMOTEENN technique demonstrated exceptional classification performance. Notably, ensemble-based methods yielded the most competitive results. Specifically, RF and XGB achieved near-perfect scores across all evaluation metrics. This highlights their robustness in mitigating class imbalance and their capacity for modeling complex decision boundaries.

The detailed performance metrics following BO are presented in Table 8, with a corresponding graphical representation in Figure 3. These results underscore the superior performance of the ensemble models. XGB attained the highest scores across all metrics, with an accuracy, precision, recall, and F1-score of 99.811%. RF followed closely with highly comparable performance metrics. Both models achieved a near-perfect Area Under the Curve (AUC) of 99.999%, indicating exceptional discriminative capability.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
RF	99.795	99.795	99.795	99.795	99.999
XGB	99.811	99.811	99.811	99.811	99.999
KNN	93.923	93.943	93.923	93.901	99.773
LR	78.047	83.261	78.047	76.071	99.117
MLPC	95.815	95.961	95.815	95.808	99.817

Table 8. The BO result for SMOTEENN

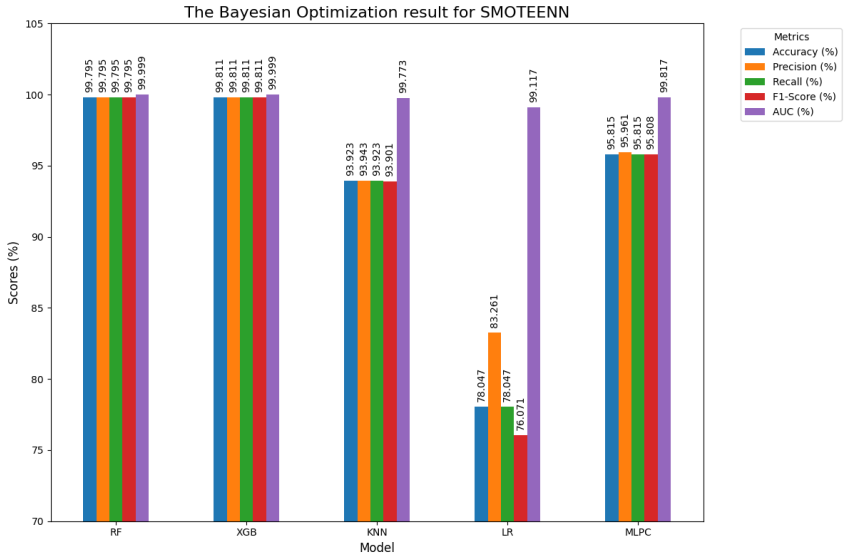

Figure 3. The BO result for SMOTEENN

BO identified the optimal configuration for RF as comprising 224 estimators with a maximum depth of 18. The use of the entropy criterion and max_features = 'log2' facilitated the development of diverse and minimally correlated trees, thereby enhancing the ensemble's robustness. Furthermore, setting min_samples_leaf = 3 and min_samples_split = 10 effectively mitigated overfitting by ensuring adequate sample representation in each terminal node and controlling the tree splitting process. This configuration enabled the RF model to achieve a weighted F1-score of 0.9977, reflecting strong predictive performance across all classes.

For XGB, the optimal setup consisted of 224 estimators, a maximum depth of 25, and a learning rate of 0.1357, achieving a balance between model complexity and convergence speed. Parameters such as colsample_bytree = 0.7224 and subsample = 0.8517 promoted ensemble diversity by limiting feature and sample usage per tree. The regularization terms reg_alpha = 0.7516 and reg_lambda = 3.2294 were instrumental in preventing overfitting by constraining model complexity. Additionally, gamma = 0.0524 and min_child_weight = 4.5367 regulated partitions to avoid excessive granularity. The scale_pos_weight = 1.4732 parameter addressed any residual class imbalance. This comprehensive tuning resulted in a weighted F1-score of 0.9981, marginally surpassing RF and demonstrating its outstanding classification power.

In contrast, the non-ensemble models demonstrated lower, albeit still effective, performance. The KNN classifier achieved a respectable F1-score of 93.901% and an AUC of 99.773%. LR yielded more moderate results, with an F1-score of 76.071% and an AUC of 99.117%, indicating reasonable but comparatively lower effectiveness. In summary, the findings confirm that for IoMT intrusion detection, ensemble methods, particularly XGB and RF, when coupled with the SMOTEENN balancing technique, deliver the most accurate, robust, and generalizable classification performance.

4.3. Performance on the SMOTETomek-Balanced Dataset

Models trained on the dataset balanced using the SMOTETomek algorithm demonstrated consistently high classification performance. The results affirmed the superiority of ensemble-based approaches, which delivered the most competitive outcomes. Both RF and XGB achieved near-perfect scores across all evaluation metrics, validating their robustness in handling class imbalance and their capability to model intricate decision boundaries.

The detailed outcomes of the BO for hyperparameter tuning are cataloged in Table 9 and visualized in Figure 4. Consistent with the findings from the SMOTEENN dataset, ensemble methods emerged as the top performers. XGB was the leading model, achieving the highest weighted F1-score of 98.892%, marginally surpassing RF, which secured an F1-score of 98.752%. Both models also attained exceptional AUC values, indicating robust discriminative power.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC (%)
RF	98.754	98.784	98.754	98.752	99.975
XGB	98.897	98.899	98.897	98.892	99.983
KNN	91.959	91.943	91.959	91.901	99.647
LR	76.719	81.287	76.719	74.773	98.861
MLPC	94.212	94.985	94.212	94.057	99.694

Table 9. The BO result for SMOTETomek

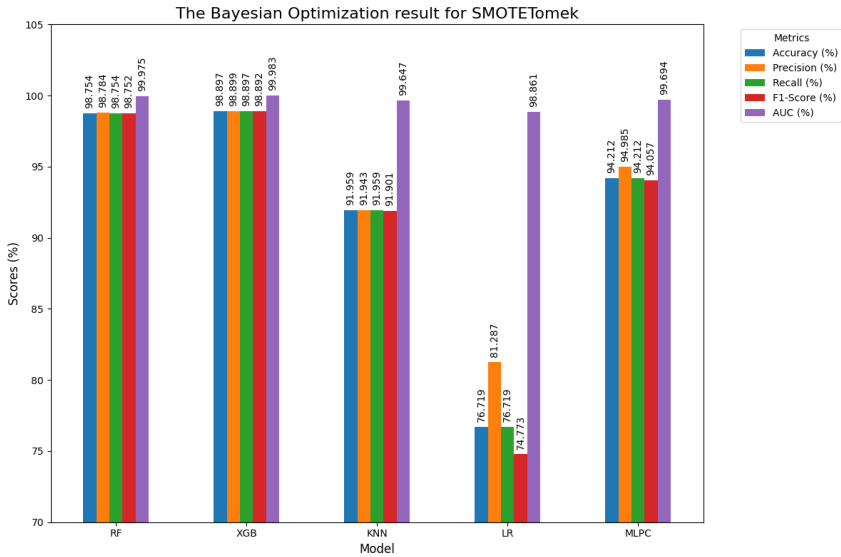

Figure 4. The BO result for SMOTETomek

BO identified the optimal configuration for RF vs 224 estimators with a maximum depth of 18. The selection of the entropy criterion alongside max_features = 'log2' was instrumental in fostering the growth of diverse and minimally correlated trees, which enhanced the overall robustness of the ensemble. To prevent overfitting, min_samples_leaf = 3 and min_samples_split = 10 were implemented, ensuring that each terminal node was supported by a sufficient number of samples while controlling the frequency of splits. This carefully tuned configuration enabled RF to achieve strong predictive accuracy and balanced performance.

For XGB, the optimal hyperparameter set included 224 estimators, a maximum depth of 25, and a learning rate of 0.1357, a combination designed to strike a balance between the model's predictive power and its convergence stability. Overfitting was further mitigated by setting colsample_bytree = 0.7224 and

subsample = 0.8517, which promoted ensemble diversity by restricting the proportion of features and samples used for each tree. The regularization terms (reg_alpha = 0.7516, reg_lambda = 3.2294) and split-control parameters (gamma = 0.0524, min_child_weight = 4.5367) effectively constrained model complexity. Finally, scale_pos_weight = 1.4732 was applied to address any residual class imbalance. This optimized setup confirmed XGB's potent and well-balanced classification capability.

The other models also yielded noteworthy results. KNN and the MLPC delivered solid performance, with F1-scores of 91.901% and 94.057%, respectively. In contrast, LR showed comparatively lower predictive accuracy, suggesting a greater sensitivity to the data structure produced by the SMOTETomek process. In conclusion, these findings reinforce that for IoMT intrusion detection, ensemble methods—particularly XGB and RF —when paired with the SMOTETomek balancing technique, provide the most accurate, robust, and generalizable classification performance.

4.4. Comparative Analysis of proposed techniques

The integration of IG for feature selection and PCA for dimensionality reduction proved highly effective in refining the dataset. By prioritizing the most informative features and eliminating redundant or irrelevant attributes, this preprocessing strategy substantially reduced data complexity while preserving critical discriminative information. This step not only enhanced computational efficiency but also contributed to improving model generalization capability. In addition, the application of the SMOTEENN and SMOTETomek balancing methods successfully mitigated the issue of class distribution imbalance. These techniques ensured that the models remained unbiased toward majority classes, thereby improving their ability to accurately detect minority-class attack instances—a crucial requirement for intrusion detection in IoMT environments. As shown in Table 10, our proposed XGB model achieved an accuracy of 99.811% on the CIC-IoMT 2024 dataset. This performance surpasses several benchmarks, including the RF model from Dadkhah et al. (73.3%) and the XGB model from Mohsin & Jony (99.20%). Notably, our model's accuracy is highly comparable to other state-of-the-art methods, such as those by Lucia Hernandez-Jaimes et al. (99.83%) and Ghourabi & Alkhalil (99.85%), validating the effectiveness of the optimized XGB pipeline proposed in this study.

Model	Accuracy
RF with 19 Class (Dadkhah et al., 2024)	73.3%
RF with 2 Class + RandomOverSampler (Abdiwi, 2024)	99.7%
XGB with 19 Class (Mohsin & Jony, 2024)	99.20%
XGB with 6 Class + SMOTE (Doménech et al., 2025)	99.83%
XGB with 2 Class + Federated Learning + BO (Ghourabi & Alkhalil, 2025)	99.85%
Proposed XGB with 19 Class + IG-PCA + SMOTEENN + BO (this paper)	99.811%

Table 10. Comparison best accuracy result for CIC-IoMT 2024 dataset

4.5. Organizational Implications and Practical Adoption

Beyond its technical accuracy, the practical viability of the proposed IDS is determined by its integration within a healthcare organization's existing operational, governance, and strategic frameworks. This section discusses the key organizational implications for its successful adoption.

For adoption and implementation, the proposed framework is designed as a software-based solution that can be deployed on a dedicated server or virtual machine within the hospital's data center. Integration into the existing infrastructure would involve configuring the system to monitor network traffic from core switches or gateways, a task manageable by the organization's network and cybersecurity teams. While a data scientist may be beneficial for initial model fine-tuning, the lightweight nature of the IG-PCA pipeline minimizes the need for specialized, high-cost hardware, making it accessible for institutions with limited resources. The primary requirement is a moderately provisioned server and collaboration between existing IT personnel for setup and maintenance, ensuring a low barrier to entry.

Effective governance is crucial for translating the system's alerts into actionable responses. Responsibility for the IDS would typically reside with the Information Security team, under the direct oversight of the Chief Information Security Officer (CISO). A clear incident response protocol must be established: upon detecting a potential threat, the IDS would generate an alert, ideally feeding into a central Security Information and Event

Management (SIEM) system (González-Granadillo et al., 2021). Tier-1 security analysts would then be responsible for initial validation. Confirmed critical threats would be escalated to a dedicated incident response team, with the CISO holding the authority to approve decisive actions, such as isolating compromised medical devices to contain the threat and protect the wider network.

The strategic benefits for the healthcare organization extend far beyond technical threat detection. By providing robust security, the IDS directly enhances patient trust in the hospital's digital services, which is critical for the adoption of telehealth and online patient portals. Furthermore, implementing such an advanced security measure helps the organization demonstrate due diligence and achieve compliance with stringent data protection regulations and cybersecurity accreditation standards. Proactively identifying threats reduces the risk of operational downtime in critical clinical systems, preventing significant financial losses and ensuring continuity of patient care. Ultimately, the system serves as a strategic asset, providing the CISO with valuable threat intelligence to inform risk management policies and justify security investments (Ramezan, 2025).

5. Conclusion

This study proposed and validated a comprehensive framework for intrusion detection tailored to the unique challenges of the IoMT, including high-dimensional feature spaces and pronounced class imbalance. The methodology integrated a multi-stage data preprocessing pipeline consisting of IG for feature selection, PCA for dimensionality reduction, and hybrid resampling through SMOTEENN and SMOTETomek to achieve data balance. This approach effectively enhanced data quality and improved overall model robustness. Furthermore, BO was systematically employed to fine-tune the hyperparameters of five classifiers—RF, KNN, XGB, LR, and MLPC—resulting in consistent performance gains across all models. Notably, the XGB classifier combined with the SMOTEENN technique achieved a peak accuracy of 99.811% on the CIC-IoMT2024 dataset. This performance is highly competitive with state-of-the-art approaches and surpasses several existing benchmarks. Although the proposed XGB model did not achieve the absolute highest accuracy, the difference (0.039%) is marginal and statistically insignificant given the 19-class complexity of the dataset. Importantly, the proposed framework achieved this result with lower computational overhead and improved class balance, underscoring its scalability and deployment potential in real-world IoMT intrusion detection systems.

Overall, the findings demonstrate the efficacy and practical viability of the proposed framework in developing accurate, efficient, and scalable intrusion detection systems suitable for resource-constrained IoMT environments. This research contributes a practical tool that enhances the cybersecurity posture of healthcare institutions. By enabling reliable and efficient threat detection, the framework strengthens operational resilience, protects critical medical data, and supports governance structures essential for secure digital transformation in the healthcare sector. Future work will focus on two key directions. First, the scalability and robustness of the framework will be validated using larger and more heterogeneous IoMT datasets. Second, lightweight deep learning architectures will be explored to further reduce detection latency while adhering to the strict energy efficiency and performance constraints inherent in IoMT devices.

Acknowledgement

This research is a part of Fundamental Research funded under contract number 012/MOU/LPPM/UNAMA/1V/2025, and supported by Yayasan Dinamika Bangsa.

References

Abdiwi, F. G. (2024). Hybrid Machine Learning and Blockchain Technology for Early Detection of Cyberattacks in Healthcare Systems. International Journal of Safety and Security Engineering, 14(6), 1883–1893. https://doi.org/10.18280/ijsse.140622

Ahmed, S. F., Alam, M. S. Bin, Afrin, S., Rafa, S. J., Rafa, N., & Gandomi, A. H. (2024). Insights into Internet of Medical Things (IoMT): Data fusion, security issues and potential solutions. Information Fusion, 102. https://doi.org/10.1016/j.inffus.2023.102060

Al-Abadi, A. A. J., Mohamed, M. B., & Fakhfakh, A. (2023). Enhanced Random Forest Classifier with K-Means Clustering (ERF-KMC) for Detecting and Preventing Distributed-Denial-of-Service and Man-in-the-Middle Attacks in Internet-of-Medical-Things Networks. Computers, 12(12), 262. https://doi.org/10.3390/computers12120262

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). https://doi.org/10.3390/s24113519

Alsharaiah, M. A., Almaiah, M. A., Shehab, R., Obeidat, M., El-Qirem, F. A., & Aldhyani, T. (2025). An explainable AI-driven transformer model for spoofing attack detection in Internet of Medical Things (IoMT) networks. Discover Applied Sciences, 7(5), 488. https://doi.org/10.1007/s42452-025-07071-5

Alturki, B., Abu Al-Haija, Q., Alsemmeari, R. A., Alsulami, A. A., Alqahtani, A., Alghamdi, B. M., Bakhsh, S. T., & Shaikh, R. A. (2025). IoMT landscape: navigating current challenges and pioneering future research trends. In Discover Applied Sciences (Vol. 7, Issue 1). Springer Nature. https://doi.org/10.1007/s42452-024-06351-w

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, 13, 32863–32882. https://doi.org/10.1109/ACCESS.2025.3543127

Binbusayyis, A., Alaskar, H., Vaiyapuri, T., & Dinesh, M. (2022). An investigation and comparison of machine learning approaches for intrusion detection in IoMT network. The Journal of Supercomputing, 78(15), 17403–17422. https://doi.org/10.1007/s11227-022-04568-3

Bouke, M. A., El Atigh, H., & Abdullah, A. (2024). Towards robust and efficient intrusion detection in IoMT: a deep learning approach addressing data leakage and enhancing model generalizability. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-024-19916-z

Chaganti, R., Mourade, A., Ravi, V., Vemprala, N., Dua, A., & Bhushan, B. (2022). A Particle Swarm Optimization and Deep Learning Approach for Intrusion Detection System in Internet of Medical Things. Sustainability, 14(19), 12828. https://doi.org/10.3390/su141912828

Chalichalamala, S., Govindan, N., & Kasarapu, R. (2023). Logistic Regression Ensemble Classifier for Intrusion Detection System in Internet of Things. Sensors, 23(23), 9583. https://doi.org/10.3390/s23239583

Chuma, K. G., & Ngoepe, M. (2022). Security of electronic personal health information in a public hospital in South Africa. Information Security Journal: A Global Perspective, 31(2), 179–195. https://doi.org/10.1080/19393555.2021.1893410

Dadkhah, S., Neto, E. C. P., Ferreira, R., Molokwu, R. C., Sadeghi, S., & Ghorbani, A. A. (2024). CICIoMT2024: A benchmark dataset for multi-protocol security assessment in IoMT. Internet of Things, 28, 101351. https://doi.org/10.1016/j.iot.2024.101351

Doménech, J., León, O., Siddiqui, M. S., & Pegueroles, J. (2025). Evaluating and enhancing intrusion detection systems in IoMT: The importance of domain-specific datasets. Internet of Things, 32, 101631. https://doi.org/10.1016/j.iot.2025.101631

Ghourabi, A., & Alkhalil, A. (2025). A Federated Learning Model for Detecting Cyberattacks in Internet of Medical Things Networks. IEEE Access, 13, 123018–123030. https://doi.org/10.1109/ACCESS.2025.3588808

González-Granadillo, G., González-Zarzosa, S., & Diaz, R. (2021). Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. Sensors, 21(14), 4759. https://doi.org/10.3390/s21144759

Husain, G., Nasef, D., Jose, R., Mayer, J., Bekbolatova, M., Devine, T., & Toma, M. (2025). SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models. Algorithms, 18(1), 37. https://doi.org/10.3390/a18010037

Ibrahim, M., & Al-Wadi, A. (2024). Enhancing IoMT network security using ensemble learning-based intrusion detection systems. Journal of Engineering Research (Kuwait). https://doi.org/10.1016/j.jer.2024.12.003

Kumar, A., Sangwan, S. R., Arora, A., & Menon, V. G. (2022). Depress-DCNF: A deep convolutional neuro-fuzzy model for detection of depression episodes using IoMT. Applied Soft Computing, 122, 108863. https://doi.org/10.1016/j.asoc.2022.108863

Kumari, P., & Jain, A. K. (2024). Timely detection of DDoS attacks in IoT with dimensionality reduction. Cluster Computing, 27(6), 7869–7887. https://doi.org/10.1007/s10586-024-04392-9

Mathkor, D. M., Mathkor, N., Bassfar, Z., Bantun, F., Slama, P., Ahmad, F., & Haque, S. (2024). Multirole of the internet of medical things (IoMT) in biomedical systems for managing smart healthcare systems: An

overview of current and future innovative trends. Journal of Infection and Public Health, 17(4), 559–572. https://doi.org/10.1016/j.jiph.2024.01.013

Mohsin, M., & Jony, A. I. (2024). A Comparative Analysis of Medical IoT Device Attacks Using Machine Learning Models. Malaysian Journal of Science and Advanced Technology, 429–439. https://doi.org/10.56532/mjsat.v4i4.318

Nasir, M., Javed, A. R., Tariq, M. A., Asim, M., & Baker, T. (2022). Feature engineering and deep learning-based intrusion detection framework for securing edge IoT. The Journal of Supercomputing, 78(6), 8852–8866. https://doi.org/10.1007/s11227-021-04250-0

Odhiambo Omuya, E., Onyango Okeyo, G., & Waema Kimwele, M. (2021). Feature Selection for Classification using Principal Component Analysis and Information Gain. Expert Systems with Applications, 174, 114765. https://doi.org/10.1016/j.eswa.2021.114765

Papaioannou, M., Karageorgou, M., Mantas, G., Sucasas, V., Essop, I., Rodriguez, J., & Lymberopoulos, D. (2022). A Survey on Security Threats and Countermeasures in Internet of Medical Things (IoMT). Transactions on Emerging Telecommunications Technologies, 33(6). https://doi.org/10.1002/ett.4049

Ramezan, C. A. (2025). Understanding the chief information security officer: Qualifications and responsibilities for cybersecurity leadership. Computers & Security, 152, 104363. https://doi.org/10.1016/j.cose.2025.104363

Razdan, S., & Sharma, S. (2022). Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies. In IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India) (Vol. 39, Issue 4, pp. 775–788). Taylor and Francis Ltd. https://doi.org/10.1080/02564602.2021.1927863

R.M., S. P., Maddikunta, P. K. R., M., P., Koppu, S., Gadekallu, T. R., Chowdhary, C. L., & Alazab, M. (2020). An effective feature engineering for DNN using hybrid PCA-GWO for intrusion detection in IoMT architecture. Computer Communications, 160, 139–149. https://doi.org/10.1016/j.comcom.2020.05.048

Salehpour, A., Norouzi, M., Balafar, M. A., & SamadZamini, K. (2024). A cloud-based hybrid intrusion detection framework using XGBoost and ADASYN-Augmented random forest for IoMT. IET Communications, 18(19), 1371–1390. https://doi.org/10.1049/cmu2.12833

Sarkar, M., Lee, T.-H., & Sahoo, P. K. (2024). Smart Healthcare: Exploring the Internet of Medical Things with Ambient Intelligence. Electronics, 13(12), 2309. https://doi.org/10.3390/electronics13122309

Sayed, N., Shoaib, M., Ahmed, W., Noman Qasem, S., M. Albarrak, A., & Saeed, F. (2023). Augmenting IoT Intrusion Detection System Performance Using Deep Neural Network. Computers, Materials & Continua, 74(1), 1351–1374. https://doi.org/10.32604/cmc.2023.030831

Sun, B., & Chen, H. (2021). A Survey of k Nearest Neighbor Algorithms for Solving the Class Imbalanced Problem. Wireless Communications and Mobile Computing, 2021(1). https://doi.org/10.1155/2021/5520990

Talukder, Md. A., Sharmin, S., Uddin, M. A., Islam, M. M., & Aryal, S. (2024). MLSTL-WSN: machine learning-based intrusion detection using SMOTETomek in WSNs. International Journal of Information Security, 23(3), 2139–2158. https://doi.org/10.1007/s10207-024-00833-z

Torre, D., Chennamaneni, A., Jo, J., Vyas, G., & Sabrsula, B. (2025). Toward Enhancing Privacy Preservation of a Federated Learning CNN Intrusion Detection System in IoT: Method and Empirical Study. ACM Transactions on Software Engineering and Methodology, 34(2), 1–48. https://doi.org/10.1145/3695998

Wang, J., Lim, M. K., Wang, C., & Tseng, M.-L. (2021). The evolution of the Internet of Things (IoT) over the past 20 years. Computers & Industrial Engineering, 155, 107174. https://doi.org/10.1016/j.cie.2021.107174

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 (Switzerland), 10(21). https://doi.org/10.3390/electronics10212562

Zhao, Q., Wang, F., Wang, W., Zhang, T., Wu, H., & Ning, W. (2025). Research on intrusion detection model based on improved MLP algorithm. Scientific Reports, 15(1), 5159. https://doi.org/10.1038/s41598-025-89798-0