IEEE International Conference on Internet of Things and Intelligence System (IoTaIS), Bandung, Indonesia, 2024

Efficient Machine Learning Frameworks for Strengthening Cybersecurity in Internet of Medical Things (IoMT) Ecosystems

Keshav Ramesh
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
b00088595@aus.edu

Nikita Christ Miller
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
g00089341@aus.edu

Aariz Faridi
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
b00086502@aus.edu

Dr. Fadi Aloul
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
faloul@aus.edu

Dr. Imran Zualkernan
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
izualkernan@aus.edu

Ali Reza Sajun
Computer Science and Engineering
American University of Sharjah
Sharjah, United Arab Emirates
b00068908@aus.edu

***Abstract*—As the Internet of Medical Things (IoMT) continues to transform healthcare, it also introduces new vulnerabilities to sophisticated cyberattacks that outpace conventional defenses. In response, we present a tailored Intrusion Detection System (IDS) optimized for IoMT environments, designed to operate within the constraints of resource-limited devices while addressing complex, real-world attack vectors. Leveraging the CICIoMT2024 dataset and advanced machine learning models like Random Forest and XGBoost, our approach overcomes severe class imbalance and high dimensionality. Using Recursive Feature Elimination with Cross-Validation (RFECV), we reduced the feature set by 44.45%, achieving a state-of-the-art weighted F1-score of 99.48%. Despite the superior performance of the Random Forest model, its large memory footprint poses challenges for deployment on IoMT devices with limited resources. In contrast, the XGBoost model offers a better balance between high detection accuracy and resource consumption, making it more suitable for real-world applications. Our solution offers a scalable, efficient, and deployable IDS that brings a new level of adaptability and precision to IoMT cybersecurity, ready to defend against today's threats while evolving to meet tomorrow's challenges.**

***Index Terms*—Internet of Medical Things (IoMT), Intrusion Detection Systems (IDS), Machine Learning, Cybersecurity, Dimensionality Reduction, Model Optimization**

I. Introduction

The Internet of Medical Things (IoMT) has emerged as a pivotal subset of the expanding Internet of Things (IoT) ecosystem, fundamentally reshaping healthcare services. From personal health monitors to advanced medical devices, IoMT solutions offer unprecedented opportunities to enhance patient care. However, with this rapid expansion comes heightened vulnerability to cyberattacks, posing significant threats to both patient safety and privacy. The sheer number of these devices, coupled with the increasingly sophisticated nature of attacks, increases their vulnerability. From Distributed Denial-of-Service (DDoS) and Denial-of-Service (DoS) attacks to more advanced spoofing and reconnaissance efforts, the complexity of these threats often surpasses the capacity of conventional cybersecurity measures. Consequently, intrusion detection and prevention systems must evolve to address this multifaceted threat landscape.

Intrusion detection systems (IDSs) tailored for IoMT environments are critical, but they often struggle with the resource limitations of small IoMT devices, such as constrained energy and processing capabilities. In this context, machine learning (ML) and deep learning (DL) techniques are emerging as the most effective approaches for threat detection and mitigation. Our research leverages the CICIoMT2024 dataset—a comprehensive collection of network traffic data from various IoMT devices exposed to a range of cyberattacks. This dataset, with its classification of both malicious and benign activity, provides a valuable resource for the development and evaluation of sophisticated IDS solutions customized for IoMT environments.

This paper outlines our methodology for designing an IDS using the CICIoMT2024 dataset, aiming to provide a detailed analysis of IoMT network traffic under diverse cyberattack scenarios, thereby advancing the security of healthcare IoMT systems. In an environment of continuously evolving threats, our goal is to lay the groundwork for more adaptive and robust cybersecurity solutions, enabling the development of intelligent, resilient defenses against an increasingly complex array of cyberthreats.

Our research offers two key contributions to the field of IoMT cybersecurity:

1) **Model size optimization:** We address the challenge of large model sizes associated with extensive datasets like CICIoMT2024. Recognizing the practical limitations of IoMT devices—particularly in terms of storage and processing capacity—our work emphasizes the need for model simplification. This not only facilitates efficient deployment but also fills a critical gap in the literature concerning model optimization for real-world IoMT applications.
2) **Comprehensive evaluation methodology:** Our work

further distinguishes itself by employing a more rigorous evaluation approach compared to baseline studies. We implement K-fold cross-validation (CV), which provides a more robust and accurate performance measure, particularly in scenarios with limited or imbalanced data. Additionally, we utilize F1 scores as a benchmark metric—a crucial, yet often overlooked, measure in evaluating models for imbalanced datasets. This comprehensive evaluation ensures that our IDS performs reliably across various attack scenarios, offering a more accurate assessment of real-world performance.

II. Literature Review

Remote medicine and alert systems leverage IoMT to enhance patient care in remote settings while ensuring continuity of services. This domain encompasses not only teleconsultations but also remote diagnosis and therapy, providing a comprehensive approach to healthcare delivery in challenging environments [1], [2], [3].

The integration of Intrusion Detection Systems (IDS) within the Internet of Things (IoT) and IoMT has been significantly advanced through sophisticated machine learning (ML) and deep learning (DL) models. However, the availability of datasets specifically focused on IoMT attacks remains limited. Due to this scarcity, we also review IoT-focused studies, as both IoT and IoMT devices are vulnerable to similar attack vectors, such as Distributed Denial-of-Service (DDoS) and spoofing. For instance, research utilizing the CIC IoT 2022 dataset demonstrated that various ML models achieved high accuracies, such as Decision Tree at 98.5%, AdaBoost at 98.7%, XGBoost at 98.6%, and K-Nearest Neighbors (KNN) at 95.6% [4]. These results underscore the relevance of IoT-based research in evaluating IDS technologies, as insights from IoT attacks can be effectively applied to securing IoMT environments.

Another study employed the TON-IoT telemetry dataset, using a hybrid CNN-LSTM model, which achieved an accuracy of 94%, demonstrating the effectiveness of hybrid deep learning architectures in complex IoMT scenarios [5]. Additionally, deep learning methods have been shown to be highly effective in detecting subtle intrusion activities. For example, one approach utilizing CNN and LSTM networks with the WUSTL EHMS 2020 dataset reported high accuracy and precision of 99% [6]. Whereas when another paper applied Logistic Regression and Random Forest models on this dataset, the Decision Tree achieved an accuracy of 96.56% [7].

The integration of IDS in the context of IoMT has been further explored in [8], focusing on secure Bluetooth communication in healthcare environments. The authors introduce a novel dataset tailored to IoMT topologies, with a specific emphasis on Bluetooth vulnerabilities. Their evaluation included multiple machine learning (ML) models, such as Support Vector Machines (SVM), K-Means, and Deep Neural Networks (DNN). The results demonstrated the effectiveness of these models, with the SVM achieving an accuracy of 96.3%, precision of 95.8%, recall of 96.0%, and F1-score of 95.9%. The K-Means algorithm, while less effective, still performed reasonably well with an accuracy of 89.4%. The Deep Neural Network model exhibited superior results with an accuracy of 98.5%, precision of 98.3%, recall of 98.7%, and an F1-score of 98.4%.

Handling the diverse network behaviors and attack vectors within large-scale IoT environments remains a challenge. To address this, the CICIoT2023 dataset was created, and it demonstrated exceptional performance, with Random Forest achieving accuracy, precision, recall, and F1-score all at 99.7% [9]. They also used ensemble learning techniques, such as Adaptive Boosting, achieved an accuracy of 99.18%, precision of 98.68%, and an F1-score of 98.84%, underscoring the effectiveness of various ML strategies in improving IoT security.

Additionally, swarm intelligence has been integrated into neural networks to enhance IoT security, as evidenced by a study using the ToN-IoT dataset, which achieved 99.5% accuracy [10]. A recent study focusing on this dataset proposed an IDS achieving an F1-score of 99.95% with their hybrid 1D CNN-LSTM model [11]. Plus, Another study achieved an accuracy of 89.0% by using a Deep Auto-Encoder on a dataset from the Criminal Investigation Department of Nigeria, further highlighting the potential of deep learning in IoMT security applications [12].

Other studies have evaluated a range of ML models on various datasets. For instance, research using the Bot-IoT dataset demonstrated that Random Forest performed well in binary classification with 99% accuracy, while KNN excelled in multi-class scenarios with the same accuracy [13]. In another study, multiple models such as Logistic Regression, SVM, Decision Tree, Random Forest, and Artificial Neural Networks (ANN) were applied to IoT sensors, achieving a test accuracy of 99.4%, illustrating the effectiveness of diverse ML techniques in distinguishing between normal and anomalous traffic [14]. Random Forest also excelled when applied to the NSL-KDD dataset, achieving 85.34% accuracy [15].

Additional work employing the CIC-IDS2017 dataset showed that Decision Tree and Random Forest classifiers achieved accuracies of 96.44% and 94.45%, respectively, in a self-training IDS that adapts to new threats [16]. A study focusing on IoMT environments proposed an IDS leveraging tree-based ML classifiers combined with filter-based feature selection techniques, demonstrating a detection accuracy of 98.79% on the CICIDS2017 dataset, while maintaining a low false alarm rate of 0.007 [17]. The CIC-IDS 2018 dataset has also been leveraged to dynamically retrain classifiers, with the Decision Tree achieving a notable accuracy of 96.44% [18]. Lastly, in [19], they use MeMalDet to extract optimal features from memory dumps using deep autoencoders in an unsupervised manner, which helps avoid manual feature engineering and got a F1-Score of 98.72%.

These studies collectively highlight the crucial role of advanced machine learning and deep learning models in enhancing IDS for IoT and IoMT. By leveraging robust datasets and innovative architectures, these studies contribute valuable insights into improving intrusion detection capabilities, demonstrating the critical importance of ML and DL techniques in fortifying IoT and IoMT environments.

III. Dataset and Methods

A. Dataset Description

The CIC-IoMT 2024 dataset [20] contains 8,775,013 entries of network traffic data collected from Internet of Medical Things (IoMT) devices. The CICIoMT2024 dataset simulates multiple attack vectors targeting IoMT devices, focusing on five main categories: Distributed Denial-of-Service (DDoS), Denial-of-Service (DoS), reconnaissance, MQTT-specific attacks, and spoofing attacks. Firstly, DDoS and DoS attacks employed methods like ICMP floods that overwhelm devices with echo requests; SYN floods inundating targets with TCP SYN requests to exhaust resources; and TCP/UDP floods exploiting transport layer protocols to overload devices. Secondly, reconnaissance attacks included port and OS scanning, ping sweeps, and vulnerability scans to map network topologies and identify vulnerabilities. Thirdly, MQTT-specific attacks targeted the MQTT protocol—crucial for IoMT communication—using MQTT Connect and Publish floods via custom Python scripts to overwhelm brokers, and malformed data attacks that inject erroneous packets to disrupt communication between brokers and devices. Furthermore, spoofing attacks, particularly ARP spoofing, enabled man-in-the-middle scenarios where attackers intercepted and potentially altered communications. Additionally, Bluetooth Low-Energy (BLE) attacks involved overloading devices through continuous data writing and scanning scripts, resulting in varied device responses from normal operation to disruption. One significant aspect of this dataset is the class imbalance; DDoS traffic makes up 66.6% of the data, while benign traffic is only 2.6%. This imbalance presents a challenge for machine learning models, which may struggle to accurately identify less frequent attack types. To prepare the data for machine learning applications, MinMax scaling was performed on the extracted features, normalizing the data to enhance model performance. These diverse attack vectors provide a thorough basis for testing and evaluating intrusion detection and mitigation models in IoMT environments.

B. t-SNE Visualization

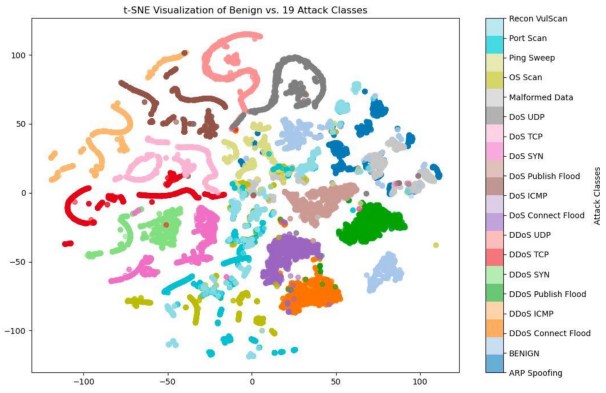

Fig. 1: t-SNE Visualization of 1000 random samples per class

To better understand the separability of attack classes and benign traffic in the CIC-IoMT 2024 dataset, a t-distributed Stochastic Neighbor Embedding (t-SNE) was performed on 1000 random samples from each class, visualizing the high-dimensional feature space in two dimensions 1. The t-SNE plot provides a qualitative assessment of the data distribution and inter-class relationships, which are crucial for developing effective machine learning-based intrusion detection systems (IDS) in IoMT environments.

The t-SNE visualization reveals a well-defined cluster for benign traffic (green) that is distinct from the various attack classes. The compact nature of the benign cluster indicates low intra-class variability, suggesting that benign network traffic exhibits consistent patterns in the feature space. This compactness is promising for classification tasks, highlighting the inherent separability between benign and malicious traffic.

The t-SNE visualization reveals a well-defined cluster for benign traffic (green) that is distinct from the various attack classes. The compact nature of the benign cluster indicates low intra-class variability, suggesting that benign network traffic exhibits consistent patterns in the feature space. This compactness is promising for classification tasks, as it highlights the inherent separability between benign and malicious traffic.

Certain attack classes, such as ”ARP Spoofing” and ”Malformed Data,” exhibit distinct, compact clusters, reflecting higher intra-class consistency and clear separability from other attack types. These attack classes may be easier to detect using conventional classification methods as they possess well-defined features. On the other hand, classes like ”DoS Publish Flood” display elongated and irregular cluster shapes, indicating high intra-class variability. This suggests the presence of multiple sub-modes of attack within the same class, potentially requiring further feature engineering or hierarchical classification techniques to improve detection accuracy.

The overlap between certain attack classes, particularly among DDoS variants, highlights the potential for misclassification in real-world IDS deployments. The positioning of related attack types, such as ”Ping Sweep” and ”Port Scan,” near one another suggests that these attacks share common features, making it difficult for a classifier to differentiate between them without additional feature refinement. This observation motivates the need for further exploration of domain-specific features and hybrid classification models that can reduce the impact of overlapping feature spaces.

C. Model Selection

Stratified K-Fold divides the dataset into K subsets, or folds, ensuring that each fold is representative of the class distribution present in the full dataset. The model is trained on K-1 folds and validated on the remaining fold. This process is repeated K times, with each fold used exactly once as the validation set. By maintaining class balance in each fold, stratified K-Fold minimizes the risk of bias toward the majority class, ensuring a more accurate and reliable evaluation of the model’s performance across both majority and minority classes.

This technique is necessary in scenarios where class imbalance skews model training, leading to poor generalization, especially for underrepresented classes such as rare attacks

in this datasets Without stratification, traditional K-Fold cross-validation might result in folds that fail to adequately represent minority classes, leading to an overestimation of the model's performance on the majority class while failing to capture the complexity of minority class predictions. Moreover, stratified K-Fold cross-validation helps mitigate issues like overfitting, providing a more comprehensive evaluation compared to a simple train-test split. Therefore, by adopting stratified K-Fold cross-validation, we ensure that our models are evaluated across different attack categories, resulting in a balanced performance evaluation across both frequent and rare attack vectors.

D. Models Used

In alignment with the methodology of [13], we benchmarked our system's performance using a variety of models, including Logistic Regression, AdaBoost, Random Forest, and XGBoost, which are well-suited for tabular data and offer a balance between accuracy and computational efficiency. Additionally, we implemented a neural network to evaluate the potential of deep learning in this context. The neural network consisted of three hidden layers with 128, 64, and 32 neurons, respectively, using ReLU activation, and incorporated dropout layers with a 30% rate to prevent overfitting. The output layer utilized softmax activation for multi-class classification. Given the resource constraints of IoMT devices, which limit the feasibility of deploying large models, Random Forest and XGBoost emerged as the most effective solutions, balancing high detection accuracy with practical deployability.

E. Evaluation Metrics

While accuracy is a commonly used metric for evaluating classification models, it can be misleading in the context of imbalanced datasets. In scenarios where one class significantly outnumbers the others, a classifier may achieve high accuracy simply by predicting the majority class for all instances. This results in an overly optimistic assessment of the model's performance, as it fails to account for the correct identification of minority class instances, which are often of greater interest in intrusion detection systems. To address this limitation, we focus on metrics that provide a more nuanced evaluation of classifier performance on imbalanced data:

a) ***Precision****:* Precision measures the accuracy of the positive predictions, indicating the proportion of true positive predictions out of all positive predictions made by the classifier. This metric is particularly important in security applications where false positives can lead to unnecessary interventions. Precision is defined as:

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \tag{1}$$

A high precision score implies that the model has a low false-positive rate, meaning it predicts fewer benign instances as malicious.

b) ***Recall****:* Recall, also referred to as sensitivity, measures the ability of the classifier to correctly identify all relevant positive instances. It reflects the proportion of true positive observations out of all actual positives in the dataset. Recall is critical in cybersecurity tasks where missing an attack can have consequences. Recall is defined as:

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \tag{2}$$

A high recall indicates that the model successfully identifies most of the malicious instances, reducing the false-negative rate.

c) ***F1-Score****:* The F1-score represents the harmonic mean of precision and recall, providing a balanced metric that combines both values. It is especially useful in cases of class imbalance, as it emphasizes the trade-off between precision and recall. The F1-score is defined as:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

This score ensures that both false positives and false negatives are considered equally when evaluating the performance of the classifier.

F. Feature Reduction Methods Used

a) ***Principal Component Analysis (PCA)****:* PCA is a popular linear dimensionality reduction technique that transforms the original features into a lower-dimensional space while retaining most of the variance in the data. By projecting the data onto a lower- dimensional subspace, PCA helps reduce computational complexity and mitigate the curse of dimensionality, particularly beneficial when dealing with high-dimensional datasets.

b) ***Recursive Feature Elimination with Cross-Validation (RFECV)****:* RFECV is an effective feature selection method that improves model performance by iteratively removing less important features. It works by training a model on subsets of features, eliminating the least significant ones at each step, and evaluating its performance through cross-validation. This process identifies the optimal feature set that best contributes to the model's predictive power.

RFECV helps enhance generalization by selecting relevant features while discarding redundant or noisy ones. The use of cross-validation ensures robustness and prevents overfitting, making it a valuable method for high-dimensional datasets where feature interactions are complex.

IV. Results

Our evaluation begins with a comparison of baseline metrics, as shown in Table **??**, which highlights the state-of-the-art results from the original CICIoMT2024 paper. These results serve as a benchmark for assessing the improvements achieved through our experimental modifications. The table also details the impact of feature selection techniques, such as Principal Component Analysis (PCA) for dimensionality reduction and Recursive Feature Elimination (RFE) for identifying optimal feature subsets. Notably, the use of stratified K-Fold cross-validation ensures a balanced evaluation across different data folds, providing a robust performance metric

TABLE I: Classifier Performance Metrics with Maximum Values per Classifier Highlighted

Classifier	Metric	Kfold Avg (Std) (Baseline)	Kfold Avg (Std) (PCA)	Kfold Avg (Std) (RFECV)
Logistic Regression	Precision	**0.7127 (0.00)**	0.6229 (0.01)	0.6930 (0.01)
	Recall	**0.7524 (0.00)**	0.7169 (0.00)	0.7496 (0.00)
	F1 Score	**0.6741 (0.00)**	0.6134 (0.00)	0.6710 (0.00)
AdaBoost	Precision	**0.5116 (0.13)**	0.4768 (0.05)	0.4713 (0.07)
	Recall	**0.4806 (0.14)**	0.3642 (0.09)	0.4292 (0.10)
	F1 Score	**0.3879 (0.14)**	0.2797 (0.07)	0.3381 (0.09)
Random Forest	Precision	**0.9981 (0.00)**	0.9189 (0.00)	**0.9981 (0.00)**
	Recall	**0.9981 (0.00)**	0.9148 (0.00)	**0.9982 (0.00)**
	F1 Score	**0.9981 (0.00)**	0.9000 (0.00)	**0.9981 (0.00)**
XGBoost	Precision	**0.9977 (0.00)**	0.9417 (0.00)	0.9976 (0.00)
	Recall	**0.9978 (0.00)**	0.9429 (0.00)	0.9977 (0.00)
	F1 Score	**0.9977 (0.00)**	0.9412 (0.00)	0.9976 (0.00)
Neural Network	Precision	**0.7514 (0.00)**	N/A	N/A
	Recall	**0.7854 (0.00)**	N/A	N/A
	F1 Score	**0.7234 (0.00)**	N/A	N/A

TABLE II: F1-Score Comparison: CICIoMT 2024 Benchmarks vs. Our Test Results

Model	CICIoMT 2024 F1-Score [20]	F1-Score on Test Set	% Improvement
Logistic Regression	0.4310	0.6561	+52.22%
AdaBoost	0.3011	0.3426	+13.78%
Random Forest	0.9074	0.9948	+9.63%
XGBoost	N/A	0.9926	N/A
Neural Network	0.5791	0.6947	+19.96%

for each model. Table **??** further illustrates the substantial improvements in F1-scores achieved by our optimized models over the original benchmarks. Ensemble methods, particularly Random Forest and XGBoost, significantly outperformed other classifiers, with Random Forest achieving an F1-score of 99.48%, marking a 9.63% improvement over the benchmark. The neural network, despite achieving an F1-score of 69.47%, showed a 19.96% improvement over its benchmark, demonstrating some effectiveness but falling short of the ensemble models in handling class imbalance and complex attack vectors. These results underscore the importance of feature selection and model optimization in enhancing intrusion detection performance within resource-constrained IoMT environments.

V. Discussion of Results

The experimental outcomes presented in Tables I and II demonstrate significant enhancements in intrusion detection performance within IoMT environments. The Random Forest classifier achieved a mean F1-score of 99.48%, surpassing the CICIoMT 2024 benchmark by 9.63%, while Logistic Regression and AdaBoost exhibited notable increases of 52.22% and 13.78% in F1-score respectively. These advancements are primarily due to our adoption of comprehensive evaluation techniques. Employing stratified K-Fold cross-validation ensured that each fold accurately reflected the dataset's class distribution, enabling the models to learn effectively from both majority and minority classes. This contrasts with baseline methods that may have inadequately represented minority classes, potentially leading to biased models.

The superior performance of ensemble methods like Random Forest and XGBoost underscores their ability to capture complex, nonlinear patterns inherent in IoMT network traffic. Their high precision and recall scores, as detailed in Table I, indicate a strong capability to distinguish between benign and malicious activities, which is crucial for effective intrusion detection.

The neural network achieved an F1-score of 69.47%, improving by 19.96% over its baseline. However, this performance still lagged behind the ensemble methods, highlighting challenges in handling class imbalance and high intra-class variability. Additionally, neural networks impose substantial computational overhead and require larger memory footprints, rendering them less practical for resource-constrained IoMT environments. This finding reinforces that ensemble methods like Random Forest and XGBoost are more suitable for IoMT applications, offering a superior balance between performance and efficiency.

Implementing Recursive Feature Elimination with Cross-Validation (RFECV) further enhanced model efficiency by reducing feature dimensionality without compromising performance. For instance, the Random Forest classifier maintained its high F1-score while reducing the feature set by 44.45% (retained 25/45 original features). This reduc-

tion is particularly significant for deployment on resource-constrained IoMT devices, where computational resources are limited. In contrast, Principal Component Analysis (PCA) did not yield comparable performance improvements, suggesting that feature selection methods preserving the original feature space are more effective in this context. RFECV's ability to identify and retain the most informative features contributes to the model's interpretability and efficiency.

Despite these significant performance improvements and feature set reductions, deploying our optimized models on resource-constrained IoMT devices presents practical challenges related to memory footprint and computational requirements. Specifically, the Random Forest model, even after feature reduction, remains large—reducing from 732.90 MB to 496.47 MB—which is impractical for deployment on IoMT devices that typically have limited memory capacities ranging from tens of kilobytes to a few megabytes [21]. In contrast, the XGBoost model is considerably smaller, with the base model at 4.37 MB and the reduced model at 4.23 MB, making it more suitable for deployment on resource-constrained devices.

Therefore, while Random Forest offers superior performance, XGBoost provides a better balance between accuracy and resource consumption for real-world applications. Future work will focus on further optimizing model sizes through techniques such as model compression and pruning to facilitate deployment on devices with stringent resource constraints.

The balance achieved between high detection accuracy and reduced model complexity addresses a critical challenge in the field. Our approach facilitates the practical deployment of robust intrusion detection systems on IoMT devices by optimizing models for performance and resource efficiency.

These findings have important implications for enhancing cybersecurity in IoMT networks. The significant improvements over the CICIoMT 2024 benchmarks, particularly in F1-scores highlighted in Table II, validate the effectiveness of our methodologies. Future work could explore integrating these optimized models into real-world IoMT systems and assess their performance under dynamic network conditions.

VI. Conclusion

The rapid expansion of the Internet of Medical Things (IoMT) offers significant advancements in healthcare but also exposes critical vulnerabilities to cyber threats. Traditional security measures often fall short in protecting these devices due to resource constraints and evolving attack complexities.

In this study, we developed a customized Intrusion Detection System (IDS) tailored for IoMT environments using the CICIoMT2024 dataset. By addressing challenges like class imbalance and high dimensionality, we achieved a state-of-the-art weighted F1-score of 99.48% while reducing the feature set by 44.45%. Our optimized Random Forest and XGBoost models outperformed baseline approaches and proved suitable for deployment on resource-constrained IoMT devices. The strength of our research lies in combining effective machine learning techniques with practical implementation strategies. By employing stratified K-Fold cross-validation and Recursive Feature Elimination with Cross-Validation (RFECV), we ensured robust performance across various attack types and enhanced model efficiency without compromising accuracy.

While our optimized models demonstrate high accuracy and reduced feature sets, practical deployment on IoMT devices necessitates consideration of resource limitations. Our results indicate that XGBoost, due to its smaller model size, is more suitable for deployment on resource-limited devices compared to Random Forest. Future research will explore additional model optimization techniques to reduce memory footprint and computational requirements, ensuring efficient and effective intrusion detection in IoMT environments.

Acknowledgment

The authors would like to express their gratitude to Anna Yakoub and Sameer Alawnah for their invaluable support and insightful contributions to the development of this project. Their guidance has been instrumental in enhancing the quality and scope of this work.

References

[1] Dwivedi, R.; Mehrotra, D.; Chandra, S. Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review. Journal of oral biology and craniofacial research 2022, 12, 302–318.

[2] Mbengue, S.M.; Diallo, O.; El Hadji, M.N.; Rodrigues, J.J.; Neto, A.; Al-Muhtadi, J. Internet of medical things: Remote diagnosis and monitoring application for diabetics. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2020, pp. 583–588.

[3] Subramaniam, E.V.D.; Srinivasan, K.; Qaisar, S.M.; Pławiak, P. Interoperable IoMT Approach for Remote Diagnosis with Privacy-Preservation Perspective in Edge Systems. Sensors 2023, 23, 7474.

[4] S. Dadkhah, H. Mahdikhani, P. K. Danso, A. Zohourian, K. A. Truong and A. A. Ghorbani, "Towards the Development of a Realistic Multidimensional IoT Profiling Dataset," 2022 19th Annual International Conference on Privacy, Security & Trust (PST), Fredericton, NB, Canada, 2022, pp. 1-11, doi: 10.1109/PST55820.2022.9851966.

[5] G. Zachos, I. Essop, G. Mantas, K. Porfyrakis, J. C. Ribeiro, and J. Rodriguez, "An Anomaly-Based Intrusion Detection System for Internet of Medical Things Networks," Electronics, vol. 10, no. 21. MDPI AG, p. 2562, Oct. 20, 2021. doi: 10.3390/electronics10212562.

[6] V. Ravi, T. D. Pham and M. Alazab, "Deep Learning- Based Network Intrusion Detection System for Internet of Medical Things," in IEEE Internet of Things Magazine, vol. 6, no. 2, pp. 50-54, June 2023, doi: 10.1109/IOTM.001.2300021.

[7] A. Aljuhani, A. Alamri, P. Kumar and A. Jolfaei, "An Intelligent and Explainable SaaS-Based Intrusion Detection System for Resource-Constrained IoMT," in IEEE Internet of Things Journal, vol. 11, no. 15, pp. 25454-25463, 1 Aug.1, 2024, doi: 10.1109/JIOT.2023.3327024.

[8] M. Zubair, A. Ghubaish, D. Unal, A. Al-Ali, T. Reimann, G. Alinier, M. Hammoudeh, and J. Qadir, "Secure Bluetooth communication in smart healthcare systems: A novel community dataset and intrusion detection system," Sensors, vol. 22, no. 21, p. 8280, Oct. 2022.

[9] E. C. P. Neto, S. Dadkhah, R. Ferreira, A. Zohourian, R. Lu, and A. A. Ghorbani, "CICIoT2023: A Real-Time Dataset and Benchmark for Large-Scale Attacks in IoT Environment," Sensors, vol. 23, no. 13. MDPI AG, p. 5941, Jun. 26, 2023. doi: 10.3390/s23135941.

[10] S. Nandy, M. Adhikari, M. A. Khan, V. G. Menon and S. Verma, "An Intrusion Detection Mechanism for Secured IoMT Framework Based on Swarm-Neural Network," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 5, pp. 1969-1976, May 2022, doi: 10.1109/JBHI.2021.3101686.

[11] S. Yaras and M. Dener, "IoT-Based Intrusion Detection System Using New Hybrid Deep Learning Algorithm," Electronics, vol. 13, no. 6. MDPI AG, p. 1053, Mar. 12, 2024. doi: 10.3390/electronics13061053.

[12] J. B. Awotunde, K. M. Abiodun, E. A. Adeniyi, S. O. Folorunso, and R. G. Jimoh, "A Deep Learning-Based Intrusion Detection Technique for a Secured IoMT System," Communications in Computer and Information Science. Springer International Publishing, pp. 50–62, 2022. doi: 10.1007/978-3-030-95630-1_4.

[13] A. Churcher et al., “An Experimental Analysis of Attack Classification Using Machine Learning in IoT Networks,” Sensors, vol. 21, no. 2. MDPI AG, p. 446, Jan. 10, 2021. doi: 10.3390/s21020446.

[14] M. Hasan, Md. Milon Islam, I. Islam, and M. M. A. Hashem, “Attack and Anomaly Detection in IoT Sensors in IoT Sites Using Machine Learning Approaches,” Internet of Things, vol. 7, p. 100059, May https://doi.org/10.1016/j.iot.2019.100059.

[15] M. Anwer, S. M. Khan, M. U. Farooq, and . Waseemullah, “Attack Detection in IoT using Machine Learning”, Eng. Technol. Appl. Sci. Res., vol. 11, no. 3, pp. 7273–7278, Jun. 2021.

[16] P. Radoglou-Grammatikis, P. Sarigiannidis, G. Efstathopoulos, T. Lagkas, G. Fragulis and A. Sarigiannidis, ”A Self-Learning Approach for Detecting Intrusions in Healthcare Systems,” ICC 2021 - IEEE International Conference on Communications, Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICC42927.2021.9500354.

[17] G. Balhareth and M. Ilyas, “Optimized Intrusion Detection for IoMT Networks with Tree-Based Machine Learning and Filter-Based Feature Selection,” Sensors, vol. 24, no. 17. MDPI AG, p. 5712, Sep. 02, 2024. doi: 10.3390/s24175712.

[18] J. A. Alzubi, O. A. Alzubi, I. Qiqieh and A. Singh, ”A Blended Deep Learning Intrusion Detection Framework for Con-sumable Edge-Centric IoMT Industry,” in IEEE Transactions on Consumer Electronics, vol. 70, no. 1, pp. 2049-2057, Feb. 2024, doi: 10.1109/TCE.2024.3350231.

[19] P. Maniriho, A. N. Mahmood, and M. J. M. Chowdhury, “MeMalDet: A memory analysis-based malware detection framework using deep autoencoders and stacked ensemble under temporal evaluations,” Computers & Security, vol. 142. Elsevier BV, p. 103864, Jul. 2024. doi: 10.1016/j.cose.2024.103864.

[20] S. Dadkhah, E. Carlos Pinto Neto, R. Ferreira, R. Chukwuka Molokwu, S. Sadeghi, and A. Ghorbani, “CICIoMT2024: Attack Vectors in Healthcare devices-A Multi-Protocol Dataset for Assessing IoMT Device Security.” MDPI AG, Feb. 16, 2024. doi: 10.20944/preprints202402.0898.v1.

[21] S. Challa, M. Wazid, A. K. Das, N. Kumar, and M. Jo, ”Secure Signature-Based Authenticated Key Establishment Scheme for Future IoT Applications,” *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 2, pp. 461–474, March-April 2021.