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*, and Zhidong Shen

***Abstract*—The Internet of Medical Things (IoMT) plays a crucial role in advancing smart healthcare by facilitating the real-time collection and processing of medical data. These interconnected devices leverage artificial intelligence to assist practitioners in making data-driven decisions. However, IoMT's dependence on communication protocols exposes it to significant security vulnerabilities. In response to this challenge, we propose a novel meta-intrusion detection system (Meta-IDS) that employs a meta-learning approach to enhance the detection of both known and zero-day intrusions. Our approach seamlessly integrates signature-based and anomaly based detection techniques, incorporating privacy-preserving methods essential for handling sensitive IoMT data. We rigorously evaluated our methodology using three publicly available data sets (WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT-2021). The results demonstrate remarkable accuracy rates of 99.57%, 99.93%, and 99.99% for signature-based detection, and 99.47%, 99.98%, and 99.99% for anomaly based detection, coupled with impressively low misclassification rates of 0.0042%, 0.0006%, and 0.00004%, respectively. Through a comparative analysis with the state-of-the-art E-GraphSAGE model, considering metrics, such as accuracy, precision, recall, F1-score, time complexity, and misclassification rate, we affirm the performance and reliability of the Meta-IDS. Our approach holds significant promise in bolstering cybersecurity within the IoMT network.**

***Index Terms*—Anomaly detection, artificial intelligence (AI), cybersecurity, Internet of Medical Things (IoMT), intrusion detection systems (IDSs), meta learning, zero-day attacks.**

I. Introduction

THE PROLIFERATION of connected devices in our daily lives is continually expanding. According to [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 (APTs) to leak information from IoMT devices [60], occasionally endangering human lives. A list of abbreviations can be found in Table I. Consequently, security monitoring becomes a top priority in IoMT-based healthcare [15].

Numerous approaches, including malicious traffic injection, man-in-the-middle (MITM) attacks, Denial of Service (DoS), 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 [6], [45].

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 (IDSs) 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

Manuscript received 13 January 2024; revised 22 February 2024 and 20 March 2024; accepted 7 April 2024. Date of publication 10 April 2024; date of current version 26 June 2024. This work was supported in part by the Key Research and Development Projects in Hubei Province under Grant 2022BAA041. *(Corresponding authors: Xioahui Cui; Zhidong Shen.)*

Umer Zukaib, Xiaohui Cui, Chengliang Zheng, and Zhidong Shen are with the Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, School of Cyber Science and Engineering, Wuhan University, Wuhan 430079, China (e-mail: umerzukaib@whu.edu.cn; xcui@whu.edu.cn; chengliang@whu.edu.cn; shenzd@whu.edu.cn).

Mir Hassan is with the Department of Information Engineering and Computer Science, University of Trento, 38123 Trento, Italy (e-mail: mir.hassan@unitn.it).

Digital Object Identifier 10.1109/JIOT.2024.3387294

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

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 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.

Our main contribution in this research article is as follows.

1) Our innovative meta-intrusion detection system (Meta-IDS) technique, comprising five base learners and a meta-learner, is specifically designed to detect known attacks targeting IoMT networks.
2) 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.
3) Addressing the challenge of zero-day attacks, we introduce mean shift clustering and biased classifiers, providing robust detection capabilities against emerging threats.
4) 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 data set is introduced in Section IV. Section V details experiments and performance evaluations. Section VI with an insightful discussion. This article 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 APTs, 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. They 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.

Park et al. [44] proposed an AI-based network IDS that effectively addresses data imbalance concerns, enhancing overall performance. Their 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 software-defined networking (SDN)-orchestrated DL-based IDS, which utilize SDN architecture for reconfiguration and a bidirectional long short-term memory (Bi-LSTM) model for attack identification. Simulations on the CICIDS-2018 data set validate its superiority over recent security solutions. However, the study lacks practical implications and applicability to the diverse IoMT network. It does not 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 data sets for effective detection of known and unknown attacks. However, its applicability to the IoMT paradigm is limited, as the data set adaptation falls short of meeting critical healthcare requirements.

Kaur and Singh [29] presented 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 (FNs) and false positives (FPs). However, the study primarily focuses on Web attacks, limiting its applicability to a broader range of diverse attack scenarios.

Bovenzi et al. [10] proposed H2ID, a 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 th ToN-IoT data set, achieving improved results and proposing a real-time deployment architecture for edge-cloud-based IoMT environments. However, the study primarily focuses

TABLE II
COMPARISON OF ARTICLES IN IDS FOR IOMT NETWORKS

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

on differentiating between attack and normal traffic using signature or rule-based methods, which may not effectively detect new attacks.

Almogren [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-Grammatikis et al. [46] introduced an active learning approach to dynamically retrain supervised classifiers to perform intrusion detection tasks. 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 cyber–physical systems (CPSs). While effective in detecting various cyber threats, the study lacks information on runtime complexity and applicability to handle multiple attack patterns.

Saheed and Arowolo [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 particle swarm optimization (PSO) feature selection. However, the study does not adequately illustrate its dynamic applicability in a cross-domain IoMT network.

Khan and Akhunzada [32] introduced an SDN-based long short-term memory (LSTM) and convolutional neural network (CNN) framework, showcasing promising results in detecting malware in the IoMT network. However, the study lacks information on novel attack detection, adaptation to the IoMT environment, and its time complexity.

Zhang et al. [61] introduced the 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.

Recent algorithms, such as the GAN-based model [20], may detect synthetic data in real network traffic, enhancing IDS applicability in real-world scenarios. Leveraging unstructured actor-modeling with discrete-time Markov chains (UAM with DTMCs) and probabilistic computation tree logic (PCTL) [23], along with the TBDB algorithm [21], may aid in threat detection and enhance security in IoMT and industrial IoT (IIoT) environments. Moreover, the Com-DDPG approach [22] may demonstrate its effectiveness in task offloading for robust IDS deployment across diverse network domains.

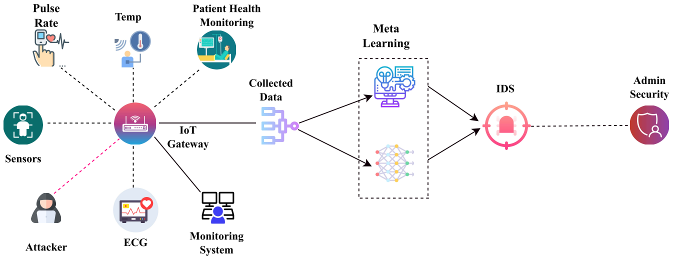


Fig. 1. IoMT-based IDS.

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 utilized network traffic and patient biometric data to improve attack detection in the IoMT. Fig. 1 illustrates the IoMT network employing ML and DL for cybersecurity.

The IDS incorporates IoT sensors, a network-traffic controller, an IoT gateway, and ML/DL pipelines for data preparation. A variety of sensors, such as pulse-rate, temperature, and ECG sensors, are utilized in this setup. The IoMT network employs MQTT, CoAP, AMQP, and DDS messaging protocols tailored to healthcare requirements. Sensor data is collected by the IoT gateway through both wired and wireless communication channels. After data preprocessing, it is forwarded to the IDS for tuning and mitigating FPs.

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	

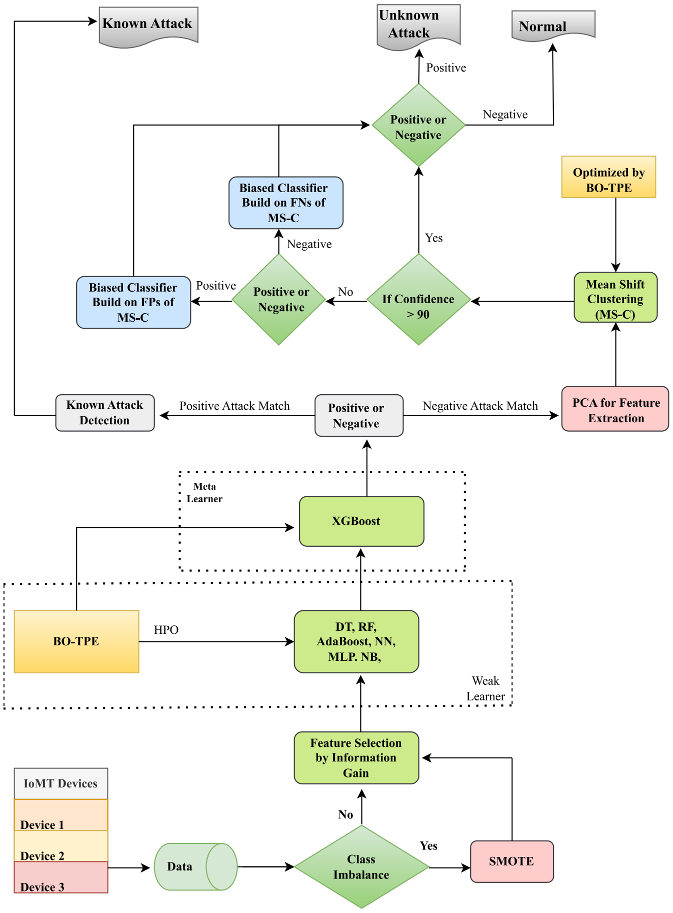


Fig. 2. Proposed framework of Meta-IDS.

A. System Architecture

The Meta-IDS, illustrated in Fig. 2, leverages meta-learning techniques for intrusion detection in the IoMT network. It initiates with the collection of data from network traffic and IoMT sensors, followed by essential feature engineering and data preprocessing steps. Class imbalance is addressed using synthetic minority over-sampling technique (SMOTE), while recursive feature elimination (RFE) and linear discriminant analysis (LDA) are employed to enhance model performance. The meta-learning process unfolds in two stages. First, hyperparameter optimization (HPO) is conducted for weak learners, including decision trees (DTs), random forests (RFs), AdaBoost, multinomial naive Bayes (MNB), neural networks (NNs), and multilayer perceptrons (MLPs), using the Bat Algorithm. Subsequently, XGBoost is utilized as the meta-learner with HPO facilitated by Bayesian-optimization-based tree-structured Parzen-estimator (BO-TPE). The meta-learner orchestrates signature-based intrusion detection, while anomaly detection is executed through mean-shift clustering (MS-CL). The combination of BO-TPE for HPO and two biased classifiers optimizes the process, minimizing FPs and FNs while ensuring accurate classification. For detailed algorithmic insights, please refer to Table III.

B. Data Preprocessing

Addressing class imbalance is crucial in IoT data, where standard samples outnumber attack samples, potentially biasing the model. Resampling techniques, such as SMOTE [12] play a vital role. Unlike random sampling, SMOTE generates new minority-class samples, ensuring balanced data sets, as mentioned in Algorithm 1.

In our approach, we employed SMOTE to tackle class imbalance, thereby ensuring a robust and unbiased model.

1) Data Normalization: After applying SMOTE to balance the data set, we proceed with data normalization. Categorical attributes are converted to numeric representations using a label encoder, a crucial step as ML models require numeric

Algorithm 1 Algorithm for Data Preprocessing

```
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 $D_{smoted}$
```

Algorithm 2 RFE

```
Input: Training set ($pKa$)
    Set of features ($F = \{f_1, \dots, f_m\}$)
    Number of features to select ($N = \{M, \dots, 10, 5\}$)
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
```

inputs. Subsequently, normalization is conducted to mitigate biases in ML results induced by large feature scales. The data is normalized to achieve a mean of 0 and a standard deviation of 1, ensuring optimal support for the ML 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

Following the initial data preprocessing, a high-quality data set is obtained; however, further refinement is essential for optimal feature selection. Feature engineering, conducted prior to feeding data into ML models, aims to eliminate noise and irrelevant features, thereby enhancing data quality. We utilize RFE based on RF (RF-RFE), from Algorithm 2 leveraging mean decrease in accuracy (MDA) as demonstrated in (2) to evaluate variable importance. This meticulous process ensures that only essential traits are retained, facilitating more accurate and efficient predictions

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

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

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 data sets by simplifying the establishment of cutoff points.

D. Proposed Meta-IDS

The IDS is responsible for classifying signature-based and anomaly based attacks. The signature-based IDS identifies known attack patterns using supervised ML techniques but faces challenges in detecting new patterns. On the other hand, the anomaly based IDS distinguishes normal behavior from unknown attacks using unsupervised learning, assuming that new attacks share statistical similarities with known attacks. The Meta-IDS (as shown in Algorithms 3 and 4) stands out for its proficiency in detecting both signature-based (known) and anomaly based (zero-day) attacks, ensuring high accuracy while maintaining minimal execution time.

1) Signature-Based IDS: Following data preprocessing and feature engineering, labeled data undergoes training using meta-learning to construct the signature-based IDS. In the meta-learning phase, a distinction is made between weak learners (DT, AdaBoost, RF, MLP, NN, and MNB) and a meta-learner (XGBoost).

DT is a tree-based model offering multiple tunable hyper-parameters, including minimum and maximum sample split, sample nodes, tree depth, sample leaf, weight fraction of leaf, etc. RF employs an ensemble learning approach that combines multiple DTs using the majority-voting rule. NN, inspired by the human brain's information processing, excels at identifying hidden patterns and improving over time. AdaBoost is a boosting technique that enhances ML algorithm performance by reassigning weights to each instance. MLP complements the feed-forward NN and is suitable for both small and large data sets. Multinomial naive Bayes is recognized for building a fast ML model with reliable predictions.

Hyper-parameter optimization for the 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:

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

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

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

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

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 $j = 1, \ldots n$, do
    for $i = 1, \ldots k$, 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 $M'$
```

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

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

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 as follows.

1) Ensemble models (RF and AdaBoost) excel with large, complex data sets, 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 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 EI,
$f(x)$ 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; and

Algorithm 4 Anomaly Based IDS

```
Input: Training dataset $D = \{(x_1, c_1), (x_2, c_2), \ldots (x_n, c_n)\}$
    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
```

$(\cdot)^+$ 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.

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 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 FPs and 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 RF, 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 FPs and FNs from the initial training phase. In our anomaly IDS, the application of MS-CL 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, an 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, and 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, and 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, various interfaces or diagnostic tools, such as DICOM, HL7, IEEE 11073, and IHE, have been utilized for data sharing or electronic control units (ECUs). Each has its pros and cons, but we have opted for HL7, a standard protocol used for the sharing, integration, exchange, and retrieval of electronic health records (EHRs), which is commonly employed for sharing healthcare data.

Fig. 3 illustrates the proposed Meta-IDS protected healthcare architecture. Different IoMT devices (IoMT cluster) gather data and pass it to the HL7 interface module, responsible for handling HL7 communication and managing data gathering. Subsequently, the data is transferred to the network via the HL7 interface module. The network TAP on the

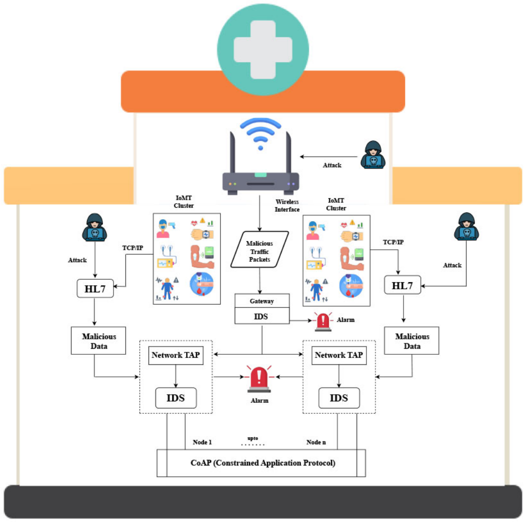


Fig. 3. Real-time architectural diagram.

network infrastructure intercepts and captures HL7 traffic, connecting the network TAP to the network. In the event of a breach in the HL7 protocol (if not adequately secured), the attacker can inject malicious data. Upon receiving the data, the network TAP processes it through our proposed Meta-IDS. If any attack is detected, it triggers an alarm. Similarly, 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 Lo et al. [38], utilize edge features (EFs) 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 data sets, utilizing only the provided EFs $e_{uv}, \forall_{uv} \in \mathcal{E}$. Since the data sets lack node features (NFs), 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

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

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.

Equation (9) shows the standard GraphSAGE model, but the E-GraphSAGE model used the aggregated embedding of the sample at adjacent edges of the kth layer, as explained in the following equation:

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

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 $z_{u,v}, \forall_{uv}, \in \mathcal{E}$
    $h^o_v \leftarrow x_v, \forall v \in \mathcal{V}$
for $k \leftarrow 1ToK$ 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
```

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:

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

The crucial distinction lies in the fact that the E-GraphSAGE algorithm computes $h^k_{\mathcal{N}(v)})$ using (11), which incorporates EF. Each graph node's k-hop neighborhood 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

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

The E-GraphSAGE model comprises three primary steps: 1) constructing a network graph using network-flow data; 2) training a supervised model with the generated data; and 3) creating edge embeddings to enable edge classification, discerning between normal and attack classes.

1) 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: 1) Sport; 2) Dport; 3) SrcAddr; and 4) 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 an NN comprising two layers of E-GraphSAGE (k=2), capturing data aggregated from a two-hop neighborhood. The mean of EFs 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, and e^{k-1}_{uv} shows EFs 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 data set 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. Data Set Description

The choice of an appropriate data set for model evaluation is a crucial step that demands careful consideration. To fulfill the objectives of experimental research on the proposed model, we utilized three data sets: 1) WUSTL-EHMS-2020 [25]; 2) IoTID20 [53]; and 3) WUSTL-IIOT-2021 [62].

These data sets were selected to comprehensively evaluate the proposed IDS across diverse scenarios. The WUSTL-EHMS-2020 data set assesses the IDS in healthcare environments, emphasizing data and system security. The IoTID20 data set evaluates the IDS in IoT networks, capturing unique challenges and attack patterns prevalent in such environments. Finally, the WUSTL-IIOT-2021 data set extends evaluation to IIoT settings, addressing cyber threats in critical infrastructure. By leveraging these data sets, we ensure the adaptability and effectiveness of the IDS across varied domains and security requirements.

A. WUSTL-EHMS-2020 Data Set

The data set, acquired 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 both sensor and network traffic data, which are analyzed to detect intrusion sources and anomalous patterns, including three types of attacks: 1) data injection; 2) spoofing; and 3) MITM attacks. This data set comprises 44 features, with 35 related to the network, and 8 to patient biometrics, along with corresponding class labels. Detailed statistical information regarding this data set is provided in Table IV.

B. IoTID20 Data Set

The IoTID20 data set, derived from IoT devices, comprises 80 features across 625 783 samples, of which 585 710 are labeled as malicious. This data set encompasses various cyberattacks, including DoS, Mirai (a botnet malware), MITM, and scan (port scanning). DoS attacks flood systems with unauthorized requests to disrupt normal operations, while Mirai transforms IoT devices into a botnet for extensive distributed DoS (DDoS) attacks. MITM attacks clandestinely intercept and potentially alter communication, while scan attacks methodically search for vulnerabilities. A summary of the data set is provided in Table V.

TABLE IV
Description of WUSTL-EHMS-2020 Data Set

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 Data Set

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

TABLE VI
Description of WUSTL-IIOT-2021 Data Set

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 %

C. WUSTL-IIOT-2021 Data Set

The WUSTL-IIOT-2021 data set, 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 data set 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.

TABLE VII
SIGNATURE-BASED EVALUATION OF THE PROPOSED META-IDS ON THREE DATA SETS

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**

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. HPO 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 Core i5-1135G7 processor with clock speeds of 2.40 and 2.42 GHz, along with 16 GB of RAM.

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), FP, and FN rates. The evaluation metrics have been computed using the following equations:

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

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

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

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

C. Evaluation of Known Intrusions' Performance

To ensure generalizability and mitigate overfitting risks in our experiments, we utilized hold-out and cross-validation (CV) methods. The procedures for model evaluation and data set split are outlined below as follows.

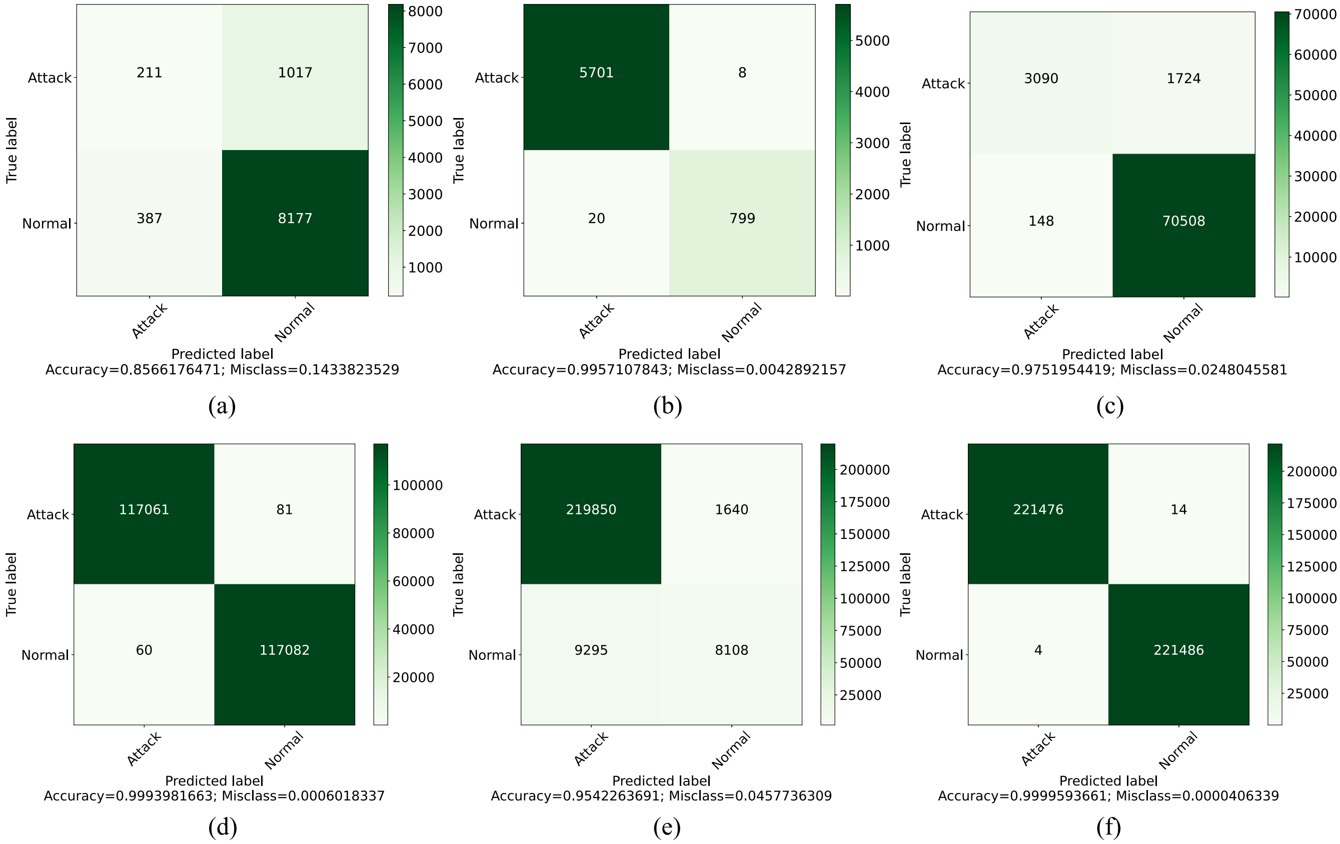
1) Seventy percent of the data was allocated for training, while the remaining 30% was set aside for testing. The test data remained unaltered throughout the hold-out validation process.
2) We assessed the model's efficacy through ten-fold CV on distinct subsets of the training data set. In each fold, 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 data set.

Employing a 70%–30% test-train split and ten-fold CV technique addresses concerns related to concept drift and overfitting [34]. This approach ensures consistent intrusion detection while mitigating the risk of overfitting or underfitting. To effectively identify unknown attacks, a hold-out validation strategy is employed. In this approach, all data samples except those corresponding to unknown attacks are utilized as training sets. The system's efficacy in identifying new attack patterns is then assessed using a validation set specifically created for unknown attacks. This evaluation involves comparing the statistical similarity of these unknown attacks 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 data sets. Results, optimized through hold-out and ten-fold CV, are summarized in Table VII.

On the WUSTL-EHMS 2020 data set, our proposed Meta-IDS was compared with several state-of-the-art techniques [3], [11], [13], [24], [33], [41], [48]. The Meta-IDS

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



(a) (b) (c)

(d) (e) (f)

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

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 s. On the WUSTL-IIOT-2021 data set, 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 [2], [13], [19], [40]. Additionally, it demonstrates minimal execution time (307.3 s). On the IoTID20 data set, our Meta-IDS showcased remarkable accuracy at 99.93%, surpassing alternative techniques [4], [5], [16], [30], [43]. Furthermore, it exhibited minimal execution time (247.3 s).

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 s. 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.

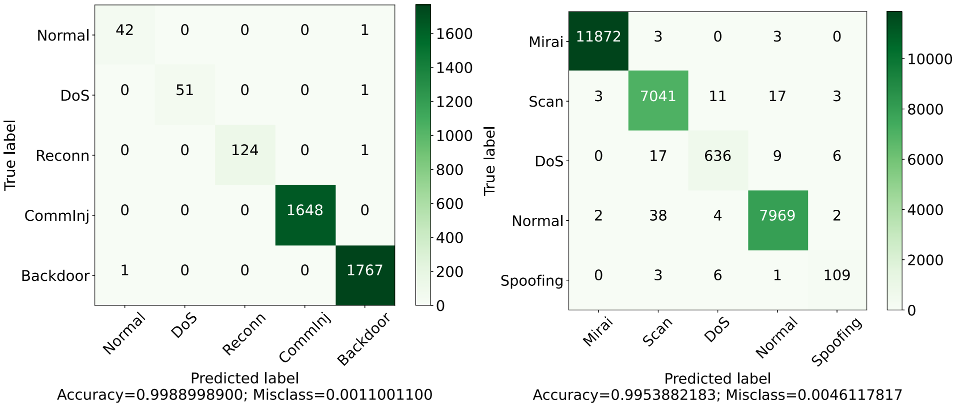
Fig. 4(a) presents the confusion matrix for E-GraphSAGE on WUSTL-EHMS-2020, achieving 85.66% accuracy with a 0.1433% misclassification rate. In contrast, Fig. 4(b) illustrates our Meta-IDS outperforming, attaining 99.57% accuracy and an impressively low 0.0042% misclassification rate.

For the IoTID20 data set [Fig. 4(c)], E-GraphSAGE achieved 97.51% accuracy with a 0.0248% misclassification rate. However, Fig. 4(d) showcases our Meta-IDS excelling, reaching 99.93% accuracy and an exceptionally low 0.0006% misclassification rate.

Turning to WUSTL-IIOT 2021 [Fig. 4(e)], E-GraphSAGE achieved 95.42% accuracy with a 0.0457% misclassification rate. Conversely, Fig. 4(f) illustrates our Meta-IDS dominance, securing 99.99% accuracy and an exceptionally low 0.00004% misclassification rate.

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

In Fig. 6(a), the AUC on WUSTL-EHMS-2020 for the Meta-IDS and weak-learners are as follows: NN (70.71%),

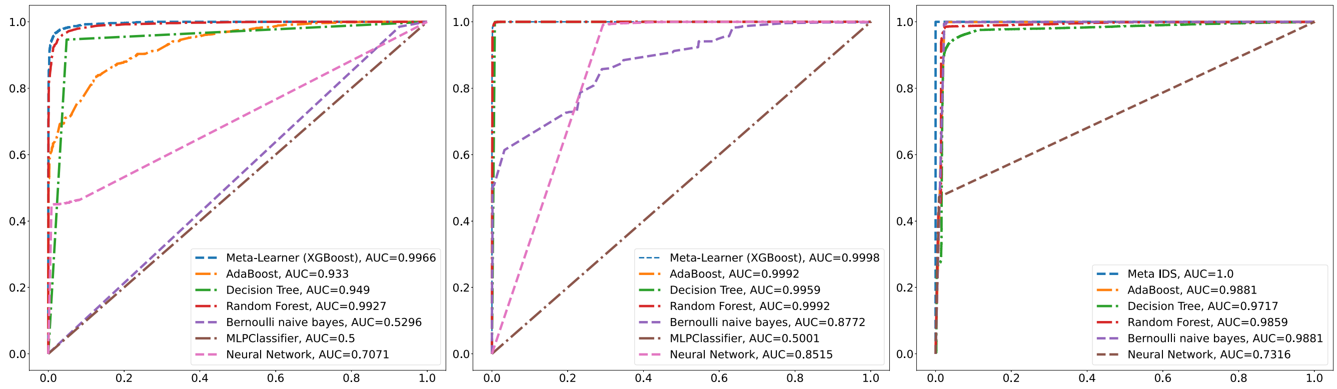


(a) (b)

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

TABLE VIII
MULTICLASS EVALUATION OF META-IDS ON WUSTL-IIOT 2021 AND IOTIO20 DATA SET

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%



(a) (b) (c)

Fig. 6. AUC on the (a) WUSTL-EHMS-2020, (b) IoTID20, and (c) WUSTL-IIOT 2021 data set.

MLP (50%), naive Bayes (53%), RF (99.2%), DT (94.9%), AdaBoost (93.3%), and meta-learner (99.66%). Fig. 6(b) 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%), and meta-learner (99.98%). In Fig. 6(c), AUC on WUSTL-IIOT 2021 for the Meta-IDS and weak-learners are as follows: NN (73.16%), naive Bayes (98.81%), RF (98.59%), DT (97.17%), AdaBoost (98.81%), and meta-learner (100%).

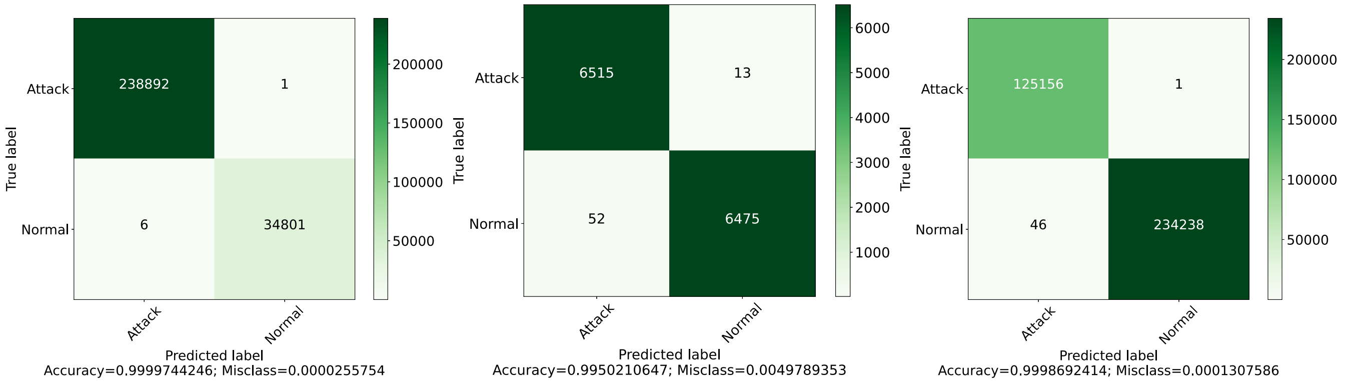
D. Performance Analysis of Anomaly Based IDS

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 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 MS-CL and biased classifiers across WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT 2021 data sets. MS-CL 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 data sets, showcasing enhanced precision, recall, and F1-score performance.

TABLE IX
EVALUATION OF ANOMALY IDS ON DIFFERENT DATA SETS

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. 7. Confusion matrix of anomaly based IDS on (a) WUSTL-IIOT 2021 data set, (b) WUSTL-EHMS-2020 data set, and (c) IoTID20 data set.

Fig. 7(a)–(c) 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 detection of known and unknown attacks in IoMT networks. Leveraging signature-based IDS for known attacks and anomaly based IDS for zero-day attacks, Meta-IDS outperforms existing models in terms of overall accuracy, precision, recall, F1-score, and execution time. We also incorporate E-GraphSAGE, an explainable AI-based graph NN model, and demonstrate its superior performance compared to Meta-IDS through detailed evaluation across three data sets, as summarized in Table VII. Further insights into model performance are provided through confusion matrices, depicting TPs, TNs, FPs, FNs, accuracy, and misclassification rates across different data sets [Fig. 4(a)–(f)].

Meta-learners demonstrate superior performance in area under the curve (AUC) compared to other weak-learners, as illustrated in Fig. 6(a)–(c), emphasizing our model's generalizability and avoidance of overfitting. Additionally, the anomaly based IDS addresses zero-day attack detection by utilizing MS-CL, initially yielding unsatisfactory results. To enhance performance, we introduce a biased classifier (random forest - RF) to accurately classify FNs and FPs, significantly improving overall detection performance (Table IX). Our experimental results confirm the efficiency and accuracy of Meta-IDS in detecting both known and unknown attacks in IoMT networks.

The optimized architecture of Meta-IDS, featuring a stack of weak learners and a meta-learner, ensures scalability, adaptability to diverse data sets, and proficiency in handling multiattack scenarios, particularly in large-scale IoMT networks. Its minimal execution time further enhances its suitability for deployment in complex network environments. However, resource-constrained IoMT and IIoT environments present challenges related to environmental variations, diverse data sources, and interoperability. Safeguarding Meta-IDS against adversarial attacks and ensuring compliance with data privacy standards are crucial for its deployment in real-world scenarios. Addressing these challenges is essential to maintain Meta-IDS's adaptability, security, and compliance within the IoMT and IIoT domains.

TABLE X
ABLATION STUDY ON DIFFERENT DATA SETS 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 %

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, an MS-CL 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 MS-CL algorithm. Tested on WUSTL-EHMS-2020, IoTID20, and WUSTL-IIOT-2021 data sets, 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 data set 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] R. Ahmad, I. Alsmadi, W. Alhamdani, and L. Tawalbeh, "A comprehensive deep learning benchmark for IoT IDS," *Comput. Secur.*, vol. 114, Mar. 2022, Art. no. 102588.

[2] M. M. Alani, "An explainable efficient flow-based Industrial IoT intrusion detection system," *Comput. Elect. Eng.*, vol. 108, May 2023, Art. no. 108732.

[3] M. M. Alani, A. Mashatan, and A. Miri, "XMeDNN: An explainable deep neural network system for intrusion detection in Internet of Medical Things," in *Proc. 9th Int. Conf. Inf. Syst. Secur. Priv.*, 2023, pp. 144–151.

[4] K. Albulayhi, Q. A. Al-Haija, S. A. Alsuhibany, A. A. Jillepalli, M. Ashrafuzzaman, and F. T. Sheldon, "Iot intrusion detection using machine learning with a novel high performing feature selection method," *Appl. Sci.*, vol. 12, no. 10, p. 5015, 2022.

[5] H. Alkahtani and T. Aldhyani, "Intrusion detection system to advance Internet of Things infrastructure-based deep learning algorithms," *Complexity*, vol. 2021, pp. 1–18, Jul. 2021.

[6] M. A. Almaiah, A. Ali, F. Hajjej, M. F. Pasha, and M. A. Alohali, "A lightweight hybrid deep learning privacy preserving model for FC-based Industrial Internet of Medical Things," *Sensors*, vol. 22, no. 6, p. 2112, 2022.

[7] A. S. Almogren, "Intrusion detection in edge-of-things computing," *J. Parallel Distrib. Comput.*, vol. 137, pp. 259–265, Mar. 2020.

[8] J. A. Alzubi, O. A. Alzubi, I. Qiqieh, and A. Singh, "A blended deep learning intrusion detection framework for consumable edge-centric IoMT industry," *IEEE Trans. Consum. Electron.*, early access, Jan. 5, 2024, doi: 10.1109/TCE.2024.3350231.

[9] 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," in *Proc. Int. Conf. Informat. Intell. Appl.*, 2022, pp. 50–62.

[10] G. Bovenzi, G. Aceto, D. Ciuonzo, V. Persico, and A. Pescapé, "A hierarchical hybrid intrusion detection approach in IoT scenarios," in *Proc. IEEE Global Commun. Conf.*, 2020, pp. 1–7.

[11] R. Chaganti, A. Mourade, V. Ravi, N. Vemprala, A. Dua, and B. Bhushan, "A particle swarm optimization and deep learning approach for intrusion detection system in Internet of Medical Things," *Sustainability*, vol. 14, no. 19, 2022, Art. no. 12828.

[12] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, 2002.

[13] A. S. Dina, A. B. Siddique, and D. Manivannan, "A deep learning approach for intrusion detection in Internet of Things using focal loss function," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100699.

[14] E. Dritsas and M. Trigka, "Machine learning methods for hypercholesterolemia long-term risk prediction," *Sensors*, vol. 22, no. 14, p. 5365, 2022.

[15] B. S. Egala, A. K. Pradhan, V. Badarla, and S. P. Mohanty, “Fortified-chain: A blockchain-based framework for security and privacy-assured Internet of Medical Things with effective access control,” *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11717–11731, Jul. 2021.

[16] S. Fenanir and F. Semchedine, “Smart intrusion detection in IoT edge computing using federated learning,” *Revue d’Intell. Artificielle*, vol. 37, no. 5, pp. 1–13, 2023.

[17] M. Fouda, R. Ksantini, and W. Elmedany, “A novel intrusion detection system for Internet of Healthcare Things based on deep subclasses dispersion information,” *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8395–8407, May 2023.

[18] R. Fu, X. Ren, Y. Li, Y. Wu, H. Sun, and M. A. Al-Absi, “Machine learning-based UAV assisted agricultural information security architecture and intrusion detection,” *IEEE Internet Things J.*, vol. 10, no. 21, pp. 18589–18598, Nov. 2023.

[19] T. Gaber, J. B. Awotunde, S. O. Folorunso, S. A. Ajagbe, E. Eldesouky, and others, “Industrial Internet of Things intrusion detection method using machine learning and optimization techniques,” *Wireless Commun. Mobile Comput.*, vol. 2023, Apr. 2023, Art. no. 3939895.

[20] H. Gao, B. Dai, H. Miao, X. Yang, R. J. D. Barroso, and H. Walayat, “A novel GAPG approach to automatic property generation for formal verification: The GAN perspective,” *ACM Trans. Multimedia Comput., Commun. Appl.*, vol. 19, no. 1, pp. 1–22, 2023.

[21] H. Gao, B. Qiu, Y. Wang, S. Yu, Y. Xu, and X. Wang, “TBDB: Token bucket-based dynamic batching for resource scheduling supporting neural network inference in intelligent consumer electronics,” *IEEE Trans. Consum. Electron.*, early access, Dec. 5, 2023, doi: 10.1109/TCE.2023.3339633.

[22] H. Gao, X. Wang, W. Wei, A. Al-Dulaimi, and Y. Xu, “Com-DDPG: Task offloading based on multiagent reinforcement learning for information-communication-enhanced mobile edge computing in the Internet of Vehicles,” *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 348–361, Jan. 2024.

[23] H. Gao, L. Zhou, J. Y. Kim, Y. Li, and W. Huang, “Applying probabilistic model checking to the behavior guidance and abnormality detection for A-MCI patients under wireless sensor network,” *ACM Trans. Sens. Netw.*, vol. 19, no. 3, pp. 1–24, 2023.

[24] K. Gupta, D. K. Sharma, K. D. Gupta, and A. Kumar, “A tree classifier based network intrusion detection model for Internet of Medical Things,” *Comput. Elect. Eng.*, vol. 102, Sep. 2022, Art. no. 108158.

[25] A. A. Hady, A. Ghubaish, T. Salman, D. Unal, and R. Jain, “Intrusion detection system for healthcare systems using medical and network data: A comparison study,” *IEEE Access*, vol. 8, pp. 106576–106584, 2020.

[26] M. Hasan, *Number of Connected IoT Devices Growing 18% to 14.4 Billion Globally*, Energy Conf. Netw., Katy, TX, USA, 2022.

[27] D. Javeed, M. S. Saeed, I. Ahmad, P. Kumar, A. Jolfaei, and M. Tahir, “An intelligent intrusion detection system for smart consumer electronics network,” *IEEE Trans. Consum. Electron.*, vol. 69, no. 4, pp. 906–913, Nov. 2023.

[28] N. Jeffrey, Q. Tan, and J. R. Villar, “A hybrid methodology for anomaly detection in cyber–physical systems,” *Neurocomputing*, vol. 568, Feb. 2024, Art. no. 127068.

[29] S. Kaur and M. Singh, “Hybrid intrusion detection and signature generation using deep recurrent neural networks,” *Neural Comput. Appl.*, vol. 32, pp. 7859–7877, Jun. 2020.

[30] I. U. Khan, M. Y. Ayub, A. Abdollahi, and A. Dutta, “A hybrid deep learning model-based intrusion detection system for emergency planning using IoT-network,” in *Proc. Int. Conf. Inf. Commun. Technol. Disaster Manage. (ICT-DM)*, 2023, pp. 1–5.

[31] M. H. Khan, A. R. Javed, Z. Iqbal, M. Asim, and A. I. Awad, “DivaCAN: Detecting in-vehicle intrusion attacks on a controller area network using ensemble learning,” *Comput. Secur.*, vol. 139, Apr. 2024, Art. no. 103712.

[32] S. Khan and A. Akhunzada, “A hybrid DL-driven intelligent SDN-enabled malware detection framework for Internet of Medical Things (IoMT),” *Comput. Commun.*, vol. 170, pp. 209–216, Mar. 2021.

[33] I. F. Kilincer, F. Ertam, A. Sengur, R. Tan, and U. Acharya, “Automated detection of cybersecurity attacks in healthcare systems with recursive feature elimination and multilayer perceptron optimization,” *Biocybern. Biomed. Eng.*, vol. 43, no. 1, pp. 30–41, 2023.

[34] D. Kim, Y. Tao, S. Kim, and A. Zeller, “Where should we fix this bug? A two-phase recommendation model,” *IEEE Trans. Softw. Eng.*, vol. 39, no. 11, pp. 1597–1610, Nov. 2013.

[35] P. Kumar, G. P. Gupta, and R. Tripathi, “An ensemble learning and fog-cloud architecture-driven cyber-attack detection framework for IoMT networks,” *Comput. Commun.*, vol. 166, pp. 110–124, Jan. 2021.

[36] B. Li, Y. Wu, J. Song, R. Lu, T. Li, and L. Zhao, “DeepFed: Federated deep learning for intrusion detection in industrial cyber–physical systems,” *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5615–5624, Aug. 2021.

[37] Y. Li, Y. Zuo, H. Song, and Z. Lv, “Deep learning in security of Internet of Things,” *IEEE Internet Things J.*, vol. 9, no. 22, pp. 22133–22146, Nov. 2022.

[38] W. W. Lo, S. Layeghy, M. Sarhan, M. Gallagher, and M. Portmann, “E-GraphSAGE: A graph neural network based intrusion detection system for IoT,” in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp.*, 2022, pp. 1–9.

[39] M. Malik and M. Dutta, “Feature engineering and machine learning framework for DDoS attack detection in the standardized Internet of Things,” *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8658–8669, May 2023.

[40] M. Mohy-eddine, A. Guezzaz, S. Benkirane, and M. Azrour, “An effective intrusion detection approach based on ensemble learning for IIoT edge computing,” *J. Comput. Virol. Hack. Techn.*, vol. 19, no. 4, pp. 469–481, 2023.

[41] F. Mosaiyebzadeh, S. Pouriyeh, R. M. Parizi, M. Han, and D. Macêdo Batista, “Intrusion detection system for ioht devices using federated learning,” in *Proc. IEEE Conf. Comput. Commun. Workshops (INFOCOM WKSHPS)*, 2023, pp. 1–6.

[42] 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,” *IEEE J. Biomed. Health Inform.*, vol. 26, no. 5, pp. 1969–1976, May 2022.

[43] D. Nguyen and K. Le, “The robust scheme for intrusion detection system in Internet of Things,” *Internet of Things*, vol. 24, Dec. 2023, Art. no. 100999.

[44] C. Park, J. Lee, Y. Kim, J. Park, H. Kim, and D. Hong, “An enhanced ai-based network intrusion detection system using generative adversarial networks,” *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2330–2345, Feb. 2023.

[45] J. Qiu, Z. Tian, C. Du, Q. Zuo, S. Su, and B. Fang, “A survey on access control in the age of Internet of Things,” *IEEE Internet Things J.*, vol. 7, no. 6, pp. 4682–4696, Jun. 2020.

[46] P. Radoglou-Grammatikis, P. Sarigiannidis, G. Efstathopoulos, T. Lagkas, G. Fragulis, and A. Sarigiannidis, “A self-learning approach for detecting intrusions in healthcare systems,” in *Proc. IEEE Int. Conf. Commun.*, 2021, pp. 1–6.

[47] S. Rani, S. H. Ahmed, R. Talwar, J. Malhotra, and H. Song, “IoMT: A reliable cross layer protocol for Internet of Multimedia Things,” *IEEE Internet Things J.*, vol. 4, no. 3, pp. 832–839, Jun. 2017.

[48] V. Ravi, T. D. Pham, and M. Alazab, “Deep learning-based network intrusion detection system for Internet of Medical Things,” *IEEE Internet Things Mag.*, vol. 6, no. 2, pp. 50–54, Jun. 2023.

[49] T. Saba, “Intrusion detection in smart city hospitals using ensemble classifiers,” in *Proc. 13th Int. Conf. Develop. eSyst. Eng. (DeSE)*, 2020, pp. 418–422.

[50] Y. K. Saheed and M. O. 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*, vol. 9, pp. 161546–161554, 2021.

[51] H. Tauqeer, M. M. Iqbal, A. Ali, S. Zaman, and M. U. Chaudhry, “Cyberattacks detection in IoMT using machine learning techniques,” *J. Comput. Biomed. Inform.*, vol. 4, no. 1, pp. 13–20, 2022.

[52] A. Thakkar and R. Lohiya, “Attack classification of imbalanced intrusion data for IoT network using ensemble learning-based deep neural network,” *IEEE Internet Things J.*, vol. 10, no. 13, pp. 11888–11895, Jul. 2023.

[53] I. Ullah and Q. H. Mahmoud, “A scheme for generating a dataset for anomalous activity detection in IoT networks,” in *Proc. Can. Conf. Artif. Intell.*, 2020, pp. 508–520.

[54] D. Unal, S. Bennbaia, and F. O. Catak, “Machine learning for the security of healthcare systems based on Internet of Things and edge computing,” in *Cybersecurity and Cognitive Science*. Amsterdam, The Netherlands, Elsevier, 2022, pp. 299–320.

[55] A. Vikram, “Anomaly detection in network traffic using unsupervised machine learning approach,” in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, 2020, pp. 476–479.

[56] Y. Wu, L. Nie, S. Wang, Z. Ning, and S. Li, “Intelligent intrusion detection for Internet of Things security: A deep convolutional generative adversarial network-enabled approach,” *IEEE Internet Things J.*, vol. 10, no. 4, pp. 3094–3106, Feb. 2023.

[57] J. Yaacoub et al., "Securing Internet of Medical Things systems: Limitations, issues and recommendations," *Future Gener. Comput. Syst.*, vol. 105, pp. 581–606, Apr. 2020.

[58] F. Yang et al., "Internet-of-Things-Enabled data fusion method for sleep healthcare applications," *IEEE Internet Things J.*, vol. 8, no. 21, pp. 15892–15905, Nov. 2021.

[59] L. Yang, A. Moubayed, and A. Shami, "MTH-IDS: A multitiered hybrid intrusion detection system for Internet of Vehicles," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 616–632, Jan. 2022.

[60] M. Zang, C. Zheng, L. Dittmann, and N. Zilberman, "Towards continuous threat defense: In-network traffic analysis for IoT gateways," *IEEE Internet Things J.*, vol. 11, no. 6, pp. 9244–9257, Mar. 2024.

[61] Z. Zhang, Y. Zhang, D. Guo, L. Yao, and Z. Li, "SecfedNIDS: Robust defense for poisoning attack against federated learning-based network intrusion detection system," *Future Gener. Comput. Syst.*, vol. 134, pp. 154–169, Sep. 2022.

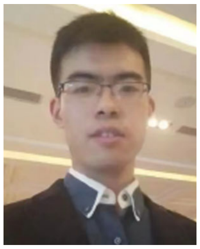
[62] M. Zolanvari, A. Ghubaish, and R. Jain, "ADDAI: Anomaly detection using distributed AI," in *Proc. IEEE Int. Conf. Netw., Sens. Control (ICNSC)*, vol. 1, 2021, pp. 1–6.

Umer Zukaib is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, Wuhan, 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 with the School of Cyber Science and Engineering, Wuhan University, Wuhan, China. His research interests include artificial intelligence, blockchain technology, and high-performance computing.



Chengliang Zheng is currently pursuing the Ph.D. degree in cyberspace security with the School of Cyber Science and Engineering, Wuhan University, Wuhan, China.

His current research interests are blockchain and machine learning.

Mir Hassan (Member, IEEE) is currently pursuing the Ph.D. degree with the University of Trento, Trento, Italy.

His current research interests are federated learning, blockchain, and IoT.

Zhidong Shen received the Ph.D. degree in computer software and theory from the School of Computer Science, Wuhan University, Wuhan, China, in 2006.

He is currently an Associate Professor with the School of Cyber Science and Engineering, Wuhan University. His research interests include artificial intelligence security, system security, and big data analysis.