Meta-IDS: Meta-Learning Based Smart Intrusion Detection System for Internet of Medical Things (IoMT) Network

Umer Zukaib , Xiaohui Cui , Chengliang Zheng , Mir Hassan , *Member, IEEE*

***Abstract*—At the forefront of the smart healthcare revolution, the Internet of Medical Things (IoMT) emerges as a transformative force, facilitating real-time data collection and processing. In this dynamic landscape, interconnected IoMT devices enable healthcare practitioners with decision-making capabilities, integrating various Artificial Intelligence (AI) models. However, the widespread use of communication protocols introduces significant security challenges, leading to concerns about data transmission and control. In response, the need for a real-time and highly accurate Intrusion Detection System (IDS) in the IoMT network becomes crucial to address emerging threats and vulnerabilities. Acknowledging this critical need, our research delves deep into understanding and navigating network traffic nuances, presenting tailored and robust cybersecurity solutions finely tuned to IoMT's distinctive challenges. In the realm of IoMT, our groundbreaking Meta-Intrusion Detection System (Meta-IDS) goes beyond conventional approaches by seamlessly integrating signature-based and anomaly-based techniques, providing a comprehensive security paradigm. Uniquely crafted to address IoMT's constraints, including limited processing power, high efficiency and adaptability. With a steadfast commitment to privacy preservation. Rigorously tested across diverse datasets, Meta-IDS excels in detecting both known and zero-day threats, making a substantial and impactful contribution to fortifying IoMT cybersecurity and demonstrate the exceptional performance, yielding promising results while maintaining minimal execution time, surpassing benchmarks set by state-of-the-art counterparts.**

***Index Terms*—Intrusion Detection Systems, Meta Learning, Internet of Medical Things, Artificial Intelligence, Cyber Security, zero-day attacks, Anomaly detection**

I. Introduction

THe proliferation of connected devices in our daily lives is continually expanding. According to Ahmad et al. [1], there will be 27 billion smart devices with internet connections by 2025, more than double the number of IoT devices in 2021 [26]. The integration of Artificial Intelligence (AI) into medical devices has led to significant breakthroughs in healthcare. The "Internet of Medical Things" (IoMT) represents the convergence of IoT and healthcare technologies, presently constituting 30 to 40% of all IoT devices [37].

The adoption of IoT technology in the medical sector brings benefits such as improved remote health services, streamlined healthcare operations, and enhanced patient health monitoring. However, security and privacy emerge as significant concerns in IoT-enabled healthcare operations [47]. Many of these medical devices exhibit critical vulnerabilities, posing security risks on public networks and allowing adversaries to exploit sensitive information [58]. Attackers often exploit known vulnerabilities and advanced persistent threats (APT) to leak information from IoMT devices [60], occasionally endangering human lives. Consequently, security monitoring becomes a top priority in IoMT-based healthcare [15].

Numerous approaches, including malicious traffic injection, man-in-the-middle attacks, denial of service, and others, are employed to breach networks. Countermeasures such as attack detection, mitigation, and prevention are implemented to secure networks and safeguard data [56]. Researchers have developed various strategies for detecting and preventing cyberattacks, including vulnerability management, end-device monitoring, log monitoring, introduction of preventive measures, and intrusion detection [17]. Intrusion detection is the most commonly used technology in IoMT devices for identifying network attacks and security issues, utilizing techniques like signature-based rules, security policies, and network-traffic anomalies [45] [6].

Standard security-detection approaches prove ineffective as attackers constantly evolve their strategies, employing advanced hacking tactics. Reverse engineering and network monitoring can compromise security policies [54]. Machine learning (ML) and deep learning (DL) play pivotal roles in intrusion detection, with researchers leveraging these technologies to develop intelligent Intrusion Detection Systems (IDS) for identifying cyberattacks on computer networks [18] [39] [52].

Traditional strategies struggle to address real-time detection and identification of unknown attacks within network data. Their static nature limits them to recognizing only known security threats, rendering them vulnerable to new intrusions like zero-day attacks. The inherent challenges of applying these strategies in the IoMT context, such as specific features, resource constraints, and privacy preservation, motivate the development of a secure IDS solution. Our work aims to overcome these limitations by creating an IDS capable of detecting new threats and adapting to dynamic intrusions in the diverse IoMT landscape. This solution prioritizes real-world deployment, ensuring minimal execution time, and

Manuscript received April 19, 2005; revised August 26, 2015. *(Corresponding author: Xioahui Cui)*

The authors are with the Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Hubei, China and the Department of Information Engineering and Computer Science, University of Trento, Trento, 38123, Italy (email: umerzukaib@whu.edu.cn, xcui@whu.edu.cn, chengliang@whu.edu.cn, mir.hassan@unitn.it)

TABLE I: Abbreviation and definition

Abbreviation	Definition
IDS	Intrusion Detection System
APT	Advanced Persistent Threats
PSO	Particle Swarm Optimization
AMQP	Advanced Messaging Queuing Protocols
MQTT	Messaging Queuing Telemetry Transport
BO-TPE	Bayesian-optimization-based Tree-structured parzen-estimator
RF-RFE	Random forest-based Recursive-feature elimination
DICOM	Digital-Imaging and Communication in Medicine
HL7	Health-Level Seven International
IEEE 11073	Health informatics - Personal health device communication
IHE	Integrating the Healthcare Enterprise
MDG	Mean Reduction in Gini
HPO	Hyper Parameter Optimization
PHMS	Patient Health Monitoring System

efficient resource utilization, ultimately enhancing detection accuracy and improving security in the ever-evolving IoMT environment.

This work aims to develop a sophisticated intrusion detection framework tailored for diverse IoMT setups, proficient in detecting both known and unknown attacks within minimal time. Embracing a hybrid and proactive approach, our smart IDS serves as a robust defense mechanism against evolving threats. By employing five base learners and leveraging ensemble learning for meta-learning predictions, the proposed framework efficiently identifies known and unknown attacks, streamlining threat pattern analysis. The Meta-IDS stands out for its swift training capabilities, outperforming traditional statistical techniques. With a crucial role in ensuring security and privacy preservation of sensitive IoMT data. The Meta-IDS excels in accurate intrusion detection within minimal time, making it suitable for both spatial and temporal aspects, particularly effective in detecting intrusions within the complex network dynamics of inter IoMT and intra IIoT environments.

Our main contribution in this research paper is as follows:

- Our innovative Meta-IDS technique, comprising five base learners and a meta-learner, is specifically designed to detect known attacks targeting IoMT networks.
- To enhance the efficiency of our Meta-IDS models, we leverage the Bat algorithm and a parzen-estimator based on a tree structure, effectively fine-tuning hyper-parameters for optimal performance.
- Addressing the challenge of zero-day attacks, we introduce Mean Shift Clustering and biased classifiers, providing robust detection capabilities against emerging threats.
- Our approach extends to a real-time architecture, ensuring the seamless deployment of Meta-IDS in authentic healthcare settings, reinforcing its practical applicability in dynamic IoMT environments.

This manuscript unfolds with a review of related work in Section II. Our novel intrusion detection methodology for the IoMT is presented in Section III, while the underlying dataset is introduced in Section IV. Section V details experiments and performance evaluations. The paper concludes with reflections and contributions in Section VII.

II. Related Work

This section explores state-of-the-art IDS utilizing ML and DL techniques. The IoMT network consists of various IoT devices connected to patients' bodies for data collection, with the IoT gateway linked to the conventional grid. Remote monitoring of patients' health is enabled through the internet. A successful attack on the IoMT network, breaching the system's security, can have severe consequences, including the loss of patient lives. Numerous studies have focused on discovering and managing cyberattacks on the IoMT network. Traditionally, IoMT networks relied on signature and anomaly-based intrusion detection methods. While signature-based or policy-based techniques are ineffective against zero-day attacks, which exploit advanced persistent threats, anomaly-based methods can effectively detect such attacks over the network [14]. Yacoub et al. [57] explored privacy and security concerns in IoMT networks, highlighting the use of ML-based solutions for detecting network attacks. Emphasized the need for an efficient IDS, given the constraints of limited processing power and storage space in IoMT devices, where applying security protocols at the device level is challenging.

AI-based Network IDS was proposed by Park et al. [44] which effectively addresses data imbalance concerns, enhancing overall performance. The generative model skillfully generates synthetic data for minor attack traffic, surpassing results from autoencoder-driven DL models. However, the approach lacks consideration for runtime complexity, and a gap persists in methodologies effective for detecting new attacks and offering real-world solutions. Javeed et al [27] proposed SDN-orchestrated DL-based IDS, which utilize SDN architecture for reconfiguration and a Bi-LSTM model for attack identification. Simulations on the CICIDS-2018 dataset validate its superiority over recent security solutions, However, the study lacks practical implications and applicability to the diverse IoMT network. It doesn't address its potential as a hybrid IDS solution capable of identifying both known and unknown attacks in real-world scenarios.

Yang et al. [59] crafted a hybrid multitiered model, integrating both signature-based IDS and anomaly-based IDS, demonstrated on CICIDS2017 datasets for effective detection of known and unknown attacks. However, its applicability to the IoMT paradigm is limited, as the dataset adaptation falls short of meeting critical healthcare requirements. Kaur et al.[29] present D-Sign, a sophisticated DL system designed for hybrid intrusion detection and the generation of signatures for unknown web attacks. D-Sign demonstrates exceptional performance showcasing minimal False Negatives and False Positives. It is important to note, however, that the study primarily focuses on web attacks, limiting its applicability to a broader range of diverse attack scenarios. Bovenzi et al. [10] proposed H2ID, two-stage hierarchical IDS, utilizing anomaly detection with an Auto Encoder and attack classification through soft-output classifiers. However, its runtime complexity poses a limitation for real-time deployment in the critical IoMT environment.

Kumar et al. [35] introduced an ensemble method using the ToN-IoT dataset, achieving improved results and proposing a real-time deployment architecture for edge-cloud-based IoMT environments. However, the study primarily focuses on differentiating between attack and normal traffic using signature or rule-based methods, which may not effectively detect new attacks. Almogren et al. [7] introduced a deep belief network (DBN) for intrusion detection, surpassing current techniques. Nevertheless, the study is primarily centered on basic attack classification, and the computational complexity of the network poses challenges for practical real-world application.

Radoglou et al. [46] introduced an active learning approach to dynamically retrain supervised classifiers to perform intrusion detection task. While the evaluation demonstrated its effectiveness against HTTP and TCP/Modbus cyberattacks, the study's narrow focus on a single domain restricts its applicability to diverse environmental scenarios. Li et al. [36] proposed a federated learning framework, DeepFed, for privacy-preserving intrusion detection in industrial CPSs. The study introduces a secure communication protocol based on the Paillier cryptosystem. While effective in detecting various cyber threats, the study lacks information on runtime complexity and applicability to handle multiple attack patterns. Saheed et al. [50] proposed a cryptographic security solution and intrusion detection based on ML and DL for IoMT cyber-attack detection. The study demonstrates improved IDS performance through PSO feature selection. However, the study does not adequately illustrate its dynamic applicability in a cross-domain IoMT network.

Khan et al. [32] introduced an SDN-based LSTM and CNN framework, showcasing promising results in detecting malware in the IoMT network. However, the study lacks information on novel attack detection, how well the system adapts to the IoMT environment, and its time complexity. Zhang et al [61] introduced SecFedNIDS model, utilizing layer-wise relevance propagation for the detection of poisoned data based on path similarity. Although the study successfully safeguards against poisoning attacks, its exclusive emphasis on data poisoning limits its applicability to diverse cyber attacks within the IoMT network. The algorithms discussed in recent studies [20], [23], [21], and [22] have significant potential for adoption in advancing and refining IDS to operate more effectively. For instance, the GAN-based model [20] may detect synthetic data in real network traffic, enhancing IDS applicability in real-world scenarios. Similarly, leveraging Unstructured Actor-Modeling with Discrete-Time Markov Chains (UAM with DTMC) and Probabilistic Computation Tree Logic (PCTL) [23], along with the TBDB algorithm [21], may aid in threat detection and enhance security in IoMT and IIoT environments. Moreover, the Com-DDPG approach [22] may demonstrate its effectiveness in task offloading for robust IDS deployment across diverse network domains.

The pressing need for a comprehensive system that effectively addresses the challenges found in state-of-the-art research studies has led to the development of our proposed Meta-IDS. This innovative solution is designed to tackle the complexities of safeguarding sensitive health-related data in IoMT networks. By swiftly detecting both known and zero-day attacks while operating within minimal execution time and resource constraints, our Meta-IDS fills a crucial research gap. Its integration of IoMT-specific features enhances its efficacy in addressing evolving threats, offering a robust and holistic solution to the pressing security concerns in healthcare environments.

A comparison of studies based on ML/DL IoMT network attacks is provided in Table II.

III. Proposed Methodology

We employed network traffic and patient biometric data to enhance attack detection in the Internet of Medical Things (IoMT). Figure 1 depicts the IoMT network utilizing ML and DL for cybersecurity.

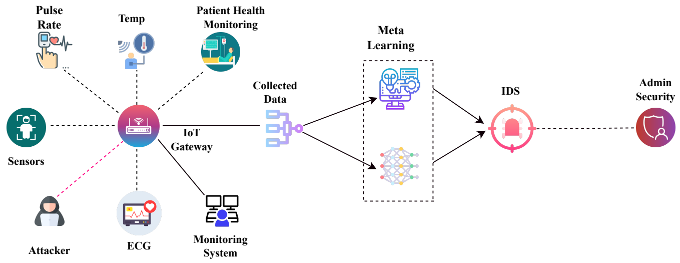


Fig. 1: IoMT-based intrusion detection system

The IDS integrates IoT sensors, a network-traffic controller, an IoT gateway, and ML/DL pipelines for data preparation. Various sensors (pulse-rate, temperature, ECG, etc.) are employed. The IoMT network utilizes MQTT, CoAP, AMQP, DDS messaging protocols based on healthcare requirements. The IoT gateway collects sensor data through wired and wireless communication, and after data-preprocessing, it is forwarded to IDS for tuning and mitigating false positives.

A. System architecture

The Meta-IDS, depicted in Figure 2, utilizes meta-learning for IoMT network intrusion detection. It starts with data collection from network traffic and IoMT sensors, followed by crucial feature engineering and data preprocessing. SMOTE handles class imbalance, and Recursive Feature Elimination (RFE) and Linear discriminant analysis (LDA) improve model performance. The meta-learning process occurs in two steps. The first involves Hyperparameter Optimization (HPO) for weak learners like Decision Trees (DT), Random Forests (RF), AdaBoost, Multinomial Naïve Bayes (MNB), Neural Networks (NN), Multi-Layer Perceptrons (MLP), using the BAT algorithm. The second step utilizes XGBoost (meta-learner) with HPO by Bayesian-Optimization Based Tree-Structured Parzen-Estimator (BO-TPE). The Meta-Learner conducts signature-based intrusion detection, and anomaly detection is performed with Mean-shift clustering. BO-TPE for HPO and two biased classifiers optimize the process, reducing false positives and false negatives while ensuring accurate classification. Table III provides detailed algorithm insights.

B. Data Pre-processing

Addressing class imbalance is crucial in IoT data, where standard samples outnumber attack samples, potentially biasing the model. Resampling techniques like SMOTE [12]

TABLE II: ML and DL-based methods for the detection of IoMT assaults

Article	Used Dataset	Technique	Results (Accuracy)	Limitation
[42]	ToN-IoT	Swarm NN	99 %	Dataset contain only network-traffic
[49]	KDDCup-99	Ensemble Classifier	93 %	The dataset isn't suitable for attacks using IoMT.
[35]	ToN-IoT	Ensemble Classifier	96.3 %	Only have data about network traffic
[46]	CICIDS 2017	Random Forest	94.45 %	Dataset only contains network-traffic features
[51]	BoT-IoT	Swarm NN	99 %	The dataset isn't suitable for attacks using IoMT.
[50]	NSL-KDD	PSO-RF	99.7 %	Dataset contain network-traffic data, not applicable for IoMT-attacks
[9]	NF ToN-IoT	Swarm NN	89 %	Results need improvement

TABLE III: ML and DL-based methods for the detection of IoMT assaults

Steps	Algorithms	Descriptions	Impact on performance
Data pre-processing	SMOTE	SMOTE addresses class imbalance and prevents misclassification by generating high-quality samples for the minority class.	Improves data quality and improves classification accuracy
Feature engineering	RFE	A user-friendly and effective feature-selection method adept at identifying the most relevant features in a dataset for accurate predictions	Improve model performance
	LDA	Project high-dimensional data into a low-dimensional space to mitigate the curse of dimensionality and minimize resource costs.	Improve training efficiency and model performance
Meta-learning	Weak learner (DT, RF, AdaBoost, NN, MLP, MNB)	Employed diverse supervised algorithms as base-learners for attack classification	Overcome the single classifier's limitations.
	Bat algorithm	Used a bat algorithm, a meta-heuristic algorithm simulating bat eco-location behavior, to globally optimize weak-learner hyperparameters for improved performance.	It improves the performance and accuracy
	Meta learner (XGBoost)	Meta-learner attains higher accuracy by learning from pre-trained weak learner models	Mitigates false predictions and enhances performance
Zero-day attack detection	Mean shift clustering	Utilized mean shift clustering for detecting unknown attacks in IoMT network, specifically targeting zero-day attacks in new samples	Perform well for detection of unknown attacks
	BO-TPE	We used tree structure parzen estimator for HPO of mean shift clustering algorithm	Improve the performance of model
	Biased classifier	Utilized mean-shift clustering algorithm's outcomes to train two biased classifiers, minimizing errors and enhancing the model's capability to detect novel attacks	

are vital. Unlike random sampling, SMOTE generates new minority-class samples, ensuring balanced datasets.

We utilized SMOTE to address class imbalance, ensuring a robust and unbiased model in our approach.

Algorithm 1 Algorithm for Data Pre-processing

```
Function SMOTE $(D_{minority}, N_{percent}, k)$
$D_{smoted} \leftarrow []$
for $(i \leftarrow 1$ to $nrowD_{smoted})$ do
    $nn \leftarrow kNN(D_{minority}, N_{percent}, k)$
    $N_i \leftarrow D_{minority}/100$
    while $N_i \neq 0$ do
        $neighbour \leftarrow select - random(nn)$
        $gap \leftarrow range - random(0, 1)$
        $diff \leftarrow neighbour - D_i$
        $synth \leftarrow D_i + gap * diff$
        $D_{smoted} \leftarrow append(D_{smoted}, synth)$
        $N_i \leftarrow N_i - 1$
    end while
end for
return $D_{smoted}$
```

1) Data normalization : After applying SMOTE for balancing the dataset, we perform data normalization. Utilizing a label encoder, categorical attributes are converted to numeric representations. This step is essential as machine learning models require numeric inputs. Normalization is then carried out to prevent biases in ML results caused by large feature scales. The data is normalized to have a mean of 0 and a standard deviation of 1, ensuring optimal support for the machine learning model.

$$x_n = \frac{x - \mu}{\sigma} \tag{1}$$

Where x is a component of the original feature and μ and σ represent sample means and variances.

C. Feature Engineering

After initial data preprocessing, a high-quality dataset is generated, but further refinement is crucial for optimal feature selection. Feature engineering, performed before feeding data into ML models, eliminates noise and irrelevant features, enhancing data quality. We employ Recursive Feature Elimination based on Random Forest (RF-RFE), utilizing mean decrease in accuracy (MDA) as demonstrated in Equation (2) to assess variable importance. This process ensures that only essential traits are retained for more accurate and efficient predictions.

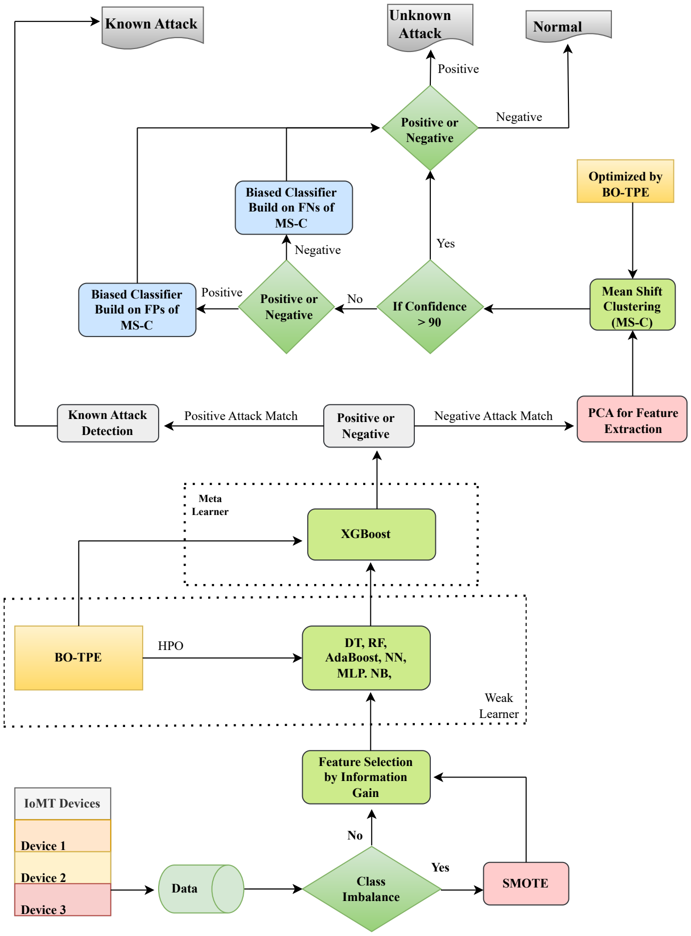


Fig. 2: Proposed framework of Meta-IDS

$$W_R\left(X_i\right) = \frac{\sum_{t\in\beta} VI^{(t)}\left(X_i\right)}{n \text{ tree}} \tag{2}$$

Where β shows out-of-bag observations for a tree (t), VI shows the variable–importance X_i in the tree (t)

Algorithm 2 Recursive feature elimination

```
Input: Training set $(pKa)$
    Set of features $(F = \{f_1, \ldots, f_m\})$
    Number of features to select $(N = \{M, \ldots, 10, 5\})$
Output: Set of features providing highest accuracy $\rightarrow F'$
for each $n$ in $N$ do
    Perform RFE and select $n$ features $\rightarrow F^n$
    Train with RF using $F^n$
    Compute accuracy of model without bag prediction $\rightarrow r^2_{N^i}$
    if $r^2_{N^i} > r^2_{N^{i-1}}$ then
        $F' \leftarrow F^n$
    end if
    $F = F^n$
end for
```

1) Linear Discriminant Analysis: For effective dimensionality reduction, we adopt LDA, chosen for its capability to reduce data dimensionality while improving class separability, making it highly effective for classification datasets by simplifying the establishment of cutoff points.

D. Proposed Meta-IDS

The Intrusion Detection System (IDS) classifies signature-based and anomaly-based attacks. Signature-based IDS identifies known attack patterns using supervised ML, but struggles with new patterns. Anomaly-based IDS distinguishes normal from unknown attacks using unsupervised learning, assuming new attacks share statistical similarities with known attacks. The Meta-IDS excels in detecting both signature-based (known) and anomaly-based (zero-day) attacks, ensuring high accuracy, with minimal execution time.

1) Signature-based IDS: Following data pre-processing and feature engineering, labeled data is trained using meta-learning to construct the signature-based Intrusion Detection System (IDS). In the meta-learning phase, weak learners (DT, AdaBoost, RF, MLP, NN, MNB) and a meta-learner (XGBoost) are distinguished.

Decision Tree (DT) is a tree-based model with multiple tunable hyper-parameters, such as minimum and maximum sample split, sample nodes, tree depth, sample leaf, weight fraction of leaf, etc. Random Forest (RF) is an ensemble learning approach that integrates multiple decision trees using the majority-voting rule. Neural Network (NN) is inspired by the human brain's information processing, identifying hidden patterns and improving over time. AdaBoost is a boosting technique that enhances machine learning algorithm performance by reassigning weights to each instance. Multi-Layer Perceptron (MLP) complements the feed-forward NN and is suitable for both small and large datasets. Multinomial Naive Bayes (MNB) is known for building a fast machine learning model with reliable predictions.

Algorithm 3 Meta-IDS

```
Input: Training dataset $T = \{(x_1, c_1), (x_2, c_2), \ldots (x_n, c_n)\}$
    Base level classifier $L_1, \ldots L_k$
    Meta learner classifier $L'$
Output: Train meta learner $M'$ for accurate classification of
    IoMT attacks
    BEGIN
Train base learner by applying $L_i$ to dataset $T$
for $i = 1, \ldots k$, do
    $B_i = \mathfrak{L}_1(T)$
end for
construct training set $T$ for meta-learner
for $j$ = 1, ... $n$, do
    for $i$ = 1, ... $k$, do
        % use $B_i$ to classify training example $x_j$
        $z_{ij} = B_i(x_j)$
    end for
    $T' = \{Z_j, c_j\}$, where $Z_j = \{z_{1j}, z_{2j}, \ldots z_{nj}\}$
end for
Train a meta level classifier $M'$
    $M'$ = $L'$ $(T')$
return $M'$
```

Hyper-parameter optimization (HPO) for weak learners has been performed using the Bat Algorithm, a nature-inspired meta-heuristic algorithm inspired by bat echolocation behavior, where bats exhibit varying loudness, frequencies, and pulse emission rates during flight. The Bat Algorithm initializes a population of bats in an n-dimensional search space. The position of bat i is denoted by $x_i(t)$, and its velocity at time t is $v_i(t)$. The new position is defined as $x_i(t+1)$, and the new velocity is $v_i(t+1)$ at the time stamp $t+1$, determined as follows:

$$x_i(t+1) = x_i(t) + x_i(t+1) \tag{3}$$

$$v_i(t+1) = v_i(t) + (x_i(t) - p(t)).f_i \tag{4}$$

$$f_1 = f_{\text{min}} + (f_{\text{max}} - f_{\text{min}}) \cdot \beta \tag{5}$$

depicted by a random vector with uniform distributions in the range [0, 1], while $p(t)$ represents the current global-optimal solution, where $f_{\text{min}} = 0$ and $f_{\text{max}} = 1$.

The Bat Algorithm demonstrates adaptive local and global search strategies as follow:

$$x_i(t+1) = \quad \overrightarrow{\mathrm{P}}(t) + \epsilon \overline{A}(t) \tag{6}$$

where ϵ represents a random number in the range [-1, 1], and $\overline{A}(t)$ denotes the loudness of the population.After Bat Algorithm optimization, weak-learner models are fine-tuned with the obtained optimal parameters.

The main rationales for selecting the base learning algorithms are:

1) Ensemble models (RF, AdaBoost) excel with large, complex datasets, while MNB, MLP, and NN perform well with substantial data.
2) They enable parallel execution, reducing training time and enhancing efficiency.
3) During training, they compute critical features, aiding feature engineering.
4) The Meta-Learning technique introduces randomness, enhancing the model's robustness and generalizability.

After base learners provide results, we ensemble and apply XGBoost to enhance performance, mitigating individual base learners' mistakes. XGBoost serves as a powerful meta-learner in our proposed system. For meta-learner HPO, we use Bayesian Optimization with Tree-structured Parzen Estimator (BO-TPE), maximizing the Expected Improvement (EI) acquisition function.

$$EI(x) = \begin{cases} (f(x) - f(x_{\text{best}}) - \xi)^+ & \text{if } \sigma(x) > 0, \\ 0 & \text{if } \sigma(x) = 0, \end{cases} \tag{7}$$

where, x is the input point to calculate expected improvement

$f(x)$ is the estimated objective function value at x

x_{best} is the current best point

ξ is a tunable exploration-exploitation parameter

$sigma(x)$ is the estimated standard deviation of the objective function at x

$(\cdot)^+$ denotes, positive function, i.e. $(a)^+ = max(0, a)$

The formula is most effective at capturing improvements over the best current values, adjusted by uncertainty and truncated to zero when there are no improvements. BO-TPE creates two functions, $h(i)$ and $o(i)$, known as generative models, which differentiate between poor and good results based on a threshold j^*. The TPE model is represented as follows:

$$P(i \mid j, D) = \begin{cases} h(i), & \text{if } j < j^* \\ o(i), & \text{if } j > j^* \end{cases} \tag{8}$$

Where $h(i)$ and $o(i)$ represent the likelihood that the subsequent hyper-parameter will be found in regions with poor and high performance. BO-TPE detects the optimal hyper-parameters by maximizing the $h(i)/o(i)$ ratio. TPE is organized as a tree structure that efficiently optimizes hyper-parameters of ML models.

Algorithm 4 Anomaly Based IDS

```
Input: Training dataset $D = \{(x_1, c_1), (x_2, c_2), \dots (x_n, c_n)\}$
    Mean-shift clustering $MS - CL$
    Random-Forest as $B_1$ and $B_2$
Output: Train $B_1$ and $B_2$ for accurate classification of normal
    and unknown IoMT attacks
    BEGIN
Split $D$ using mean-shift clustering $MS - CL$
Label each Cluster $CL$ as Normal-Cluster $C_n$ or
    Unknown-Attack-Cluster $C_{un}$
for each sample $i$ in $CL$ do
    Calculate clustering probability $P_i$ for each $CL$
end for
Optimize no's of $CL$
Collect False-Negative $FN$ and False-Positive $FP$ from
    $MS - CL$
Construct two biased classifiers based on $B_1$ and $B_2$
for each $FN$ do
    Train $B_1$
    if $FN \leftarrow reduce$ then
        Data = normal
    end if
end for
for each $FP$ do
    Train $B_2$
    if $FP \leftarrow reduce$ then
        Data = unknown attack
    end if
end for
return
```

2) Anomaly-based IDS: The signature-based IDS excels in detecting known attacks, but it falls short against zero-day attacks. To address this, we introduced an anomaly-based IDS optimized by RF-RFE and LDA in our system. The mean-shift clustering (MS-CL) technique is then applied to distinguish between average and attack data.

The MS-CL method separates the data into clusters, each assigned a class label, ”attack” or ”normal.” Test set instances are categorized based on the cluster label, determining if they represent attack or normal data. The ”clustering probability” p_i of the class with the highest probability in each cluster is computed for each sample in the test set i.

The primary objectives of MS-CL are as follows:

1) Divide the data samples into a suitable number of clusters.
2) Assign the class label "attack" or "normal" to each cluster based on the majority of instances.
3) Calculate the clustering probability (p_i) for each test instance (i) in a cluster by determining the percentage of instances in that cluster belonging to the majority class.

MS-CL effectively differentiates between average and attack data, offering superior performance to k-means clustering. It excels in computational efficiency with a time complexity of ($O(nkt)$), where n is the data size, k is the number of clusters, and t is the number of iterations. The introduction of mini-batches has significantly reduced training time per cycle, ensuring rapid convergence and adaptability to new sample sets.

To improve MS-CL's detection rate and reduce false accuracy, two biased classifiers were added, leading to a decrease in false positives (FPs) and false negatives (FNs).

The biased classifiers serve the following purposes:

1) Retrieve the FPs and FNs from the training set of the MS-CL model.
2) Utilize Random Forest, a supervised learning model, to construct the biased classifiers.
3) Train the initial biased classifier $B1$ by employing all FNs along with an equal number of randomly selected instances of normal data to mitigate the FN rate.
4) Train another biased classifier $B2$ with an equivalent number of randomly selected instances of attack data and all FPs to decrease the FP rate.

After implementing the MS-CL model, instances with clustering probability pi lower than the threshold $\hat{p}$ are considered ambiguous. The continuous variable $\hat{p}$ has been optimized by BO-TPE, and its value can range up to 0.90. Once the two biased classifiers are trained, ambiguous samples are passed to $B1$ (if MS-CL classifies them as normal) or $B2$ (if MS-CL classifies them as attacks) to obtain the final outcome. The biased classifier in construction utilized only the false positives (FPs) and false negatives (FNs) from the initial training phase. In our anomaly IDS, the application of Mean-Shift clustering with biased classifiers enhances its proficiency compared to other methods like One-Class SVM (OC-SVM) and Isolation-Forest (i-Forest) [55], enabling effective detection of new attack methods and efficient handling of unlabeled data.

Our MS-CL method with biased classifiers provides distinct advantages.

1) The MS-CL model, unlike OC-SVM and i-Forest, can accurately simulate data samples with normal and attack patterns, offering superior generalizability and data-pattern modeling capabilities.
2) MS-CL adapts the number of clusters automatically, according to the complexity of the data pattern.
3) Biased classifiers mitigate FPs and FNs, enhancing the detection of challenging patterns in misclassified data samples by MS-CL.
4) The clustering probability (p_i streamlines model efficiency by directing uncertain samples (with low probability) to biased classifiers for further processing, while new instances resembling normal or attack patterns (with higher confidence) are directly labeled.
5) Mini-batch MS-CL is used to shorten Meta-IDS execution times to satisfy IoMT's real-time requirements.

E. Runtime Complexity

The Meta-IDS was trained on a high-speed GPU to achieve rapid processing, ensuring applicability in real healthcare settings. Its streamlined runtime complexity enables real-time performance, minimizing IoMT network latency. During implementation, the Meta-IDS undergoes evaluation with six ML models, a MS-CL model, and a biased classifier. The runtime complexity of RF, AdaBoost, and XGBoost is $O(dtf)$, where d denotes the maximum tree depth, f denotes the features, t denotes the number of trees, while the runtime complexity of DT is $O(df)$. The NN and MLP have $O(nt*(ij+jk+kl))$ runtime complexity, where n denotes the number of epochs, t denotes training, i, j, k and l denote the nodes. The runtime complexity of naive Bayes is $O(n*f*c)$, where n shows the number of data points, f shows the features, and c shows the number of classes.

The proposed MS-CL in anomaly-IDS has a time complexity of $O(fk)$ where k shows the number of clusters. A tree-based model with an $O(dft)$ time complexity serves as the biased classifier in anomaly-IDS. The suggested Meta-IDS has an overall runtime complexity of $O(2dft+fk+(nt*(ij+jk+kl)+(n*f*c))$ at the low level.

F. Real-World Applicability of Proposed Meta-IDS

Our Meta-IDS revolutionizes IoMT security by dynamically adapting to evolving intrusion patterns through innovative meta-learning. Its hierarchical framework specializes base learners for distinct intrusion scenarios, ensuring effective generalization across various attacks. With superior accuracy, efficiency, and adaptive interpretability, it excels in real-world IoMT security, serving as a versatile IDS gateway for network traffic monitoring and threat detection. Additionally, its adaptability to cloud and fog nodes makes it ideal for infrastructure as a service (IaaS) and software as a service (SaaS) deployment, offering robust intrusion detection capabilities in dynamic healthcare environments.

In the healthcare domain, different interfaces or diagnostic tools have been used as data sharing or electronic control units (ECUs) like DICOM, HL7, IEEE 11073, IHE, etc. Each has its pros and cons, but we have used HL7, which is a standard protocol used for the sharing, integration, exchange, and retrieval of electronic health records (EHRs) and is commonly used for sharing healthcare data.

Figure 3 shows the proposed Meta-IDS protected healthcare architecture. Different IoMT devices (IoMT Cluster) gather the data and pass data to the HL7 interface module, which is responsible for handling HL7 communication and managing data gathering. Then the data is transferred to the network by the HL7 interface module. The network TAP on the network infrastructure intercepts and captures HL7 traffic and connects the network TAP to the network. The attacker can breach the HL7 protocol (if it is not secured enough) and inject malicious

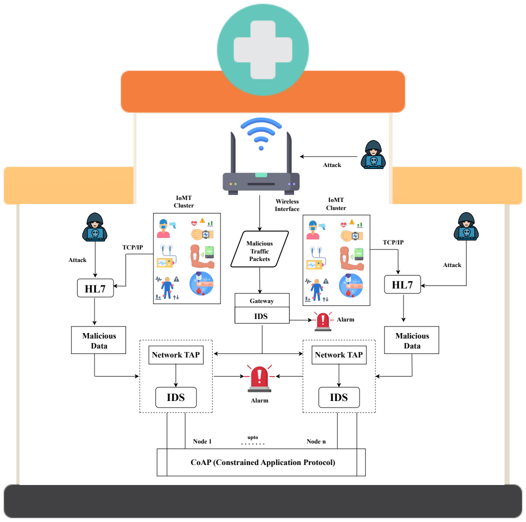


Fig. 3: Real-time architectural diagram

data. When data is received by the network TAP, it will be processed by our proposed Meta-IDS. If any attack is found, it triggers the alarm. The same is the case for the wireless interface – if the attacker injects malicious data, it will be filtered out at the gateway.

G. Comparison of the proposed Meta-IDS with Explainable AI-based E-GraphSAGE Model

E-GraphSAGE, a renowned explainable AI algorithm introduced by Weng et al. [38], utilize edge features to aggregate graph information through the input, aggregator function, and message passing mechanisms. We applied the E-GraphSAGE algorithm to the WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT-2021 datasets, utilizing only the provided edge features (EF) $e_{uv}, \forall_{uv} \in \mathcal{E}$. Since the datasets lack node features (NF), we initialized NF using the constant vector $X_v = 1, \ldots, 1$ as per Algorithm 5, maintaining the vector dimensions consistent with the number of EF.

$$h^k_{\mathcal{N}(\sqsubseteq)} = AGG_k(\{e^{k-1}_{u,v}, \forall u \in \mathcal{N}(v)\}) \tag{9}$$

The embedding of every node u in the vicinity of node v was combined to create embedding-node v at layer k, where $h^N_k(v)$ displays node embedding (u) in the preceding layer.

Eq. 9 shows the standard GraphSAGE model, but the E-GraphSAGE model used the aggregated embedding of the sample at adjacent edges of the k-th layer, as explained in the following equation.

$$h^k_{\mathcal{N}(\sqsubseteq)} = AGG_k(\{e^{k-1}_{u,v}, \forall u \in \mathcal{N}(v), uv \in \mathcal{E}\}) \tag{10}$$

where the edge-features uv from the $\mathcal{N}(v)$, sample adjacent-nodes v at layer $k-1$ are displayed by e^{k-1}_{uv}. Additionally, the edge-sample in the adjacent of $\mathcal{N}(v)$ is shown by $\{\forall u \in \mathcal{N}(v), uv \in \mathcal{E}\}$. the embedding-node v at k layer calculated by the original GraphSAGE algorithm is shown as follows.

$$h^k_v = \sigma(W^k.CONCAT(h^{k-1}_v, h^k_{\mathcal{N}(v)})) \tag{11}$$

The crucial distinction lies in the fact that the E-GraphSAGE algorithm computes $h^k_{\mathcal{N}(v)})$ using equation 11, which incorporates EF. Each graph node's k-hop neighbourhood provides the edge and topological information that is gathered and aggregated in the network-flow graph [38]. By concatenating the embedding of nodes u and v, it is possible to calculate the edge-embedding z^K_{uv} of edges uv, as illustrated below.

$$z^K_{uv} = CONCAT(z^K_u, z^K_v)uv \in \mathcal{E} \tag{12}$$

The E-GraphSAGE model comprises three primary steps: constructing a network graph using network-flow data, training a supervised model with the generated data, and creating edge embeddings to enable edge classification, discerning between normal and attack classes.

1) The construction of the network graph:: The network-flow data, comprising source and destination details alongside additional fields like data bytes and packet counts, naturally translates into a graph format. We define graph edges using four flow fields: Sport, Dport, SrcAddr, and DstAddr are used to identify nodes in the graph. DstAddr and Dport denote the destination node, while SrcAddr and Sport indicate the source node, exemplified by data exchange between 10.0.1.172:58059 and 10.0.1.150:1111. For network graph creation, we randomly selected source IP addresses ranging from 172.15.0.1 to 172.33.0.1 to obfuscate potential attack vectors. During graph construction, all other flow fields were assigned to the edge, resulting in featureless graph nodes. Each node received a vector containing all one values, as outlined in the algorithm.

2) E-GraphSAGE training:: Implemented a neural network comprising two layers of E-GraphSAGE (k=2), capturing data aggregated from a 2-hop neighborhood. The mean of edge features (EF) within the neighborhood sample is computed using the mean function, as defined below:

$$h^k_{\mathcal{N}(v)} = \sum_{u \in \mathcal{N}(v), uv \in \mathcal{E}} \frac{e^{k-1}_{uv}}{|\mathcal{N}(v)|_e} \tag{13}$$

Where $|\mathcal{N}(v)|_e$ shows the number of edges, e^{k-1}_{uv} shows edge features at layer $k-1$.

Utilized two E-GraphSAGE layers with 128 hidden units, ReLU activation, 20% dropout regularization, Adam optimizer, cross-entropy loss, and a learning rate of 0.001 for backpropagation. In the final layer, node embeddings were converted to edge embeddings by concatenating two node embeddings. These edge embeddings were then passed through a softmax layer during backpropagation to fine-tune trainable parameters by comparing them to dataset labels.

3) Edge classification:: After training and parameter tuning, the model categorizes unseen samples during evaluation. Test-flow records are transformed into graphs and passed through the trained E-GraphSAGE model. Edge embeddings are then generated, and class probabilities are computed using softmax and compared to actual class labels.

IV. Dataset Description

The choice of an appropriate dataset for model evaluation is a crucial step that demands careful consideration. To meet the objectives of conducting experimental research on the

Algorithm 5 E-GraphSAGE edge embedding

```
Input: Graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ :
    input edge features $\{e_{uv}, \forall_{uv}, \in \mathcal{E}\}$ :
    input node features $x_u = \{1, \ldots, 1\}$ :
    depth $K$ :
    weight matrices $W^k, \forall k \in \{1, \ldots, K\}$ :
    non-linearity $\sigma$ :
    differentiable aggregator function $AGG_k$ :
Output: Edge embeddings $z_{u,v}, \forall_{uv}, \in \mathcal{E}$
    $h_v^o \leftarrow x_v, \forall v \in \mathcal{V}$
for $k \leftarrow 1 To K$ do
    for $v \in \mathcal{V}$ do
        $h^k_{\mathcal{N}(\sqsubseteq)} \leftarrow AGG_k(\{e^{k-1}_{u,v}, \forall u \in \mathcal{N}(v), uv \in \mathcal{E}\})$
        $h^k_v \leftarrow \sigma(W^k.CONCAT(h^{k-1}_v, h^k_{\mathcal{N}(v)}))$
    end for
end for
$z_v = h^k_v$
for $uv \in \mathcal{E}$ do
    $z^K_{uv} \leftarrow CONCAT(z^K_u, z^K_v)$
end for
```

TABLE IV: Description of WUSTL-EHMS-2020 dataset

Class Label	Original Sample	Training Samples	Test Samples
Normal Data	14272	9990	4281
Attack Data	2046	1432	613
Total Samples	16318	11422	4899

TABLE V: Description of IoTID20 dataset

Label	Description	Samples Counts
Normal Data	No suspicious activity	40,073
DoS	Denial of service attack	59,391
Mirai	Mirai botnet attack	415,677
MITM	Man in the middle attack	35,377
Scan	Scan attack	75,265
Total		625,783

proposed model, three datasets were utilized: WUSTL-EHMS-2020 [25], IoTID20 [53], and WUSTL-IIOT-2021 [62].

The datasets were selected to comprehensively evaluate the proposed IDS across diverse scenarios. The WUSTL-EHMS-2020 dataset assesses the IDS in healthcare environments, prioritizing data and system security. The IoTID20 dataset evaluates the IDS in IoT networks, capturing unique challenges and attack patterns. Lastly, the WUSTL-IIOT-2021 dataset extends evaluation to industrial IoT settings, addressing cyber threats in critical infrastructure. These datasets ensure the IDS's adaptability and effectiveness across varied domains and security requirements.

A. WUSTL-EHMS-2020 dataset

The dataset, obtained from a live health monitoring testbed [25], integrates patient-worn sensors, a network gateway, and a software-defined network controller for monitoring data transmissions. It encompasses sensor and network traffic data, analyzed to detect intrusion sources and anomalous patterns, including three types of attacks: data-injection, spoofing, and man-in-the-middle. Comprising 44 features—35 related to the network and 8 to patient biometrics, along with class labels—this dataset's statistical details are outlined in Table IV.

TABLE VI: Description of WUSTL-IIOT-2021 dataset

Samples or Traffic Type	No. of samples or %
No. of samples	1,194,464
No. of features	41
No. of attack samples	87,016
No. of normal samples	1,107,448
Normal traffic	92.72 %
Total attack-traffic	7.28 %
command-injection traffic	0.31 %
DoS traffic	89.89 %
Reconnaissance Traffic	9.46 %
Backdoor Traffic	0.25 %

B. IoTID20 dataset

The IoTID20 dataset, sourced from IoT devices, comprises 80 features across 625,783 samples, with 585,710 labeled as malicious. It encompasses various cyberattacks, including DoS (Denial of Service), Mirai (a botnet malware), MITM (Man-in-the-Middle), and Scan (port scanning). DoS floods systems with unauthorized requests to disrupt normal operations, while Mirai transforms IoT devices into a botnet for extensive DDoS attacks. MITM clandestinely intercepts and potentially alters communication, while Scan methodically searches for vulnerabilities. See Table V for an overview of the dataset.

C. WUSTL-IIOT-2021 dataset

The WUSTL-IIOT-2021 dataset, designed to mimic real-world industrial systems, contains 41 features across 1,194,464 samples, with 1,107,448 labeled as normal and 87,016 as attack. It includes various attacks like command injection, DoS, reconnaissance, and backdoor incidents. Command injection manipulates system behavior, while DoS overwhelms with requests. Reconnaissance gathers system info, and backdoor incidents provide unauthorized access. See Table VI for dataset summary.

V. Experiment and Performance Evaluation

In this section, we delineate the experimental methodology, present the setup, and analyze performance results. We further compare these results with benchmarks to assess the effectiveness of the proposed approach.

A. Experimental Setup

In our proposed method, we implemented features engineering and machine-learning algorithms using Python with the well-known Pandas library, Scikit-learn, and XgBoost. Hyperparameter optimization was conducted using Skopt, hyperOpt, and Sklearn Nature-Inspired Algorithm [1]. For our experiments, we utilized the Lenovo Yoga Slim 14ITL05, featuring an 11th generation Intel(R) Core(TM) i5-1135G7 processor with clock speeds of 2.40 GHz and 2.42 GHz, along with 16 gigabytes (GB) of RAM.

[1] The code is available at: https://github.com/UmerZu/Code-files

B. Evaluation metrics

The suggested model is evaluated using a variety of measures, including Accuracy (Acc), precision (P), recall (R), and F1-score, that have been calculated by the true-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) rates. The evaluation metrics have been computed using the following equations.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{14}$$

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \tag{17}$$

C. Evaluation of Known Intrusions' Performance:

To ensure generalizability and mitigate overfitting risks in our experiments, we utilized hold-out and cross-validation methods. The procedures for model evaluation and dataset split are outlined below:

1) Seventy percent of the data was allocated for training, while the remaining thirty percent was set aside for testing. The test data remained unaltered throughout the hold-out validation process.
2) We assessed the model's efficacy through 10-fold cross-validation (CV) on distinct subsets of the training dataset. In each fold, 90% of the data was utilized for training, and the remaining 10% for validation.
3) The model from Step 2 underwent evaluation on a separate, untouched dataset.

Employing a 70-30% test-train split and 10-fold cross-validation addresses concept-drift and over-fitting [34]. It ensures consistent intrusion detection, preventing over-fitting or under-fitting. To identify unknown attacks, a hold-out validation is used, with all data samples except those corresponding to unknown attacks as training sets. The system's effectiveness in identifying new attack patterns is measured using a validation set created for unknown attacks, comparing their statistical similarity to previously identified attack data.

The proposed Meta-IDS is assessed for signature-based intrusion detection using WUSTL-EHMS 2020, IoTID20, and WUSTL-IIOT-2021 datasets. Results, optimized through hold-out and 10-fold cross-validation, are summarized in Table VII.

On the WUSTL-EHMS 2020 dataset, our proposed Meta-IDS was compared with [13], [48], [24], [33], [41], [11], and [3]. The Meta-IDS exhibited outstanding performance with 99.57% accuracy, 99.56% precision, 99.57% recall, and a 99.56% F1-score. Additionally, it achieved the shortest execution time of 77.2 seconds. On the WUSTL-IIOT-2021 dataset, the proposed Meta-IDS achieved an impressive accuracy of 99.99%. Notably excelling in distinguishing normal and attack data patterns, our Meta-IDS outperforms other techniques [19], [40], [2], [13]. Additionally, it demonstrates minimal execution time (307.3 seconds). On the IoTID20 dataset, our Meta-IDS showcased remarkable accuracy at 99.93%, surpassing alternative techniques [16], [5], [43], [30], [4]. Furthermore, it exhibited minimal execution time (247.3 seconds). Leveraging ConvNeXt-wavelet for feature enrichment, our Meta-IDS excels in accuracy, efficiency, and overall model robustness, surpassing ConvNeXt-Wavelet's limitations, particularly in balancing accuracy with timely responsiveness. This underscores the groundbreaking strides of our proposed methodology, solidifying its status as an exceptionally effective and efficient solution for deployment in healthcare systems. For a detailed breakdown of the results, consult Table VII.

In crucial multivariate benchmarking, our Meta-IDS outperforms other intrusion detection models, excelling in known and zero-day attack detection, dynamic adaptability, and minimizing execution time. DivaCAN [31] is effective in CAN bus intrusion detection, but, lacks zero-day attack coverage and has an extended execution time of 406 seconds. Similarly, 1C-KNN [28] and the blended IDS [8] focus on specific domains, lacking real-world applicability details. Meta-IDS emerges as the superior choice for comprehensive intrusion detection in diverse environments.

Figure 4a presents the confusion matrix for E-GraphSAGE on WUSTL-EHMS-2020, achieving 85.66% accuracy with a 0.1433% misclassification rate. In contrast, Figure 4b illustrates our Meta-IDS outperforming, attaining 99.57% accuracy and an impressively low 0.0042% misclassification rate. For the IoTID20 dataset (Figure 4c), E-GraphSAGE achieved 97.51% accuracy with a 0.0248% misclassification rate. However, Figure 4d showcases our Meta-IDS excelling, reaching 99.93% accuracy and an exceptionally low 0.0006% misclassification rate. Turning to WUSTL-IIOT 2021 (Figure 4e), E-GraphSAGE achieved 95.42% accuracy with a 0.0457% misclassification rate. Conversely, Figure 4f illustrates our Meta-IDS dominance, securing 99.99% accuracy and an exceptionally low 0.00004% misclassification rate.

Table VIII offers a thorough multi-class assessment of Meta-IDS on the WUSTL-IIOT 2021 and IOTIO20 datasets, highlighting its precision in predicting normal data and different attack classes. Further insights into the multi-class performance are provided through detailed confusion matrices in Figure 5a and Figure 5b.

In Figure 6a, the AUC on WUSTL-EHMS-2020 for the Meta-IDS and Weak-Learners are as follows: NN (70.71%), MLP (50%), Naive Bayes (53%), RF (99.2%), DT (94.9%), AdaBoost (93.3%), Meta-Learner (99.66%). Figure 6b shows the AUC on IoTID20 for the Meta-IDS and Weak-Learners: NN (85.15%), MLP (50.01%), Naive Bayes (87.72%), RF (99.92%), DT (99.59%), AdaBoost (99.92%), Meta-Learner (99.98%). In Figure 6c, AUC on WUSTL-IIOT 2021 for the Meta-IDS and Weak-Learners are: NN (73.16%), Naive Bayes (98.81%), RF (98.59%), DT (97.17%), AdaBoost (98.81%), Meta-Learner (100%).

D. Performance analysis of anomaly-based IDS

The Anomaly-based IDS is trained on instances labeled as binary categories (attack and normal). Following the evaluation by the Signature-based IDS, all known-attack samples

TABLE VII: Signature-based evaluation of the proposed Meta-IDS on three datasets

Dataset	Model	Accuracy	Precision	Recall	F1-Score	Ex-Time (S)
WUSTL-EHMS 2020	CNN-Focal [13]	93.08	94.23	73.38	79.63	125.6
	CNN [48]	95.00	94.00	85.00	88.00	145.2
	FNN-Focal [13]	93.26	95.24	73.69	80.11	194.4
	DNN-FL [41]	91.40	65.05	61.42	61.05	-
	MLP [3]	97.57	97.60	97.50	97.60	-
	Tree Classifiers [24]	93.00	93.00	93.00	93.00	-
	RFE-MLP [33]	96.20	96.23	96.20	96.20	-
	PSO-DNN [11]	96.00	96.00	96.00	96.00	-
	ConvNeXt-Wavelet transformation	77.89	76.45	76.42	77.40	420.0
	EGraphSAGE	85.66	85.67	85.66	85.64	115.8
	Meta-IDS	**99.57**	**99.56**	**99.57**	**99.56**	**77.2**
WUSTL-IIOT 2021	BA-RF [19]	99.90	99.60	93.60	99.60	911.6
	IFPCC-RF [40]	99.12	89.59	99.50	94.29	566.19
	FNN-Focal [13]	98.95	77.22	64.06	68.48	756.5
	CNN-GRU-10 [2]	97.74	97.74	97.69	97.74	402.4
	CNN-Focal [13]	98.21	88.54	66.51	70.50	589.4
	ConvNeXt-Wavelet transformation	88.32	88.21	86.76	88.41	1480
	EGraphSAGE	95.42	95.43	95.43	95.41	947.2
	Meta-IDS	**99.99**	**99.99**	**99.99**	**99.99**	**307.3**
IoTID20	LSTM [16]	99.00	99.00	99.00	99.00	1851.6
	CNN-LSTM [5]	98.00	98.40	77.40	98.80	1876.9
	Multistaged DT-SAINT [43]	94.41	92.31	94.40	92.30	1756.5
	Hybrid DL Model [30]	99.70	99.80	99.60	99.70	-
	Ensemble UMF [4]	99.70	99.60	99.70	99.50	919.91
	ConvNeXt-Wavelet transformation	76.96	74.04	75.54	76.23	940.0
	EGraphSAGE	97.51	97.52	97.51	97.51	826.4
	Meta-IDS	**99.91**	**99.93**	**99.91**	**99.91**	**247.3**

TABLE VIII: Multi-class evaluation of Meta-IDS on WUSTL-IIOT 2021 and IOTIO20 dataset

Dataset	Classes	Accuracy	Precision	Recall	F1-Score
WUSTL-IIOT 2021	Normal	99.99%	99.98%	99.99%	99.99%
	DoS	99.99%	100%	100%	99.99%
	Reconn	99.99%	100%	100%	100%
	CommInj	99.99%	100%	99.99%	100%
	Backdoor	100%	100%	100%	99.99%
IOTIO20	Mirai	99.99%	100%	100%	100%
	Scan	99.99%	99.99%	100%	99.99%
	DoS	99.97%	99.96%	99.97%	99.96%
	Normal	99.99%	100%	99.99%	99.99%
	Spoofing	99.97%	99.97%	99.93%	99.95%

are identified, while the remaining samples are labeled as ”suspicious.” Subsequently, this data is input to the Anomaly IDS to detect the presence of any unknown attacks.

Table IX presents the results of the Anomaly-based IDS, utilizing Mean-Shift Clustering and biased classifiers across WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT 2021 datasets. Mean-Shift Clustering achieved accuracy scores of 62.86%, 81.31%, and 94.68%, revealing limitations in handling complex data. With the integration of biased classifiers, the proposed IDS demonstrated substantial improvements, achieving 99.50%, 99.98%, and 99.99% accuracy on the respective datasets, showcasing enhanced precision, recall, and F1-Score performance.

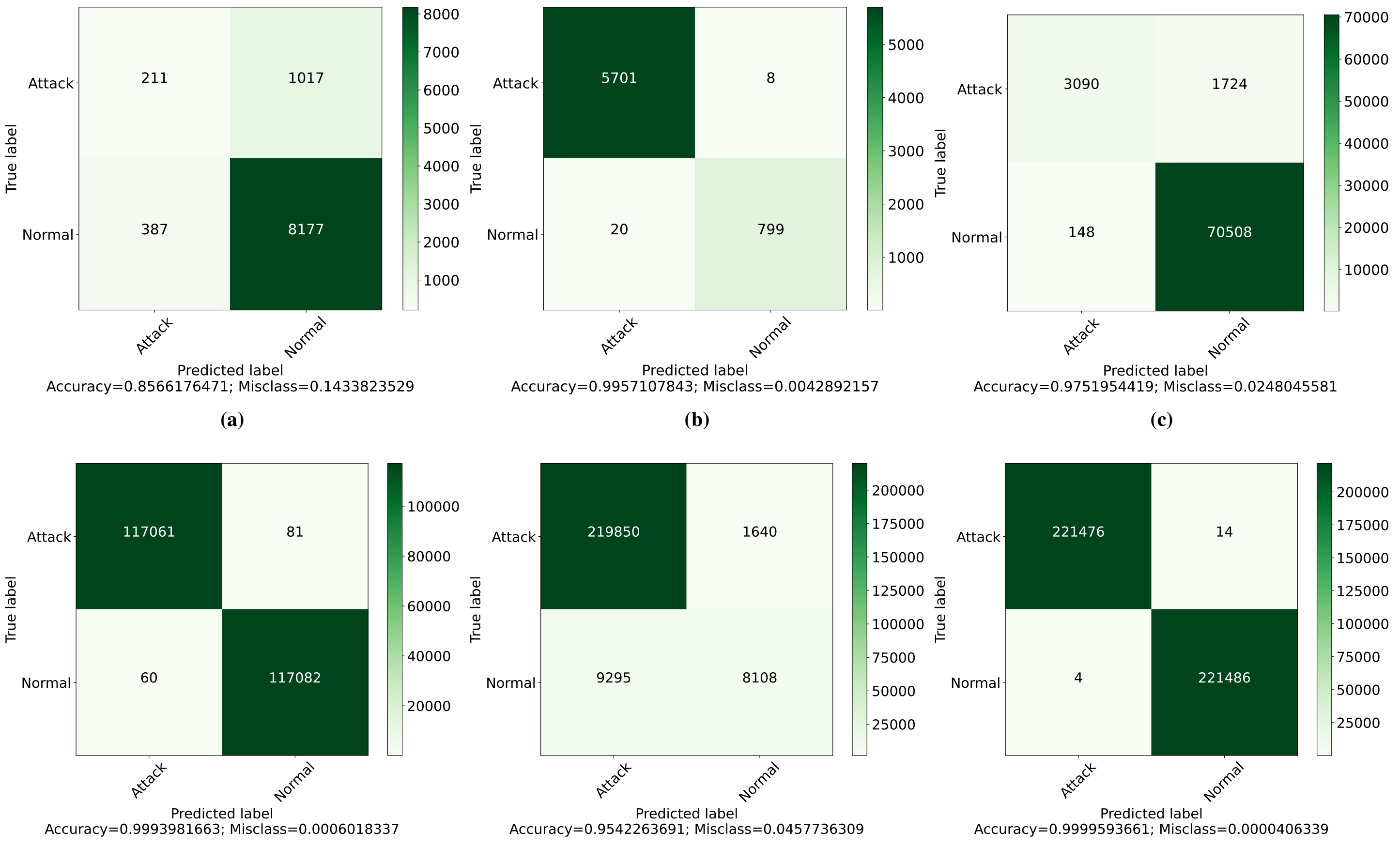
Figures 7a, 7b, and 7c display the confusion metrics of the Anomaly-based IDS. Achieving 99.99% accuracy with a 0.00002% misclassification rate on WUSTL-IIOT 2021, 99.50% accuracy with a 0.0049% misclassification rate on WUSTL-EHMS-2020, and 99.98% accuracy with a 0.0001% misclassification rate on IoTID20.

E. Ablation study:

The ablation study systematically disables base learners and the meta-learner, providing insights into their impact on the IDS. Results highlight the meta-learner's indispensability, guiding refinement for a more robust and accurate IDS. In EHMS-2020, the full model achieves 99.57% accuracy; omitting base learners minimally affects accuracy, with the meta-learner's absence resulting in 98.83%. For IoTID20, the complete model reaches 99.91% accuracy, with negligible changes upon deactivating base learners. The meta-learner's omission results in 98.64% accuracy. In WUSTL-IIOT 2021, the full model achieves 99.99% accuracy, with slight fluctuations when disabling base learners. The absence of the meta-learner leads to 97.24% accuracy. The detailed discussion of results can be found in Table X.

VI. DISCUSSION

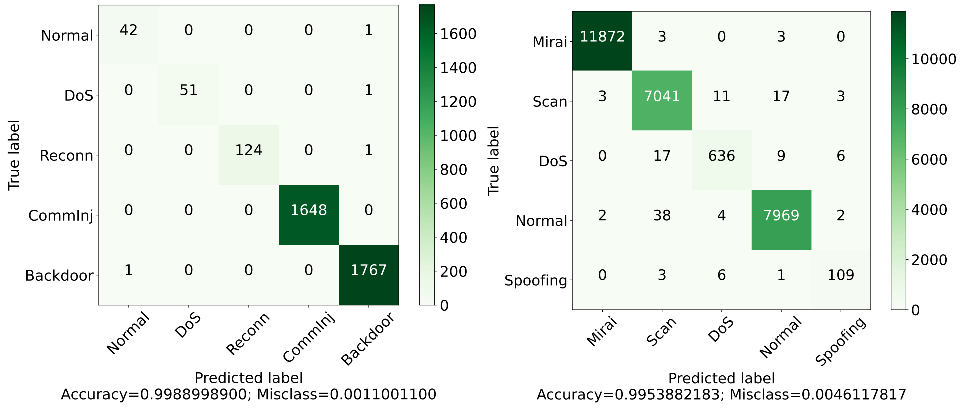
Our proposed Meta-IDS comprises two subsystems tailored for effective known and unknown attack detection in IoMT networks. For known attack detection, we employ the



(a) **(b)** **(c)**

(d) **(e)** **(f)**

Fig. 4: Confusion metrics for (a) E-GraphSAGE on WUSTL-EHMS-2020 dataset (b) Meta-IDS on WUSTL-EHMS-2020 dataset (c) E-GraphSAGE on IoTID20 dataset (d) Meta-IDS on IoTID20 dataset (e) E-GraphSAGE on WUSTL-IIOT 2021 dataset (f) Meta-IDS on WUSTL-IIOT 2021 dataset

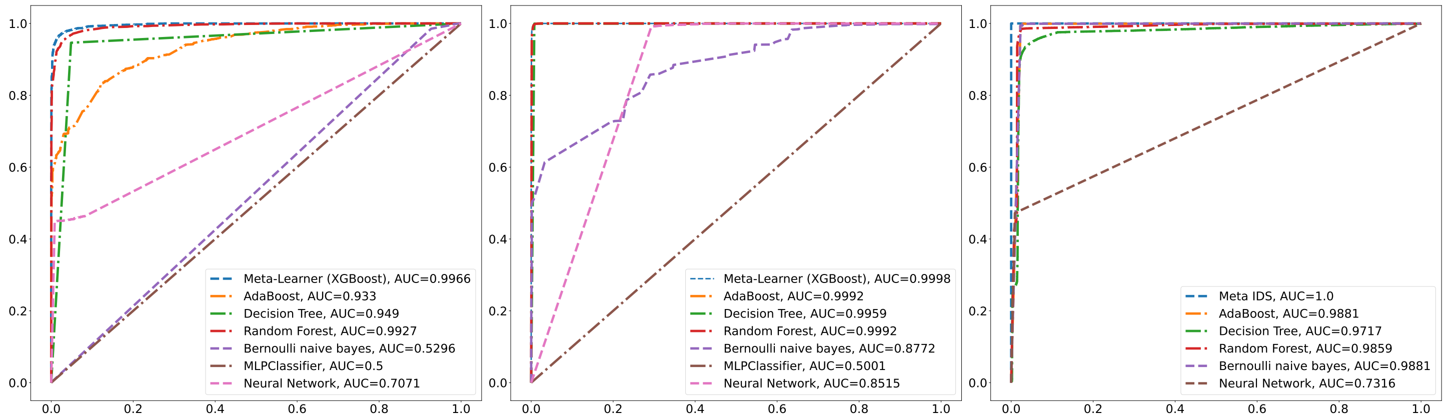


(a) **(b)**

Fig. 5: Confusion matrices for (a) Multi-class CM of Meta-IDS on WUSTL-IIOT 2021 and (b) Multi-class CM of Meta-IDS on IoTIO20

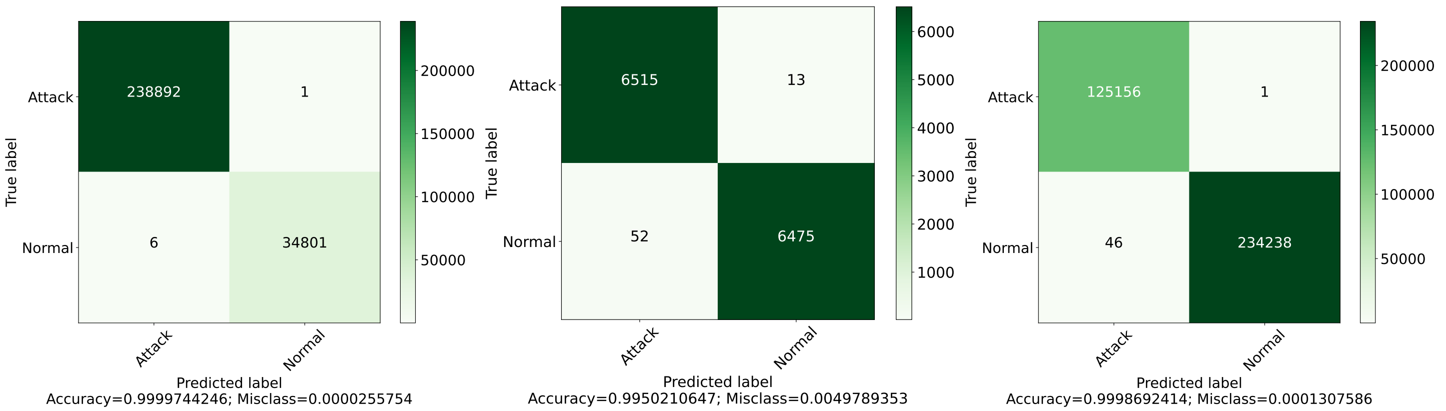
TABLE IX: Evaluation of Anomaly IDS on Different Datasets

Datasets	**Mean-Shift Clustering**				**Biased Classifier**			
	Accuracy	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**Precision**	**Recall**	**F1-Score**
WUSTL-EHMS-2020	62.86%	64.54%	62.86%	61.75%	99.50%	99.50%	99.50%	99.50%
IoTID20	81.31%	81.67%	81.31%	81.44%	99.98%	99.98%	99.98%	99.98%
WUSTL-IIOT 2021	94.68%	94.80%	94.68%	94.73%	99.99%	99.99%	99.99%	99.99%



(a) (b) (c)

Fig. 6: Area under the curve (AUC) on the (a) WUSTL-EHMS-2020, (b) IoTID20, and (c) WUSTL-IIOT 2021 dataset



(a) (b) (c)

Fig. 7: Confusion matrix of Anomaly-based IDS on (a) WUSTL-IIOT 2021 dataset (b) WUSTL-EHMS-2020 dataset (c) IoTID20 dataset

TABLE X: Ablation Study on Different Datasets and Outcomes

Dataset	Configuration	DT	RF	Adaboost	Naive Bayes	MLP	NN	Meta-Learner	Accuracy
EHMS-2020	**Full Model**	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	**99.57 %**
	DT	**Disabled**	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	99.32 %
	RF	Enabled	**Disabled**	Enabled	Enabled	Enabled	Enabled	XGBoost	99.25 %
	Adaboost	Enabled	Enabled	**Disabled**	Enabled	Enabled	Enabled	XGBoost	99.18 %
	Naive Bayes	Enabled	Enabled	Enabled	**Disabled**	Enabled	Enabled	XGBoost	99.23 %
	MLP	Enabled	Enabled	Enabled	Enabled	**Disabled**	Enabled	XGBoost	99.11 %
	NN	Enabled	Enabled	Enabled	Enabled	Enabled	**Disabled**	XGBoost	99.36 %
	No Meta-Learner	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	**No Meta-L**	98.83 %
IoTID20 dataset	**Full Model**	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	**99.91 %**
	DT	**Disabled**	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	99.86 %
	RF	Enabled	**Disabled**	Enabled	Enabled	Enabled	Enabled	XGBoost	99.78 %
	Adaboost	Enabled	Enabled	**Disabled**	Enabled	Enabled	Enabled	XGBoost	99.69 %
	Naive Bayes	Enabled	Enabled	Enabled	**Disabled**	Enabled	Enabled	XGBoost	99.88 %
	MLP	Enabled	Enabled	Enabled	Enabled	**Disabled**	Enabled	XGBoost	99.87 %
	NN	Enabled	Enabled	Enabled	Enabled	Enabled	**Disabled**	XGBoost	99.83 %
	No Meta-Learner	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	**No Meta-L**	98.64 %
WUSTL-IIOT 2021	**Full Model**	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	**99.99 %**
	DT	**Disabled**	Enabled	Enabled	Enabled	Enabled	Enabled	XGBoost	99.92 %
	RF	Enabled	**Disabled**	Enabled	Enabled	Enabled	Enabled	XGBoost	99.89 %
	Adaboost	Enabled	Enabled	**Disabled**	Enabled	Enabled	Enabled	XGBoost	99.44 %
	Naive Bayes	Enabled	Enabled	Enabled	**Disabled**	Enabled	Enabled	XGBoost	99.86 %
	MLP	Enabled	Enabled	Enabled	Enabled	**Disabled**	Enabled	XGBoost	99.85 %
	NN	Enabled	Enabled	Enabled	Enabled	Enabled	**Disabled**	XGBoost	99.86 %
	No Meta-Learner	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	**No Meta-L**	97.24 %

signature-based IDS technique, while zero-day attacks are targeted using the anomaly-based IDS. Meta-IDS outperforms existing models, achieving an overall accuracy, precision, recall, F1-score and it boasts the fastest execution time. Additionally, we implement E-GraphSAGE, an Explainable AI based Graph Neural Network model, and compare its results with our proposed Meta-IDS, demonstrating the superior performance of our model. Detailed results across three datasets are provided in Table VII. Confusion matrices and comprehensive details of True Positives (TPs), True Negatives (TNs), False Positives (FPs), False Negatives (FNs), accuracy, and misclassification rates for the E-GraphSAGE model and the proposed Meta-IDS on three different datasets are presented in Figure 4a, 4b, 4c, 4d, 4e, 4f.

Meta-learners exhibit superior performance in Area Under the Curve (AUC) compared to other weak-learners, as showcased in Figure 6a, Figure 6b, and Figure 6c, depicting AUC values of weak-learners and meta-learners on three datasets. Our model achieves high results with excellent generalizability, avoiding overfitting through the adoption of comprehensive feature-engineering methods that eliminate misleading and irrelevant features.

The proposed anomaly-based IDS addresses zero-day attack detection. In this approach, data processing involves marking known-attack samples, labeling other normal samples as "suspicious," and feeding the data to a mean-shift clustering model. However, the initial results were unsatisfactory, as indicated in Table IX. To enhance results, we introduced a biased classifier (Random Forest - RF) that accurately classifies False Negatives (FNs) and False Positives (FPs), significantly improving the overall performance of the anomaly-based IDS. Table IX presents the improved overall results of the anomaly-based IDS, while Figure 7a, 7b, and 7c illustrate the confusion matrices. Experimental results affirm that our proposed Meta-IDS efficiently and accurately detects both known and unknown attacks in IoMT networks.

Meta-IDS exhibits scalability through its optimized architecture, featuring a stack of weak-learners and a meta-learner. With its robust capabilities in detecting both known and zero-day attacks, along with its adaptability to diverse datasets and proficiency in handling multi-attack scenarios, Meta-IDS stands out as a highly scalable solution, especially in large-scale IoMT networks. Furthermore, its minimal execution time further enhances its scalability, making it well-suited for deployment in complex and expansive network environments.

However, Meta-IDS may encounter limitations and challenges due to its computational demands, potentially limiting deployment in resource-constrained IoMT and IIoT environments. Transitioning from research to IIoT prototypes introduces hurdles related to environmental variations, diverse data sources, and interoperability. Safeguarding Meta-IDS against adversarial attacks is imperative in IIoT's critical infrastructure. Privacy regulations mandate robust measures for intrusion detection in industrial settings, emphasizing stringent adherence to data privacy standards. Addressing these challenges is crucial for Meta-IDS to be adaptable, secure, and compliant within IoMT and IIoT domains.

VII. Conclusion

To enhance IoMT security, this work have proposed a Meta-IDS model for accurately detecting known and zero-day attacks in IoMT networks. The proposed methodology consists of different steps, data-preprocessing and feature-engineering, which help to improve the data quality. Second, applied six ML models as a weak learner and one meta-learner. Third, used HPO to optimize the model's hyper-parameters, and achieve better performance of ML models. Fourth a mean-shift clustering algorithm, an unsupervised algorithm that detects zero-day attacks. And finally, used biased classifiers which not only improved the attack detection rate but also improved the performance of the mean-shift clustering algorithm. Tested on WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT-2021 datasets, Meta-IDS achieves high accuracy with minimal execution time. Also proposed the feasibility of Meta-IDS in a real-time environment of healthcare security.Limitations of proposed model includes sensitivity to network environment variations, dependence on dataset quality, susceptibility to evolving cyber threats, resource-intensive training, and challenges in optimizing detection accuracy and computational efficiency, hindering its deployment in cross-domain industrial scenarios. In the future, we will enhance the anomaly-based IDS with further improvements by using the unsupervised algorithms and the integration of advanced anomaly detection techniques to enhance the model's adaptability to emerging threats and further investigate the model's scalability for large-scale IoT healthcare industry.

References

[1] Rasheed Ahmad, Izzat Alsmadi, Wasim Alhamdani, and Lo'ai Tawalbeh. A comprehensive deep learning benchmark for iot ids. *Computers & Security*, 114:102588, 2022.

[2] Mohammed M Alani. An explainable efficient flow-based industrial iot intrusion detection system. *Computers and Electrical Engineering*, 108:108732, 2023.

[3] Mohammed M Alani, Atefeh Mashatan, and Ali Miri. Xmednn: An explainable deep neural network system for intrusion detection in internet of medical things. In *ICISSP*, pages 144–151, 2023.

[4] Khalid Albulayhi, Qasem Abu Al-Haija, Suliman A Alsuhibany, Ananth A Jillepalli, Mohammad Ashrafuzzaman, and Frederick T Sheldon. Iot intrusion detection using machine learning with a novel high performing feature selection method. *Applied Sciences*, 12(10):5015, 2022.

[5] Hasan Alkahtani and Theyazn HH Aldhyani. Intrusion detection system to advance internet of things infrastructure-based deep learning algorithms. *Complexity*, 2021:1–18, 2021.

[6] Mohammed Amin Almaiah, Aitizaz Ali, Fahima Hajjej, Muhammad Fermi Pasha, and Manal Abdullah Alohali. A lightweight hybrid deep learning privacy preserving model for fc-based industrial internet of medical things. *Sensors*, 22(6):2112, 2022.

[7] Ahmad S Almogren. Intrusion detection in edge-of-things computing. *Journal of Parallel and Distributed Computing*, 137:259–265, 2020.

[8] Jafar A Alzubi, Omar A Alzubi, Issa Qiqieh, and Ashish Singh. A blended deep learning intrusion detection framework for consumable edge-centric iomt industry. *IEEE Transactions on Consumer Electronics*, 2024.

[9] Joseph Bamidele Awotunde, Kazeem Moses Abiodun, Emmanuel Abidemi Adeniyi, Sakinat Oluwabukonla Folorunso, and Rasheed Gbenga Jimoh. A deep learning-based intrusion detection technique for a secured iomt system. In *Informatics and Intelligent Applications: First International Conference, ICIIA 2021, Ota, Nigeria, November 25–27, 2021, Revised Selected Papers*, pages 50–62. Springer, 2022.

[10] Giampaolo Bovenzi, Giuseppe Aceto, Domenico Ciuonzo, Valerio Persico, and Antonio Pescapé. A hierarchical hybrid intrusion detection approach in iot scenarios. In *GLOBECOM 2020-2020 IEEE global communications conference*, pages 1–7. IEEE, 2020.

[11] Rajasekhar Chaganti, Azrour Mourade, Vinayakumar Ravi, Naga Vemprala, Amit Dua, and Bharat Bhushan. A particle swarm optimization and deep learning approach for intrusion detection system in internet of medical things. *Sustainability*, 14(19):12828, 2022.

[12] Nitesh V Chawla, Kevin W Bowyer, Lawrence O Hall, and W Philip Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[13] Ayesha S Dina, AB Siddique, and D Manivannan. A deep learning approach for intrusion detection in internet of things using focal loss function. *Internet of Things*, 22:100699, 2023.

[14] Elias Dritsas and Maria Trigka. Machine learning methods for hypercholesterolemia long-term risk prediction. *Sensors*, 22(14), 2022.

[15] Bhaskara S Egala, Ashok K Pradhan, Venkataramana Badarla, and Saraju P Mohanty. Fortified-chain: a blockchain-based framework for security and privacy-assured internet of medical things with effective access control. *IEEE Internet of Things Journal*, 8(14):11717–11731, 2021.

[16] Samir Fenanir and Fouzi Semchedine. Smart intrusion detection in iot edge computing using federated learning. *Revue d'Intelligence Artificielle*, 37(5), 2023.

[17] Marwa Fouda, Riadh Ksantini, and Wael Elmedany. A novel intrusion detection system for internet of healthcare things based on deep subclasses dispersion information. *IEEE Internet of Things Journal*, 2022.

[18] Rui Fu, Xiaojun Ren, Ye Li, Yongtang Wu, Hao Sun, and Mohammed Abdulhakim Al-Absi. Machine learning-based uav assisted agricultural information security architecture and intrusion detection. *IEEE Internet of Things Journal*, 2023.

[19] Tarek Gaber, Joseph B Awotunde, Sakinat O Folorunso, Sunday A Ajagbe, Esraa Eldesouky, et al. Industrial internet of things intrusion detection method using machine learning and optimization techniques. *Wireless Communications and Mobile Computing*, 2023, 2023.

[20] Honghao Gao, Baobin Dai, Huaikou Miao, Xiaoxian Yang, Ramon J Duran Barroso, and Hussain Walayat. A novel gapg approach to automatic property generation for formal verification: The gan perspective. *ACM Transactions on Multimedia Computing, Communications and Applications*, 19(1):1–22, 2023.

[21] Honghao Gao, Binyang Qiu, Ye Wang, Si Yu, Yueshen Xu, and Xinheng Wang. Tbdb: Token bucket-based dynamic batching for resource scheduling supporting neural network inference in intelligent consumer electronics. *IEEE Transactions on Consumer Electronics*, 2023.

[22] Honghao Gao, Xuejie Wang, Wei Wei, Anwer Al-Dulaimi, and Yueshen Xu. Com-ddpg: task offloading based on multiagent reinforcement learning for information-communication-enhanced mobile edge computing in the internet of vehicles. *IEEE Transactions on Vehicular Technology*, 2023.

[23] Honghao Gao, Lin Zhou, Jung Yoon Kim, Ying Li, and Wanqiu Huang. Applying probabilistic model checking to the behavior guidance and abnormality detection for a-mci patients under wireless sensor network. *ACM Transactions on Sensor Networks*, 19(3):1–24, 2023.

[24] Karan Gupta, Deepak Kumar Sharma, Koyel Datta Gupta, and Anil Kumar. A tree classifier based network intrusion detection model for internet of medical things. *Computers and Electrical Engineering*, 102:108158, 2022.

[25] Anar A Hady, Ali Ghubaish, Tara Salman, Devrim Unal, and Raj Jain. Intrusion detection system for healthcare systems using medical and network data: A comparison study. *IEEE Access*, 8:106576–106584, 2020.

[26] M Hasan. Number of connected iot devices growing 18% to 14.4 billion globally. 2022.

[27] Danish Javeed, Muhammad Shahid Saeed, Ijaz Ahmad, Prabhat Kumar, Alireza Jolfaei, and Muhammad Tahir. An intelligent intrusion detection system for smart consumer electronics network. *IEEE Transactions on Consumer Electronics*, 2023.

[28] Nicholas Jeffrey, Qing Tan, and José R Villar. A hybrid methodology for anomaly detection in cyber–physical systems. *Neurocomputing*, 568:127068, 2024.

[29] Sanmeet Kaur and Maninder Singh. Hybrid intrusion detection and signature generation using deep recurrent neural networks. *Neural Computing and Applications*, 32:7859–7877, 2020.

[30] Inam Ullah Khan, Muhammad Yaseen Ayub, Asrin Abdollahi, and Arijit Dutta. A hybrid deep learning model-based intrusion detection system for emergency planning using iot-network. In *2023 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*, pages 1–5. IEEE, 2023.

[31] Muneeb Hassan Khan, Abdul Rehman Javed, Zafar Iqbal, Muhammad Asim, and Ali Ismail Awad. Divacan: Detecting in-vehicle intrusion attacks on a controller area network using ensemble learning. *Computers & Security*, page 103712, 2024.

[32] Soneila Khan and Adnan Akhunzada. A hybrid dl-driven intelligent sdn-enabled malware detection framework for internet of medical things (iomt). *Computer Communications*, 170:209–216, 2021.

[33] Ilhan Firat Kilincer, Fatih Ertam, Abdulkadir Sengur, Ru-San Tan, and U Rajendra Acharya. Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization. *Biocybernetics and Biomedical Engineering*, 43(1):30–41, 2023.

[34] Dongsun Kim, Yida Tao, Sunghun Kim, and Andreas Zeller. Where should we fix this bug? a two-phase recommendation model. *IEEE transactions on software Engineering*, 39(11):1597–1610, 2013.

[35] Prabhat Kumar, Govind P Gupta, and Rakesh Tripathi. An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for iomt networks. *Computer Communications*, 166:110–124, 2021.

[36] Beibei Li, Yuhao Wu, Jiarui Song, Rongxing Lu, Tao Li, and Liang Zhao. Deepfed: Federated deep learning for intrusion detection in industrial cyber–physical systems. *IEEE Transactions on Industrial Informatics*, 17(8):5615–5624, 2020.

[37] Yuxi Li, Yue Zuo, Houbing Song, and Zhihan Lv. Deep learning in security of internet of things. *IEEE Internet of Things Journal*, 9(22):22133–22146, 2021.

[38] Wai Weng Lo, Siamak Layeghy, Mohanad Sarhan, Marcus Gallagher, and Marius Portmann. E-graphsage: A graph neural network based intrusion detection system for iot. In *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, pages 1–9. IEEE, 2022.

[39] Manisha Malik, Maitreyee Dutta, et al. Feature engineering and machine learning framework for ddos attack detection in the standardized internet of things. *IEEE Internet of Things Journal*, 2023.

[40] Mouaad Mohy-eddine, Azidine Guezzaz, Said Benkirane, and Mourade Azrour. An effective intrusion detection approach based on ensemble learning for iiot edge computing. *Journal of Computer Virology and Hacking Techniques*, 19(4):469–481, 2023.

[41] Fatemeh Mosaiyebzadeh, Seyedamin Pouriyeh, Reza M Parizi, Meng Han, and Daniel Macêdo Batista. Intrusion detection system for ioht devices using federated learning. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE, 2023.

[42] Sudarshan Nandy, Mainak Adhikari, Mohammad Ayoub Khan, Varun G Menon, and Sandeep Verma. An intrusion detection mechanism for secured iomt framework based on swarm-neural network. *IEEE Journal of Biomedical and Health Informatics*, 26(5):1969–1976, 2021.

[43] Dat-Thinh Nguyen and Kim-Hung Le. The robust scheme for intrusion detection system in internet of things. *Internet of Things*, 24:100999, 2023.

[44] Cheolhee Park, Jonghoon Lee, Youngsoo Kim, Jong-Geun Park, Hyunjin Kim, and Dowon Hong. An enhanced ai-based network intrusion detection system using generative adversarial networks. *IEEE Internet of Things Journal*, 10(3):2330–2345, 2023.

[45] Jing Qiu, Zhihong Tian, Chunlai Du, Qi Zuo, Shen Su, and Binxing Fang. A survey on access control in the age of internet of things. *IEEE Internet of Things Journal*, 7(6):4682–4696, 2020.

[46] Panagiotis Radoglou-Grammatikis, Panagiotis Sarigiannidis, Georgios Efstathopoulos, Thomas Lagkas, George Fragulis, and Antonios Sarigiannidis. A self-learning approach for detecting intrusions in healthcare systems. In *ICC 2021-IEEE International Conference on Communications*, pages 1–6. IEEE, 2021.

[47] Shalli Rani, Syed Hassan Ahmed, Rajneesh Talwar, Jyoteesh Malhotra, and Houbing Song. Iomt: A reliable cross layer protocol for internet of multimedia things. *IEEE Internet of things Journal*, 4(3):832–839, 2017.

[48] Vinayakumar Ravi, Tuan D Pham, and Mamoun Alazab. Deep learning-based network intrusion detection system for internet of medical things. *IEEE Internet of Things Magazine*, 6(2):50–54, 2023.

[49] Tanzila Saba. Intrusion detection in smart city hospitals using ensemble classifiers. In *2020 13th International Conference on Developments in eSystems Engineering (DeSE)*, pages 418–422. IEEE, 2020.

[50] Yakub Kayode Saheed and Micheal Olaolu Arowolo. Efficient cyber attack detection on the internet of medical things-smart environment based

on deep recurrent neural network and machine learning algorithms. *IEEE Access*, 9:161546–161554, 2021.

[51] Haseeb Tauqeer, Muhammad Munwar Iqbal, Aatka Ali, Shakir Zaman, and Muhammad Umar Chaudhry. Cyberattacks detection in iomt using machine learning techniques. *Journal of Computing & Biomedical Informatics*, 4(01):13–20, 2022.

[52] Ankit Thakkar and Ritika Lohiya. Attack classification of imbalanced intrusion data for iot network using ensemble learning-based deep neural network. *IEEE Internet of Things Journal*, 2023.

[53] Imtiaz Ullah and Qusay H Mahmoud. A scheme for generating a dataset for anomalous activity detection in iot networks. In *Canadian conference on artificial intelligence*, pages 508–520. Springer, 2020.

[54] Devrim Unal, Shada Bennbaia, and Ferhat Ozgur Catak. Machine learning for the security of healthcare systems based on internet of things and edge computing. In *Cybersecurity and Cognitive Science*, pages 299–320. Elsevier, 2022.

[55] Aditya Vikram et al. Anomaly detection in network traffic using unsupervised machine learning approach. In *2020 5th International Conference on Communication and Electronics Systems (ICCES)*, pages 476–479. IEEE, 2020.

[56] Yixuan Wu, Laisen Nie, Shupeng Wang, Zhaolong Ning, and Shengtao Li. Intelligent intrusion detection for internet of things security: A deep convolutional generative adversarial network-enabled approach. *IEEE Internet of Things Journal*, 2021.

[57] Jean-Paul A Yaacoub, Mohamad Noura, Hassan N Noura, Ola Salman, Elias Yaacoub, Raphaël Couturier, and Ali Chehab. Securing internet of medical things systems: Limitations, issues and recommendations. *Future Generation Computer Systems*, 105:581–606, 2020.

[58] Fan Yang, Qilu Wu, Xiping Hu, Jiancong Ye, Yuting Yang, Haocong Rao, Rong Ma, and Bin Hu. Internet-of-things-enabled data fusion method for sleep healthcare applications. *IEEE Internet of Things Journal*, 8(21):15892–15905, 2021.

[59] Li Yang, Abdallah Moubayed, and Abdallah Shami. Mth-ids: A multitiered hybrid intrusion detection system for internet of vehicles. *IEEE Internet of Things Journal*, 9(1):616–632, 2021.

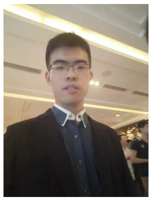
[60] Mingyuan Zang, Changgang Zheng, Lars Dittmann, and Noa Zilberman. Towards continuous threat defense: In-network traffic analysis for iot gateways. *IEEE Internet of Things Journal*, 2023.

[61] Zhao Zhang, Yong Zhang, Da Guo, Lei Yao, and Zhao Li. Secfednids: Robust defense for poisoning attack against federated learning-based network intrusion detection system. *Future Generation Computer Systems*, 134:154–169, 2022.

[62] Maede Zolanvari, Ali Ghubaish, and Raj Jain. Addai: anomaly detection using distributed ai. In *2021 IEEE International Conference on Networking, Sensing and Control (ICNSC)*, volume 1, pages 1–6. IEEE, 2021.

Umer Zukaib is currently pursuing a Ph.D. degree in Cyberspace Security at the School of Cyber Science and Engineering, Wuhan University in China. His current research interests are threat detection, monitoring, mitigation, blockchain and Machine learning.

Xiaohui Cui received the Ph.D. degree in computer science and engineering from the University of Louisville, Louisville, KY, USA, in 2004. He is currently a Professor at the School of Cyber Science and Engineering, Wuhan University. His research interests include Artificial Intelligence, Blockchain Technology, and High-performance Computing.



Chengliang Zheng is currently pursuing a Ph.D. degree in Cyberspace Security at the School of Cyber Science and Engineering, Wuhan University in China. His current research interests are blockchain and Machine learning.

Mir Hassan is currently pursuing a Ph.D. degree at the University of Trento, Italy. His current research interests are Federated Learning, blockchain, and IoT.