ICT Express 12 (2026) 1075–1084

Contents lists available at ScienceDirect

ICT Express

journal homepage: www.elsevier.com/locate/icte

A robust cross-domain IDS using BiGRU-LSTM-attention for medical and industrial IoT security

Afrah Gueriani [a,c], Hamza Kheddar [a,c], Ahmed Cherif Mazari [a], Mohamed Chahine Ghanem [b,c,*]

[a] Laboratory of Advanced Electronic Systems (LSEA), University of Medea, Medea 26000, Algeria
[b] Cybersecurity Institute, Department of Computer Science, University of Liverpool, Liverpool, L693BX, UK
[c] Cyber Security Research Centre, London Metropolitan University, London, N78DB, UK

ARTICLE INFO

Keywords:
Transformers
Intrusion detection
Anomalies detection
Cyber-security
Deep learning
Internet of Medical Things
Industrial Internet of Things

ABSTRACT

The increased Internet of Medical Things (IoMT) and the Industrial Internet of Things (IIoT) interconnectivity has introduced complex cybersecurity challenges, exposing sensitive data, patient safety, and industrial operations to advanced cyber threats. To mitigate these risks, this paper introduces a novel transformer-based intrusion detection system (IDS), termed BiGAT-ID—a hybrid model that combines bidirectional gated recurrent units (BiGRU), long short-term memory (LSTM) networks, and multi-head attention (MHA). The proposed architecture is designed to effectively capture bidirectional temporal dependencies, model sequential patterns, and enhance contextual feature representation. Extensive experiments on two benchmark datasets; CICIoMT2024 (medical IoT) and EdgeIIoTset (industrial IoT); demonstrate the model's cross-domain robustness, achieving detection accuracies of 99.13% and 99.34%, respectively. Additionally, the model exhibits exceptional runtime efficiency, with inference times as low as 0.0002 seconds per instance in IoMT and 0.0001 seconds in IIoT scenarios. Coupled with a low false positive rate, BiGAT-ID proves to be a reliable and efficient IDS for deployment in real-world heterogeneous IoT environments.

1. Introduction

The rapid expansion of Internet of Things (IoT) technologies has significantly transformed various sectors, including healthcare, industry, agriculture, transportation, and smart cities, by offering enhanced services and improving operational efficiency [1]. With the continuous advancement of technology, the number of connected devices has grown substantially. This growth is largely driven by the integration of IoT and its associated communication systems, which now play a crucial role in both everyday applications and critical infrastructures across multiple domains. Within the healthcare sector, a specialized branch of IoT—commonly referred to as the Internet of Medical Things (IoMT), smart healthcare, or Healthcare 4.0, has emerged as a transformative force. IoMT enables continuous, real-time monitoring through connected medical devices, supporting applications such as telemedicine, wearable sensors, and remote diagnostics. These technologies have significantly improved patient care by enabling proactive intervention and seamless healthcare delivery.

In both IIoT and IoMT, the system architecture is typically structured into three primary layers: device layer, network layer, and application layer [2]. The *device layer* (or perception layer) comprises various sensors and smart devices responsible for data collection and environmental sensing. In IoMT, this includes patient monitoring devices, implantable sensors, and telemedicine equipment, while in IIoT it involves sensors from smart homes, vehicles, and industrial equipment [2]. The *network layer* ensures secure and efficient data transmission between devices and platforms using protocols specific to each domain. In both IIoT and IoMT environments, communication relies on lightweight and domain-specific protocols designed to support interoperability, real-time monitoring, and system control. Typical protocols include Modbus, MQTT, and LoRaWAN in industrial applications, as well as Bluetooth low energy (BLE), Zigbee, and Wi-Fi in healthcare systems. These communication technologies play a critical role in enabling seamless data exchange between devices and platforms, facilitating remote monitoring, and supporting the deployment of digital twins for smart factories, smart hospitals, and intelligent building management systems. At the top of the architecture, the *application layer* in both IIoT

* Corresponding author at: Cybersecurity Institute, Department of Computer Science, University of Liverpool, Liverpool, L693BX, UK.
E-mail addresses: gueriani.afrah@univ-medea.dz (A. Gueriani), kheddar.hamza@univ-medea.dz (H. Kheddar), mazari.ahmedcherif@univ-medea.dz (A.C. Mazari), mohamed.chahine.ghanem@liverpool.ac.uk (M.C. Ghanem).

https://doi.org/10.1016/j.icte.2025.08.011
Received 25 June 2025; Received in revised form 17 August 2025; Accepted 28 August 2025
Available online 6 September 2025
2405-9595/© 2025 The Authors. Published by Elsevier B.V. on behalf of The Korean Institute of Communications and Information Sciences. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

and IoMT systems enables intelligent decision-making, automation, and user interaction. In healthcare, this layer supports functionalities such as electronic health records, clinical decision support systems, and remote patient monitoring platforms [3]. Similarly, in industrial environments, the application layer integrates supervisory control and data acquisition (SCADA) systems, human-machine interfaces (HMI), and tools for real-time industrial process monitoring [4]. Across both domains, this layer transforms raw sensor data into actionable insights, supporting the deployment of digital twins and enabling proactive, data-driven responses to dynamic environments.

1.1. Motivation and research gap

IoMT and IIoT technologies have enhanced interconnectivity and operational efficiency; however, this growing reliance on interconnected networks has also heightened their vulnerability to sophisticated and targeted cyber-attacks. Each layer of the IIoT and IoMT architecture presents specific security challenges and is exposed to various cyber threats. The *device layer* is vulnerable to physical attacks such as node tampering, fake node injection, replay attacks, and eavesdropping, targeting sensor data integrity and device functionality. The *network layer* is a primary target for communication-based attacks, including denial-of-service (DoS), distributed DoS (DDoS), man-in-the-middle (MITM) attacks, and unauthorized access, aiming to disrupt data transmission and network availability [5]. Meanwhile, the *application layer* faces software-level threats such as malware, ransomware, SQL injection, and privacy breaches, which can compromise sensitive data, critical services, and decision-making processes.

Motivated by these critical needs, intrusion detection system (IDS) solutions have been developed for industrial control environments, including control network analysis, protocol analysis, and traffic mining based on machine learning or deep learning (DL) techniques [6]. However, traditional IDS approaches often fall short in addressing the increasing complexity and sophistication of modern cyber-attacks. This limitation is particularly critical in sensitive environments such as healthcare and industry, where security breaches may result in severe consequences, including data leakage, service disruption, operational downtime, and safety risks. These challenges underscore the urgent need for advanced IDS frameworks capable of real-time threat detection and rapid response, in order to safeguard both IoMT and IIoT networks against emerging and evolving threats.

1.2. Our contribution

This study proposes a hybrid DL-based IDS, named bidirectional GRU and Attention-based Transformer for intrusion detection *(BiGAT-ID)*. Within the proposed IDS, IoT traffic is modeled as a *non-stationary* sequence where attack semantics often depend on short transients (e.g., flag flips, rapid port scans) embedded in longer session patterns. To address this, a BiGRU encoder is employed for efficient *past–future context capture*, while an LSTM layer ensures *long-range* attack persistence through stable gradient flow that mitigates vanishing gradients. To enhance discrimination, a multi-head attention (MHA) mechanism is applied, improving detection accuracy and reducing false positives by *focus on salient patterns* within the traffic sequence [7].

The main contributions of this study are summarized as follows:

- To the best of the authors' knowledge, BiGAT-IDS uniquely combines BiGRU, LSTM, and MHA to achieve outstanding intrusion detection performance.
- BiGAT-IDS detects attacks across IIoT and IoMT domains, proving its versatility and strong generalization.
- Leave-one-attack-out (LOAO) zero-day strategy testing confirms that BiGAT-IDS detects both signature-based and zero-day threats across distinct attack patterns.
- BiGAT-IDS outperforms state-of-the-art models on both datasets, validating its superior detection accuracy.

2. Related works

Numerous techniques have been developed to combat intrusions in IoMT networks. For example, the work in [8] introduces a CNN-based IDS for IoMT networks that outperforms traditional ML methods in accuracy and efficiency, demonstrating strong robustness and real-world applicability for healthcare cybersecurity. Moreover, both [9,10] utilize deep neural networks (DNN) for intrusion detection but differ in focus and design. It integrates DNN within a federated learning framework for decentralized IoT, emphasizing scalability and privacy, while [10] employs a self-attention-based DL model for IoMT anomaly detection. Notably, [9] outperforms [10] in terms of F1-score, demonstrating superior effectiveness across the same dataset. LSTM-based approaches form the foundation of several studies. The DAG-LSTM model [19] enhances intrusion detection in IoMT by incorporating feature optimization techniques, achieving an accuracy of 92%. In contrast, the EFL-LSTM model [20] leverages federated learning to enable decentralized detection, reaching a higher accuracy of 97.16%. Both schemes are evaluated solely on the ECU-IoHT benchmark dataset, which limits their ability to detect a wide range of unseen IoT attacks. Similarly, an advanced LSTM-based intrusion detection scheme is proposed in [11] to secure IoMT environments. By learning from diverse equipment and attack patterns, the model effectively captures complex threat behaviors to safeguard sensitive medical data against evolving cybersecurity challenges. However, the scheme has not been evaluated against IIoT-specific attacks. Stacking ensemble models combining LSTM, CNN, and DNN in [21] achieve high performance on the ECU-IoHT dataset for IoMT detection. However, the approach incurs high computational cost and lacks evaluation on IIoT or general IoT datasets, raising concerns about its generalizability and real-world adaptability. Innovative architectures include CKANs, which embed the Kolmogorov-Arnold theorem with self-attention mechanisms to balance accuracy and interpretability in IoMT, though with trade-offs in resource efficiency [22]. Yet, CKANs aim to balance interpretability and accuracy; this trade-off might not hold across all datasets or intrusion types, leading to potential performance drops. Meanwhile, FlowID [12] leverages a hypergraph-based DL framework combined with dual-contrastive self-supervised learning to effectively capture high-order traffic interactions. This design enhances its robustness and generalization across diverse network scenarios and benchmark datasets.

Researchers have shown increasing interest in developing techniques to counter intrusions in IIoT networks compared to IoMT. For example, in [13,14], CNN-LSTM architectures are proposed for IIoT intrusion detection. The first employs LSTM only, while the second incorporates a self-attention layer to enhance accuracy. Both models are evaluated on the Edge-IIoTset dataset, demonstrating strong performance in binary and multiclass classification. The self-attention-based model achieves superior accuracy and low inference time, making it well-suited for real-time and robust IIoT cybersecurity applications. The simplest approach is observed in [23], which explores standard DL architectures such as artificial neural networks (ANN), LSTM, and GRU for attack classification. These models are applied independently and lack attention for contextual focus. A step further, [15] employs a dual recurrent approach by combining Bi-GRU with LSTM to improve the modeling of sequential dependencies in network traffic. This fusion enhances the learning of long-range temporal features while maintaining a focus on recurrent architectures. More advanced still, [24] proposes the Magru-IDS model, which integrates GRU with MHA to refine the extraction of relevant temporal features. This addition of attention mechanisms improves the model's ability to selectively focus on critical time-dependent patterns. However, the model has limited sequential dependency modeling capability. Another strategy, like [25] introduces a scheme that augments a conventional CNN by capturing global feature dependencies through self-attention. This work also applies data cleaning and mutual information-based feature

Table 1
Comparative summary of state-of-the-art IDS solutions for IoMT and IIoT networks.

	Ref	Dataset	DL method	BP (%)	Limitations	Employing		
						Bi/GRU	LSTM	MHA
Medical domain	[8]	CICIoMT2024	CNN	Acc = 99.00 F1 = 98.00	Absence of comparison with prior works, and evaluation conducted on a single dataset	✗	✗	✗
	[9]	CICIoMT2024	DNN	Acc = 99.56 Pr = 94.59	Lacks temporal awareness and generalizes poorly on sequences, and use imbalanced dataset	✗	✗	✗
	[10]	CICIoMT2024	Att-DNN	Pr = 84.43 F1 = 91.02	Protocol diversity and data imbalance hinder robust cross-domain model generalization.	✗	✗	✓
	[11]	CICIoMT2024	L2D2	Acc = 98.00, F1 = 98.00	Not evaluated for delays, zero-day threats, or IIoT generalization.	✗	✓	✗
	[12]	CIC-IOMT2024	CNN-GNN	Acc = 95.27 F1 = 90.36	High complexity limits deployment on resource-constrained devices.	✓	✗	✗
Industrial domain	[9]	CICIoT2023	DNN	Acc = 99.09 Pr = 91.56	Lacks temporal awareness and generalizes poorly on sequences, and use imbalanced dataset	✗	✗	✗
	[10]	MQTT-IoT-IDS	Att-DNN	Pr = 92.14 F1 = 95.53	Protocol diversity and data imbalance hinder robust cross-domain model generalization.	✗	✗	✓
	[13]	EdgeIIoTset	CNN-LSTM	Acc = 98.68	Evaluation based on limited metrics and a single dataset affects overall generalizability.	✗	✓	✗
	[14]	EdgeIIoTset	LSTM-CNN-Att	F1 = 99.04, FPR = 0.002	Lacks mechanism to emphasize key temporal patterns and cross-domain IIoT–IoMT applicability.	✗	✓	✓
	[15]	Edge-IIoTset	BiGRU-LSTM	Acc = 98.32, FPR = 0.046	Lacks attention mechanism to highlight important temporal features and patterns.	✓	✓	✗
	[16]	EdgeIIoTset	CNN-GRU	Acc = 98.70 FPR = 0.7	Valid for single dataset, no advanced comparison, and high FPR.	✓	✗	✗
	[17]	EdgeIIoTset	CNN-LSTM-GRU	Acc = 97.44	Limited metrics, single unbalanced dataset, no dynamic focus on temporal patterns.	✓	✓	✗
	[18]	Edge-IIoTset	BiGRU-CNN	Acc = 94.7 Pr = 94.8	Limits capturing long-range dependencies, weakening performance on extended or noisy sequences.	✓	✗	✓

Abbreviations: Best performance (BP); Attention (Att).

filtering, enhancing the model's feature selection without adding sequence modeling. However, the scheme suffer from weak temporal sequence learning ability. Increasing in complexity. Similarly, in [16], which combines CNN and GRU. The CNN component captures spatial patterns, while GRU processes temporal sequences, allowing the model to handle both static and dynamic data characteristics effectively. A more complex hybrid strategy is presented in [17], where CNN, LSTM, and GRU are combined into a triple-ensemble model. This structure leverages the individual strengths of each DL block, spatial, sequential, and gated temporal processing, resulting in a more expressive and robust intrusion detection framework. Finally, the most sophisticated architecture appears in [18], which integrates self-attention mechanisms, BiGRU, and Inception-CNN. Beyond model complexity, this work incorporates advanced data handling techniques, including mixed sampling, denoising, and feature selection based on Pearson correlation and Random Forest. According to Table 1, many existing IDS models demonstrate strong performance but suffer key limitations. Most fail to integrate BiGRU, LSTM, and MHA in a unified framework, reducing their capacity to capture comprehensive temporal and contextual features. Additionally, many are domain-specific, limiting cross-environment generalization between IoMT and IIoT systems. Crucial deployment aspects such as inference time and runtime efficiency are often neglected, undermining real-world applicability. Finally, explainability is rarely addressed, despite its importance for transparency and trust in critical settings like healthcare. These limitations motivate the development of more holistic, generalizable, and interpretable IDS frameworks.

3. Proposed methodology

The proposed model synergistically combines BiGRU for capturing forward–backward dependencies, LSTM for deep sequential feature extraction, and MHA for extracting long dependencies salient temporal features. Unlike existing models evaluated on single-domain datasets with shallow temporal encoding, BiGAT-ID enables robust intrusion detection across both IoMT and IIoT environments. This composite architecture enhances model interpretability, minimizes false alarms, and maintains low inference latency, ensuring scalability and real-time applicability in heterogeneous operational settings. The following subsections detail the preprocessing pipeline and architectural components. The architecture of the proposed BiGAT-ID, illustrated in Fig. 1, is organized into two main components: the data preprocessing pipeline and the model architecture.

3.1. Pre-processing

The preprocessing stage plays a crucial role in converting raw, heterogeneous network traffic data into a structured format suitable for DL-based IDS. The process typically involves several essential steps:

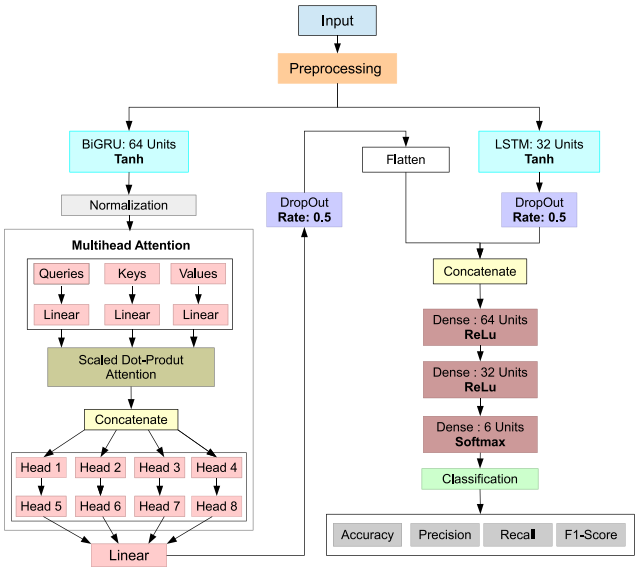


Fig. 1. The proposed BiGAT-IDS model.

- **Data cleaning and label encoding:** Dataset structure and labels are analyzed; categorical fields are numerically encoded using `LabelEncoder` for model compatibility.
- **Feature extraction and reshaping:** Selected features are reshaped into 3D matrices to align with sequential model input and capture temporal dependencies.
- **Class imbalance mitigation:** Random over sampler (RoS) addresses IoMT sparsity by duplicating existing minority samples, while SMOTE enhances IIoT class balance by synthetically generating new minority samples through feature-space interpolation, improving generalization compared to simple duplication [26]. Focal loss was further incorporated during training to down-weight well-classified examples and emphasize hard and minority-class samples, thereby improving sensitivity to subtle, rare attack patterns [27].
- **Dataset splitting:** Data is split into 80% training and 20% testing sets using stratified sampling via `train_test_split` for balance.
- **Label transformation:** Response labels are one-hot encoded using `to_categorical` to support categorical cross-entropy loss in multiclass classification.

3.2. Model architecture

The BiGAT-ID model processes sequential data shaped as (83,1) for IoMT and (60,1) for IIoT. Its dual-branch architecture (Fig. 1) captures complementary features. Table 2 outlines the model structure and hyperparameters.

- Branch 1: BiGRU with MHA: This branch uses a BiGRU layer (64 units) followed by Layer Normalization, multi-head attention (8 heads, 64 key dim), and Dropout (0.5). It captures bidirectional dependencies and highlights salient temporal patterns. Output is flattened for fusion.

- Branch 2: LSTM with Dropout: A single LSTM layer (32 units, `return_sequences=False`) extracts condensed temporal features, followed by Dropout (0.5). This setup ensures a compact embedding of input sequences suitable for classification while reducing overfitting through neuron deactivation during training.

- Fusion and classification: Outputs from both branches are concatenated and passed through Dense layers (64, 32 units with ReLU). A final Dense layer (6 units, softmax) generates class probabilities for multiclass classification, leveraging learned features from both processing paths.

Table 2
BiGAT-ID model summary and hyperparameters.

	DL Layer	Unit	Output Shape	Connected to
1	Input	–	IoMT: (None, 83, 1) IIoT:(None, 60, 1)	–
2	BiGRU	64	(None, 83, 128)	Input_Layer[0][0]
3	LayerNorm.	–	(None, 83, 128)	BiGRU
4	MHA	–	(None, 83, 128)	LayerNorm.
5	Dropout_1	–	(None, 83, 128)	MHA (8, 64)
6	Flatten	–	(None, 10624)	Dropout_1
7	LSTM	32	(None, 32)	Input_Layer[0][0]
8	Dropout_2	–	(None, 32)	LSTM
9	Concatenate	–	(None, 10656)	Flatten, Dropout_2
10	Dense	–	(None, 64)	Concatenate
11	Dense	–	(None, 32)	Dense
12	Dense	–	(None, 6)	Dense
Total parameters:			978,470	
Trainable parameters:			978,470	
Non-trainable parameters:			0	

- Model compilation and validation: The model is compiled with `Adam` optimizer and categorical cross-entropy loss. Performance is evaluated using accuracy, FPR, precision, recall, F1-score, and inference time [28], ensuring robust validation of classification effectiveness.

4. Experimentation

4.1. Datasets

This study employs two publicly available benchmark datasets to evaluate the proposed BiGAT-ID model, summarized in Table 3.

1. **CICIoMT2024**[1]: Designed to strengthen IoMT security research, this dataset includes 18 attack types grouped into five (5) categories, captured from a testbed of 40 devices (25 physical and 15 simulated). It features key IoMT protocols such as Wi-Fi, MQTT, and Bluetooth [29].
2. **Edge-IIoTset**[2]: Collected from a 7-layer IIoT testbed with 10 smart devices, this dataset contains 14 attack types across six (6) categories. It focuses on IIoT-specific traffic and protocols, using 61 selected features out of 1,176 [30].
3. **TON_IoT**[3]: Encompasses heterogeneous data sources derived from telemetry records of IoT/IIoT services, operating system logs, and network traffic within an IoT environment. It comprises 44 features and incorporates benign, and nine (9) distinct categories of cyberattacks, namely scanning, DoS, DDoS, ransomware, backdoor, data injection, XSS, password cracking, and MITM; executed against various IoT and IIoT sensors across the IIoT network, captured from a testbed of more than 10 devices. This dataset is publicly accessible through the TON_IoT repository [31].

4.2. Obtained results

- Accuracy and loss graph: Fig. 2 illustrates the training and validation curves for the multiclass attack detection model. Subfigure (a) depicts the accuracy trends, while subfigure (b) shows the loss trajectories across epochs for both datasets.

- **Accuracy curves** (Fig. 2(a)): For the CICIoMT2024 dataset, the training accuracy rapidly approaches 100%, and the validation accuracy follows closely, stabilizing at 99.13%, which indicates excellent generalization capability. For the EdgeIIoTset dataset, both training and

[1] https://www.unb.ca/cic/datasets/tabular-iot-attack-2024.html.
[2] https://www.kaggle.com/datasets/cnrieiit/mqttset.
[3] https://www.kaggle.com/datasets/alaaelmor/ton-iot-train-test-network.

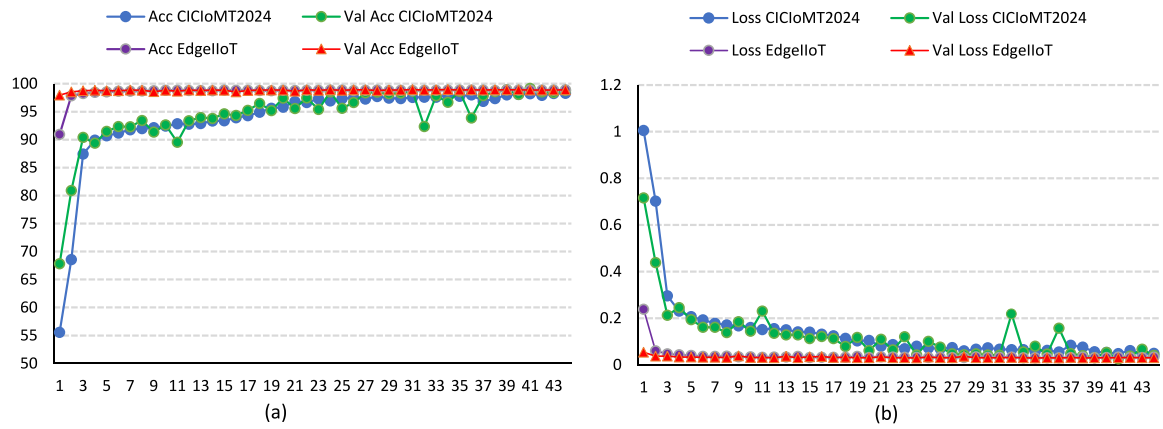


Fig. 2. Accuracy and loss metrics for the proposed BiGAT-ID model across both datasets. (a): Accuracy, (b): Loss.

Table 3
Datasets utilized in the study, along with their respective attack categories, types, and the number of samples.

Data	Category	Samples	Attack types
CICIoMT2024	Normal	32 620	–
	DDoS UDP	2576	DDoS SYN, DDoS TCP, DDoS ICMP, DDoS UDP
	DoS UDP	3115	DoS SYN, DoS TCP, DoS ICMP, DoS UDP
	MITM	1053	ARP Spoofing
	MQTT	953	Malformed Data, DoS Connected Flood, DDoS Publish Flood, DoS Publish Flood
	Recon	8321	Ping Sweep, Recon VulScan, OS Scan, Port Scan
EdgeIIoTset	Normal	24 301	–
	DDoS	49 396	DDoS_UDP, DDoS_TCP, DDoS_ICMP, DDoS_HTTP, DDoS_HTTP
	Info gathering	21 148	Port Scanning, Fingerprinting, Vulnerability_Scanner
	MITM	1214	DNS Spoofing, ARP Spoofing
	Injection	30 632	SQL_Injection, XSS, UpLoading
	Malware	31 109	Ransomeware, Backdoor, Password

Table 4
Classification report for the proposed BiGAT-ID model on CICIoMT2024 and EdgeIIoTset datasets in percentage (%).

CICIoMT2024 dataset				EdgeIIoTset dataset			
Attack type	Pr	Rc	F1	Attack type	Pr	Rc	F1
Normal	98	98	98	Normal	100	100	100
DDoS UDP	100	100	100	DDoS	98	96	97
DoS UDP	97	98	98	Info. gath.	98	99	99
MITM	100	100	100	MITM	100	100	100
MQTT	100	100	100	Injection	97	98	98
Recon.	100	100	100	Malware	100	100	100
Accuracy			99	Accuracy			99
Macro avg	99	99	99	Macro avg	98	98	98
Weighted avg	99	99	99	Weighted avg	99	99	99

Abbreviation: Validation data (VD).

validation accuracy curves follow a consistently upward trajectory with minor oscillations, achieving a peak validation accuracy of 99.34%. These results confirm the model's strong classification ability in both medical and industrial IoT environments.

- **Loss curves** (Fig. 2(b)): The training and validation loss curves for CICIoMT2024 demonstrate a steady decline, converging to a minimal loss value of 0.0257%. A similar downward trend is observed for the EdgeIIoTset dataset, where the final validation loss reaches 0.0158%. The consistent reduction in loss across both domains reflects effective learning and minimal overfitting.

- Classification report analysis: Table 4 presents a detailed classification report for the proposed BiGAT-ID model on both CICIoMT2024 and EdgeIIoTset datasets, reflecting its robust performance across multiple cyberattack classes.

- **IoMT domain:** BiGAT-ID achieves 100% F1-scores in 4 of 6 classes, and 98% in others. With 39,144 validated samples, the model shows excellent precision and recall, effectively distinguishing normal and malicious IoMT traffic.
- **IIoT domain:** On 59,276 validated samples, the model scores 100% F1 in MITM and Malware, and 97%–99% in other attacks. BiGAT-ID maintains 99% accuracy, confirming its robust detection capability across diverse IIoT attack types.

Macro and weighted F1-scores between 98%–99% confirm BiGAT-ID's robustness against class imbalance and diverse attack types.

- Confusion Matrix: Fig. 3 presents the normalized confusion matrices for the CICIoMT2024 and EdgeIIoTset datasets, shown in subfigures (a) and (b), respectively. These matrices provide detailed insights into the classification performance of the BiGAT-ID model across all classes.

- In Fig. 3(a), which presents results for the CICIoMT2024 dataset, the model achieves high accuracy across all attack types. Benign traffic attains a TPR of 98% with only 2% misclassified as MQTT, indicating strong separation between benign and malicious flows. MITM achieves a perfect TPR of 100%, reflecting excellent pattern capture. MQTT records a TPR of 97% but 2% are misclassified as benign, raising concern over false negatives and suggesting the need for improved feature extraction for this attack type. Recon, DDoS, and DoS all achieve a perfect TPR of 100%, showing error-free detection.
- In Fig. 3(b), which presents results for the EdgeIIoTset dataset, the model maintains strong classification performance. Normal traffic achieves a TPR of 100% with no misclassifications, indicating highly reliable benign traffic detection. The DDoS attack is classified with 99% accuracy, with 1% misclassified into the injection class. The injection class shows a slight drop in precision, with 3% of instances misclassified as DDoS, resulting in 97% accuracy. This minor weakness may stem from shared characteristics such as abnormal request rates between injection and volumetric attacks like DDoS. MITM, information gathering, and malware are classified with perfect accuracy of 100%, underscoring the model's precision in detecting critical malicious traffic patterns and confirming its robustness.

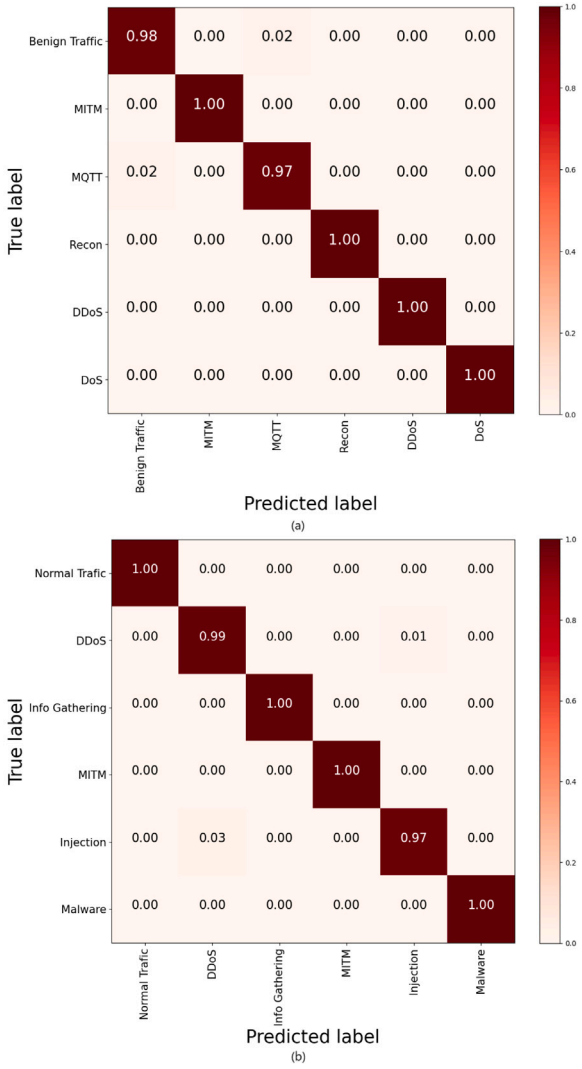


Fig. 3. The confusion matrix of the BiGAT-ID model: (a) For the CICIoMT2024 dataset; (b) For the EdgeIIoTset dataset.

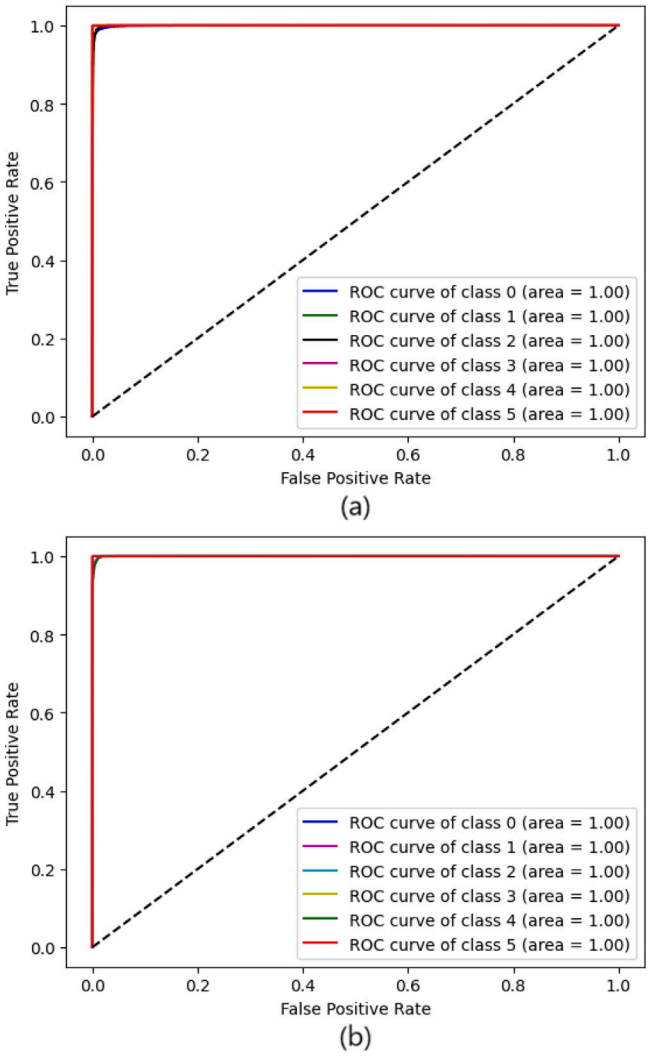


Fig. 4. The ROC curves of the proposed BiGAT-ID model. (a) For the CICIoMT2024 dataset, (b) For the EdgeIIoTset dataset.

The minimal confusion observed between certain traffic types reflects realistic challenges in intrusion detection while still showcasing high generalization and discriminative power across diverse IoT/IIoT attack surfaces, with confusion in CICIoMT2024 limited to benign–MQTT cases and in EdgeIIoTset limited to the injection class. These confusions are minimal and acceptable given the complexity of multi-class classification in a medical IoT environment. The BiGAT-ID model demonstrates highly accurate detection on CICIoMT2024, especially for MITM, Recon, DDoS, and DoS, and performs excellently across the EdgeIIoTset dataset.

Fig. 4 illustrates the ROC curves, derived from the confusion matrix, for multi-class classification performance on two datasets. In both subfigures, the curves show almost perfect separation for all six attack classes, tightly concentrated in the top-left corner, with each class achieving an AUC of 1.00. In Fig. 4(a) for the CICIoMT2024 dataset, this performance reflects ideal sensitivity and specificity across all classes, underscoring the model's robustness in complex medical IoT scenarios where precise detection of various attack types or system states is critical. The sharp curve rise toward the top-left further indicates an extremely low false positive rate (FPRs) even at high TPRs. In Fig. 4(b), for the Edge-IIoT dataset, the uniform ROC curves across classes confirm strong generalization and the absence of class bias, which is crucial in real-world intrusion detection. These results highlight both the architecture's discriminative strength and its ability to accurately identify diverse intrusion types.

To address potential overfitting risks associated with oversampling minority classes, multiple mitigation strategies were implemented. The datasets chosen were accurately labeled and consistently structured in terms of features, while rigorous preprocessing, tailored to each dataset, was applied, including data cleaning and class imbalance handling as described in Section 3. Furthermore, a learning rate of 10^{-3} and a dropout rate of 0.5 were adopted to enhance generalization. Collectively, these measures ensured stable and robust performance, with no observable signs of overfitting throughout the experiments.

- Inference time: Inference time is vital for real-time IDS deployment. BiGAT-ID achieves impressive efficiency with 0.0002 s per instance in IoMT and 0.0001 s in IIoT environments. These results ensure rapid threat detection without latency, which is critical in healthcare systems to protect patient safety and in industrial systems to prevent costly disruptions. The low inference times highlight the model's suitability for real-world, resource-constrained settings where timely intrusion response is essential for operational continuity and data security.

Table 5
Performance of different variants of the proposed models in multiclass classification, with the endices indicating the hyperparameter values of each block.

	Model	Before balancing data					
		CICIoMT2024			EdgeIIoTset		
		Acc. (%)	Loss (%)	FPR (%)	Acc. (%)	Loss (%)	FPR (%)
#1	$BiGRU_{64}+MHA_8$	94.60	0.1804	0.0406	98.87	0.0364	0.0016
#2	$LSTM_{32}+MHA_8$	94.61	0.1884	0.038	98.33	0.0600	0.0023
#3	$BiGRU_{64}$-($LSTM_{32}+MHA_8$)	95.30	0.1665	0.0243	99.01	0.0284	0.0016
#4	**($BiGRU_{64}+MHA_8$)-$LSTM_{32}$**	96.19	0.1368	0.0211	99.04	0.0262	0.0013
#5	($MHA_8+BiGRU_{64}$)-$LSTM_{32}$	86.12	0.4893	0.0983	97.57	0.0841	0.0024
#6	$BiGRU_{64}$-(MHA_8+LSTM_{32})	95.88	0.1400	0.0200	99.11	0.0227	0.0013
#7	($BiGRU_{128}+MHA_8$)-$LSTM_{256}$	94.62	0.1811	0.0456	97.73	0.0734	0.0079
#8	($BiGRU_{64}+MHA_2$)-$LSTM_{32}$	94.21	0.1938	0.0478	98.55	0.0365	0.0038
#9	($BiGRU_{64}+MHA_4$)-$LSTM_{32}$	94.41	0.1823	0.0356	98.85	0.0300	0.0178
#10	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.3 D"	94.92	0.1688	0.0294	98.68	0.0522	0.0020
#11	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.7 D"	93.96	0.1892	0.0486	99.00	0.0278	0.0016
#12	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.2 D"	93.72	0.2068	0.0449	98.94	0.0303	0.0043
		After balancing data					
#1	$BiGRU_{64}+MHA_8$	96.15	0.0828	0.007	99.32	0.0165	0.0013
#2	$LSTM_{32}+MHA_8$	98.41	0.0440	0.0026	98.88	0.0304	0.0023
#3	$BiGRU_{64}$-($LSTM_{32}+MHA_8$)	94.73	0.1100	0.0093	99.12	0.0194	0.0020
#4	**($BiGRU_{64}+MHA_8$)-$LSTM_{32}$**	99.13	0.0257	0.0013	99.34	0.0158	0.0013
#5	($MHA_8+BiGRU_{64}$)-$LSTM_{32}$	65.69	0.7986	0.0625	99.04	0.0260	0.0013
#6	$BiGRU_{64}$-(MHA_8+LSTM_{32})	93.74	0.1295	0.0106	99.32	0.0159	0.0013
#7	($BiGRU_{128}+MHA_8$)-$LSTM_{256}$	90.74	0.2156	0.0323	60.27	1.1627	0.4231
#8	($BiGRU_{64}+MHA_2$)-$LSTM_{32}$	93.54	0.1344	0.0146	99.21	0.0214	0.0016
#9	($BiGRU_{64}+MHA_4$)-$LSTM_{32}$	95.64	0.0944	0.0090	99.26	0.0217	0.0013
#10	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.3 D"	95.49	0.0978	0.0089	99.30	0.0230	0.0013
#11	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.7 D"	94.18	0.1290	0.0106	99.31	0.0166	0.0013
#12	($BiGRU_{64}+MHA_8$)-$LSTM_{32}$ "0.2 D"	94.06	0.1221	0.0623	99.20	0.0211	0.0020

Abbreviations: Dropout (D).

4.3. Zero-day attack assessment

To further evaluate the generalization capability of the proposed model in detecting previously unseen cyber threats, we conducted a LOAO zero-day simulation across the datasets considered in our study. This experimental setup involved completely excluding a specific attack class during the training phase, while retaining all other classes. The proposed model demonstrated strong zero-day detection capabilities across all evaluated datasets. In our experiments, the model achieved a maximum validation accuracy of 98.50% with a corresponding validation loss of 0.04 on the EdgeIIoT dataset. For the CICIoMT2024 dataset, the proposed BiGAT-ID model attained a maximum validation accuracy of 99.23% with a loss of 0.01. On the TON_IoT dataset, the BiGAT-ID model reached a maximum validation accuracy of 98.07% with a loss of 0.06. These findings highlight the model's resilience in detecting unseen attack behaviors and underscore its effectiveness for real-world deployment scenarios where novel attacks are likely to emerge.

4.4. Ablation study

Table 5 presents the comparative performance of ten architectural variants of the proposed model across both CICIoMT2024 and EdgeIIoTset datasets, evaluated before and after data balancing. The core objective of this ablation study is to identify the configuration that maintains high performance when the model is deployed across both the medical (IoMT) and industrial (IIoT) domains, as well as consistency under data imbalance, a frequent characteristic in real-world intrusion detection scenarios.

The results reveal that configuration #4 (`BiGRU+MHA)-LSTM`, denoted as BiGAT-ID, consistently outperforms other variants. Specifically, before data balancing, it achieves an accuracy of **96.19%** (CICIoMT2024) and **99.04%** (EdgeIIoTset) with corresponding FPRs of **0.0211%** and **0.0013%**, respectively. After balancing, the same model further improves, reaching **99.13%** accuracy on CICIoMT2024 and **99.34%** on EdgeIIoTset, while maintaining an exceptionally low FPR of **0.0013%** for both datasets. Although other configurations, such as #3 (BiGRU-LSTM-MHA) and #6 (BiGRU-(MHA+LSTM)), also show strong results, they exhibit either higher FPRs or less stability under balancing conditions. For instance, configuration #3 suffers a noticeable drop in performance on CICIoMT2024 after balancing (from **95.30%** to **94.73%**) and shows a higher FPR of **0.0093%**, making it less optimal for deployment in sensitive domains like healthcare.

Furthermore, additional experiments (cases #7–#12) explored the impact of varying attention heads, dropout rates, and unit configurations. We empirically evaluated configurations with 2, 4, and 8 heads, finding that 8 heads offered the best balance between accuracy and computational efficiency. Increasing beyond this point yielded no significant gains but increased complexity. Dropout rates of 0.2, 0.3, 0.5, and 0.7 were tested, with 0.5 providing optimal regularization, effectively mitigating overfitting without inducing underfitting. Similarly, experiments with 64, 128, and 256 units showed that setting the LSTM to 32 units and the BiGRU to 64 units resulted in a balanced architecture that sustained strong performance while improving training efficiency.

4.5. BiGAT-ID explainability

The Shapley additive explanations (SHAP) summary plot in Fig. 5 illustrates the average contribution of each feature to the model's output across IIoT and IoMT datasets. Feature importance is reflected by bar length, while colors denote class-specific impacts (class 0–5). The X-axis represents the mean SHAP magnitude, and the Y-axis lists the model features [32].

In Fig. 5(a), feature 6 *(Timestamp)* emerges as the most influential, with a SHAP value magnitude around 10, strongly impacting class 0 and class 2. feature 72 *(Bwd Init Win Bytes)* is particularly relevant for class 4, while feature 20 *(Flow Bytes/s)* and feature 4 *(Dst Port)* show moderate contributions across several classes. Other features, such as feature 30 *(Fwd IAT Min)*, feature 53 *(ACK Flag Count)*, and feature 43 *(Bwd Packets/s)*, provide smaller yet meaningful support, serving as secondary discriminators. The remaining features contribute

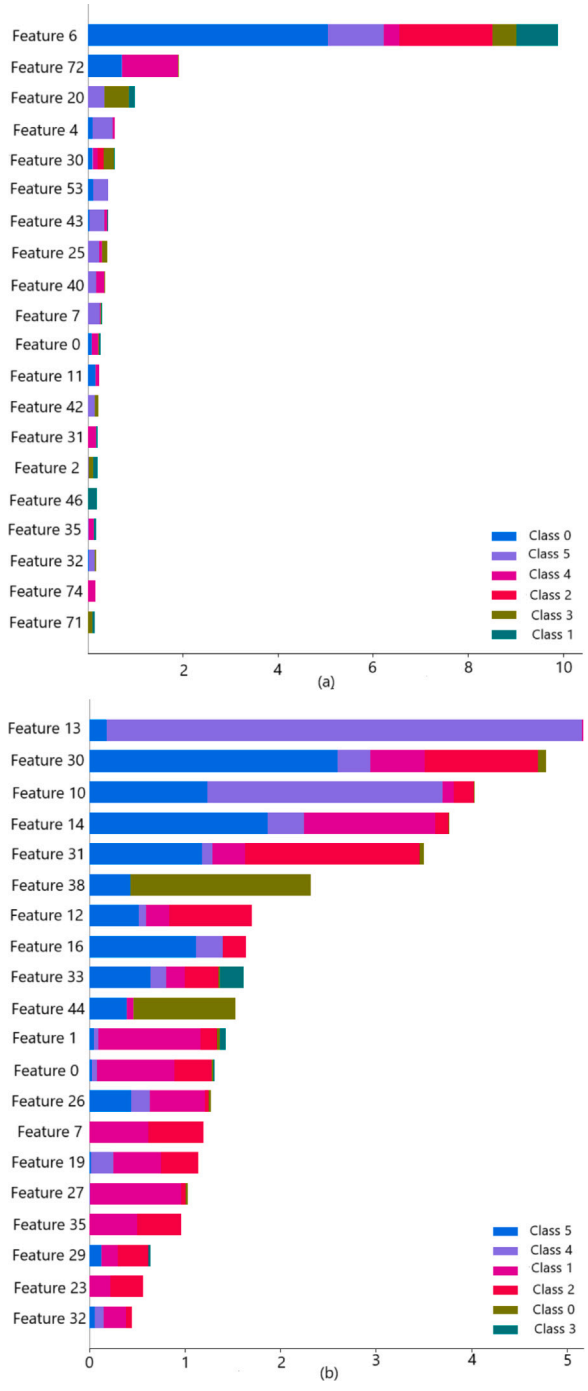


Fig. 5. Mean SAHAP value average impact on model output magnitude. (a): For the CICIoMT2024 dataset; (b): For the EdgeIIoT dataset.

minor but cumulative effects, likely refining predictions and reducing false positives through added contextual information.

In Fig. 5(b), feature 13 *(http.request.uri.query)* emerges as the most influential, particularly for class 4. Several features, such as feature 30 *(tcp.len)*, feature 10 *(icmp.unused)*, feature 14 *(http.request.method)*, and feature 31 *(tcp.options)*, also contribute significantly across multiple classes, while features 38 *(dns.qry.name)* and 44 *(dns.retransmit_request_in)* are dominant for benign traffic (class 0). Moderate features (e.g., *frame.time*, *ip.src_host*, *tcp.dstport*) strengthen predictions, whereas lower-impact features (e.g., *tcp. flags.ack*, *udp.port*) play supportive roles. Globally, the SHAP analysis highlights the interpretability of BiGAT-ID, confirming its ability to capture both global and class-specific patterns across CICIoMT2024 and EdgeIIoT datasets, without bias toward any single class.

4.6. Comparison with state-of-the-art

As summarized in Table 6, several recent IDS models demonstrate only moderate performance and are evaluated on a single dataset, limiting their generalizability. For instance, L2D2 [11] achieves 98% accuracy on CICIoMT2024, but omits loss, FPR, and inference time. Similarly, BiGRU–LSTM in [15] reaches 98.32% accuracy on EdgeIIoT, with no F1-score, loss, or runtime reported. The DNN model from [30] records an even lower accuracy of 96.01%, while providing no supporting metrics. Likewise, the XGBoost model in [33] achieves only 95.01% on CICIoMT2024 and lacks any further evaluation criteria. The absence of FPR, inference time, and loss in all these studies limits their applicability in real-time, resource-constrained environments such as healthcare and industrial IoT, where both detection quality and operational efficiency are critical. On the other hand, three notable works—[9,10], and [34]—conduct cross-domain testing across datasets representing different IoTcontexts. However, their generalization capability remains limited. For example, [34] achieves 99.88% on CICIoMT2024 but drops drastically to 33.30% on EdgeIIoT. Similarly, [9,10] report imbalanced precision–recall and omit runtime metrics, hindering deployment potential.

In contrast, the proposed BiGAT-ID model offers consistent and robust performance across domains, achieving up to 99.13% accuracy on CICIoMT2024 and 99.34% on EdgeIIoT, with exceptionally low FPRs (0.0013%), and inference times as low as 0.0002 s and 0.0001 s per instance, respectively. Notably, it is the only model to report a full set of evaluation metrics with stable performance across domains, highlighting its readiness for deployment in diverse and dynamic IoT environments.

To validate the generalizability of the proposed BiGAT-ID model, experiments were extended to the TON_IoT dataset, which comprises 10 classes of diverse IoT/IIoT attack types with varying traffic characteristics. Two preprocessing settings were applied: SMOTE with RoS and focal loss. As shown in Table 6, the SMOTE+RoS configuration achieved the best results on TON_IoT, with 98.67% accuracy, 99.36% precision, 98.94% recall, and 99.15% F1-score, while focal loss yielded 98.47% accuracy, 98.63% precision, 98.47% recall, and 98.51% F1-score. These results demonstrate the proposed BiGAT-ID model's robustness across datasets with varying characteristics and imbalance levels, showing superior performance on IIoT and IoMT datasets and maintaining competitive results on a general IoT dataset in terms of multiple DL performance metrics and inference time.

5. Conclusion

This paper introduced BiGAT-ID model, designed for intrusion detection across heterogeneous IoT environments. Validated on medical and industrial datasets, the model demonstrated consistent and high performance, aided by tailored balancing strategies to address class imbalance. Unlike previous approaches, BiGAT-ID provides a complete evaluation, including precision, latency, and FPR, confirming its readiness for real-time deployment in critical infrastructures. The study also emphasizes the importance of cross-domain generalization, an area where many existing models struggle. BiGAT-ID addresses this through architectural synergy between temporal modeling and attention mechanisms, resulting in stable outcomes across diverse IoT contexts. Given the limited availability of labeled IoT data, particularly in IoMT, future work may focus on leveraging transfer learning, enhancing interpretability, and optimizing the architecture for scalable real-world deployment. In particular, fine-tuning small language models could be explored to improve the efficiency of the proposed BiGAT-ID architecture across diverse IoT domains.

Table 6
Performance metrics of the proposed BiGAT-ID model in comparison to state-of-the-art methods for multiclass classification.

Work	Model	Dataset	Acc. (%)	Loss	Pr (%)	Rc (%)	F1 (%)	FPR (%)	Inf time (s/inst)
[11]	L2D2 (LSTM)	CICIoMT2024	98	✗	98	98	98	✗	✗
[15]	BiGRU-LSTM	EdgeIIoT	98.32	✗	98.78	97.22	✗	✗	✗
[30]	DNN	EdgeIIoT	96.01	✗	✗	✗	✗	✗	✗
[33]	XGBoost	CICIoMT2024	95.01	✗	✗	✗	✗	✗	✗
[9]	DNN	CICIoMT2024	99.56	✗	94.59	91.28	92.62	✗	✗
		CICIoT2023	99.09	✗	91.56	66.98	69.86	✗	✗
[10]	Att-DNN	CICIoMT2024	✗	✗	84.43	98.73	91.02	✗	✗
		MQTT-IoT-IDS	✗	✗	92.14	99.17	95.53	✗	✗
[34]	CNN-LSTM-ResNet-SA	CICIoMT2024	99.88	✗	99.89	99.99	99.94	✗	✗
		EdgeIIoMT	33.30	✗	33.31	100	49.97	✗	✗
Our	**BiGRU+MHA-LSTM (SMOTE+RoS)**	**CICIoMT2024**	**99.13**	**0.0257**	**99.13**	**99.13**	**99.13**	**0.0013**	**0.0002**
		EdgeIIoT	**99.34**	**0.0158**	**99.34**	**99.34**	**99.34**	**0.0013**	**0.0001**
		TON_IoT	**98.67**	**0.0476**	**98.68**	**98.67**	**98.67**	**0.0086**	**0.00016**
	BiGRU+MHA-LSTM (Focal loss)	**CICIoMT2024**	**94.98**	**0.0145**	**94.98**	**93.66**	**94.23**	**0.0293**	**0.00026**
		EdgeIIoT	**99.13**	**0.0016**	**99.13**	**99.13**	**99.13**	**0.0013**	**0.00017**
		TON_IoT	**98.47**	**0.0054**	**98.63**	**98.47**	**98.51**	**0.0068**	**0.00015**

CRediT authorship contribution statement

Afrah Gueriani: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hamza Kheddar:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Ahmed Cherif Mazari:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mohamed Chahine Ghanem:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge that the Open Access APC was funded by the University of Liverpool, United Kingdom (Grant No. 100896930).

References

[1] Ali Ghubaish, Tara Salman, Maede Zolanvari, Devrim Unal, Abdulla Al-Ali, Raj Jain, Recent advances in the internet-of-medical-things (IoMT) systems security, IEEE Internet Things J. 8 (11) (2020) 8707–8718.
[2] Hamza Kheddar, Yassine Himeur, Ali Ismail Awad, Deep transfer learning for intrusion detection in industrial control networks: A comprehensive review, J. Netw. Comput. Appl. 220 (2023) 103760.
[3] Hoe Tung Yew, Ming Fung Ng, Soh Zhi Ping, Seng Kheau Chung, Ali Chekima, Jamal A Dargham, IoT based real-time remote patient monitoring system, in: 2020 16th IEEE International Colloquium on Signal Processing & Its Applications, CSPA, IEEE, 2020, pp. 176–179.
[4] Vangelis Malamas, Fotis Chantzis, Thomas K Dasaklis, George Stergiopoulos, Panayiotis Kotzanikolaou, Christos Douligeris, Risk assessment methodologies for the internet of medical things: A survey and comparative appraisal, IEEE Access 9 (2021) 40049–40075.
[5] Mikiyas Alemayehu, Mohamed Chahine Ghanem, Karim Ouazzane, Hamza Kheddar, Marcio J Lacerda, A systematic analysis on the use of AI techniques in industrial IoT DDoS attacks detection, mitigation and prevention, 2025.
[6] Hamza Kheddar, Diana W Dawoud, Ali Ismail Awad, Yassine Himeur, Muhammad Khurram Khan, Reinforcement-learning-based intrusion detection in communication networks: A review, IEEE Commun. Surv. Tutor. (2024).
[7] Hamza Kheddar, Transformers and large language models for efficient intrusion detection systems: A comprehensive survey, Inf. Fusion (2025) 103347.
[8] Alireza Mohammadi, Hosna Ghahramani, Seyyed Amir Asghari, Mehdi Aminian, Securing healthcare with deep learning: A CNN-based model for medical IoT threat detection, in: 2024 19th Iranian Conference on Intelligent Systems, ICIS, IEEE, 2024, pp. 168–173.
[9] Duc Vu-Minh, My Duong-Tran-Tra, Luan Van-Thien, Anh Pham-Nguyen-Hai, Thuat Nguyen-Khanh, Quan Le-Trung, Performance evaluation of decentralized machine learning based network-based intrusion detection system for internet of things, in: 2024 18th International Conference on Advanced Computing and Analytics, ACOMPA, IEEE, 2024, pp. 78–85.
[10] Mireya Lucia Hernandez-Jaimes, Alfonso Martinez-Cruz, Kelsey Alejandra Ramírez-Gutiérrez, Alicia Morales-Reyes, Network traffic inspection to enhance anomaly detection in the Internet of Things using attention-driven deep learning, Integr. (2025) 102398.
[11] Gökhan Akar, Shaaban Sahmoud, Mustafa Onat, Ünal Cavusoglu, Emmanuel Malondo, L2D2: A novel LSTM model for multi-class intrusion detection systems in the era of IoMT, IEEE Access (2025).
[12] Jiajun Zhou, Wentao Fu, Hao Song, Shanqing Yu, Qi Xuan, Xiaoniu Yang, Multi-view correlation-aware network traffic detection on flow hypergraph, 2025, arXiv preprint arXiv:2501.08610.
[13] Amina Khacha, Rafika Saadouni, Yasmine Harbi, Zibouda Aliouat, Hybrid deep learning-based intrusion detection system for industrial internet of things, in: 2022 5th International Symposium on Informatics and Its Applications, ISIA, IEEE, 2022, pp. 1–6.
[14] Afrah Gueriani, Hamza Kheddar, Ahmed Cherif Mazari, Adaptive cyber-attack detection in IIoT using attention-based LSTM-CNN models, in: 2024 International Conference on Telecommunications and Intelligent Systems, ICTIS, IEEE, 2024, pp. 1–6.
[15] Danish Javeed, Tianhan Gao, Muhammad Shahid Saeed, Prabhat Kumar, An intrusion detection system for edge-envisioned smart agriculture in extreme environment, IEEE Internet Things J. (2023).
[16] Rafika Saadouni, Amina Khacha, Yasmine Harbi, Chirihane Gherbi, Saad Harous, Zibouda Aliouat, Secure iiot networks with hybrid cnn-gru model using edge-iiotset, in: 2023 15th International Conference on Innovations in Information Technology, IIT, IEEE, 2023, pp. 150–155.
[17] Dusmurod Kilichev, Dilmurod Turimov, Wooseong Kim, Next–generation intrusion detection for IoT evcs: Integrating cnn, lstm, and gru models, Math. 12 (4) (2024) 571.
[18] Kai Yang, JiaMing Wang, MinJing Li, An improved intrusion detection method for IIoT using attention mechanisms, BiGRU, and inception-CNN, Sci. Rep. 14 (1) (2024) 19339.
[19] Manish Kumar, Changjong Kim, Yongseok Son, Sushil Kumar Singh, Sunggon Kim, Empowering cyberattack identification in IoHT networks with neighborhood component-based improvised long short-term memory, IEEE Internet Things J. (2024).
[20] Manish Kumar, Sunggon Kim, Securing the internet of health things: Embedded federated learning-driven long short-term memory for cyberattack detection, Electron. 13 (17) (2024) 3461.
[21] Easa Alalwany, Bader Alsharif, Yazeed Alotaibi, Abdullah Alfahaid, Imad Mahgoub, Mohammad Ilyas, Stacking ensemble deep learning for real-time intrusion detection in IoMT environments, Sens. 25 (3) (2025) 624.
[22] Zhen Wang, Anazida Zainal, Maheyzah Md Siraj, Fuad A Ghaleb, Xue Hao, Shaoyong Han, An intrusion detection model based on convolutional Kolmogorov-arnold networks, Sci. Rep. 15 (1) (2025) 1917.
[23] Mohammed Mudassir, Devrim Unal, Mohammad Hammoudeh, Farag Azzedin, Detection of botnet attacks against industrial IoT systems by multilayer deep learning approaches, Wirel. Commun. Mob. Comput. 2022 (1) (2022) 2845446.

[24] Safi Ullah, Wadii Boulila, Anis Koubaa, Jawad Ahmad, MAGRU-IDS: A multi-head attention-based gated recurrent unit for intrusion detection in IIoT networks, IEEE Access (2023).
[25] Mohammed S Alshehri, Oumaima Saidani, Fatma S Alrayes, Saadullah Farooq Abbasi, Jawad Ahmad, A self-attention-based deep convolutional neural networks for IIoT networks intrusion detection, IEEE Access (2024).
[26] Alberto Fernández, Salvador Garcia, Francisco Herrera, Nitesh V Chawla, SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary, J. Artificial Intelligence Res. 61 (2018) 863–905.
[27] Yesi Novaria Kunang, Siti Nurmaini, Deris Stiawan, Bhakti Yudho Suprapto, Deep learning with focal loss approach for attacks classification, TELKOMNIKA (Telecommun. Comput. Electron. Control) 19 (4) (2021) 1407–1418.
[28] Afrah Gueriani, Hamza Kheddar, Ahmed Cherif Mazari, Enhancing IoT security with CNN and LSTM-based intrusion detection systems, in: 2024 6th International Conference on Pattern Analysis and Intelligent Systems, PAIS, IEEE, 2024, pp. 1–7.
[29] Sajjad Dadkhah, Euclides Carlos Pinto Neto, Raphael Ferreira, Reginald Chukwuka Molokwu, Somayeh Sadeghi, Ali Ghorbani, Ciciomt2024: Attack vectors in healthcare devices-a multi-protocol dataset for assessing iomt device security, 2024.
[30] Mohamed Amine Ferrag, Othmane Friha, Djallel Hamouda, Leandros Maglaras, Helge Janicke, Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning, IEEE Access 10 (2022) 40281–40306.
[31] Abdullah Alsaedi, Nour Moustafa, Zahir Tari, Abdun Mahmood, Adnan Anwar, TON_IoT telemetry dataset: A new generation dataset of IoT and iIoT for data-driven intrusion detection systems, Ieee Access 8 (2020) 165130–165150.
[32] Scott M. Lundberg, Su-In Lee, A unified approach to interpreting model predictions, Adv. Neural Inf. Process. Syst. 30 (2017).
[33] Fatima Sohail, Muhammad Asim Mukhtar Bhatti, Muhammad Awais, Aamna Iqtidar, Explainable boosting ensemble methods for intrusion detection in internet of medical things (IoMT) applications, in: 2024 4th International Conference on Digital Futures and Transformative Technologies, ICoDT2, IEEE, 2024, pp. 1–8.
[34] Tinshu Sasi, Arash Habibi Lashkari, Rongxing Lu, Pulei Xiong, Shahrear Iqbal, An efficient self attention-based 1D-CNN-LSTM network for IoT attack detection and identification using network traffic, J. Inf. Intell. (2024).