IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 September 2023, accepted 9 October 2023, date of publication 16 October 2023, date of current version 20 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3324657

RESEARCH ARTICLE

MAGRU-IDS: A Multi-Head Attention-Based Gated Recurrent Unit for Intrusion Detection in IIoT Networks

SAFI ULLAH[1,2], WADII BOULILA[1,3], (Senior Member, IEEE), ANIS KOUBÂA[1], AND JAWAD AHMAD[4], (Senior Member, IEEE)

[1]Robotics and Internet of Things Laboratory, Prince Sultan University, Riyadh 12435, Saudi Arabia
[2]Department of Computer Science, Quaid-i-Azam University, Islamabad 44000, Pakistan
[3]RIADI Laboratory, National School of Computer Science, University of Manouba, Manouba 2010, Tunisia
[4]School of Computing, Edinburgh Napier University, EH10 5DT Edinburgh, U.K.

Corresponding author: Wadii Boulila (wboulila@psu.edu.sa)

ABSTRACT The increasing prevalence of the Industrial Internet of Things (IIoT) in industrial environments amplifies the potential for security breaches and compromises. To monitor IIoT networks, intrusion detection systems (IDS) have been introduced to detect malicious activities within the network flow, in which machine learning (ML) and deep learning (DL) play an important role. However, existing IDSs face challenges during training when dealing with imbalanced training data and a higher number of classes. These issues can significantly reduce the IDS's performance and may result in missed network attacks, especially those with fewer training samples. To address these challenges, this paper introduces a multi-head attention-based gated recurrent unit (MAGRU) that scrutinizes IIoT network traffic to detect malicious activities. In the proposed model, the multi-head attention (MA) has the ability to enhance the learning capability of the model to handle limited sample classes. The gated recurrent unit (GRU) is employed for the detection of IIoT network behavior. The proposed MAGRU is evaluated using two publicly available datasets, namely Edge-IIoTset and MQTTset. To validate the performance of the proposed MAGRU, various ML and DL models were implemented and compared against MAGRU using the same dataset. The proposed model outperformed the other models, achieving an average precision, recall, F1-score, and accuracy of 99.62%, 99.67%, 99.64%, and 99.97%, respectively, for the aforementioned datasets. These results demonstrate optimal performance in the detection of intrusions in IIoT networks.

INDEX TERMS Deep learning, gated recurrent unit, industrial Internet of Things, intrusion detection, multi-head attention.

I. INTRODUCTION

The Industrial Internet of Things (IIoT) is a network of interconnected devices, sensors, and equipment used in industrial settings to collect, exchange, and analyze data. IIoT aims to enhance operational efficiency, enable predictive maintenance, optimize processes, and improve overall productivity in industries such as manufacturing, energy, transportation, and healthcare [1], [2]. Several IIoT devices are operating with sensitive data, and the growing prevalence of IIoT in industrial environments also amplifies the potential for security breaches and compromises [3], [4], [5]. This is primarily due to the insufficient security features found in various IIoT devices and systems, including weak authentication and a lack of regular security updates [6], [7], [8], [9]. As a result, these vulnerabilities expose them to exploitation by cyber attackers. Fig 1 presents a cyberattack scenario in which the botnet is used to launch a DDoS attack on an IIoT network, specifically targeting industrial servers.

An intrusion detection system (IDS) is a sophisticated software employed to scrutinize network traffic and detect malevolent activities within the network [10], [11], [12], [13], [14], [15], [16], [17]. Its primary purpose is to identify and raise alarms on any potentially malicious

The associate editor coordinating the review of this manuscript and approving it for publication was Vicente Alarcon-Aquino.

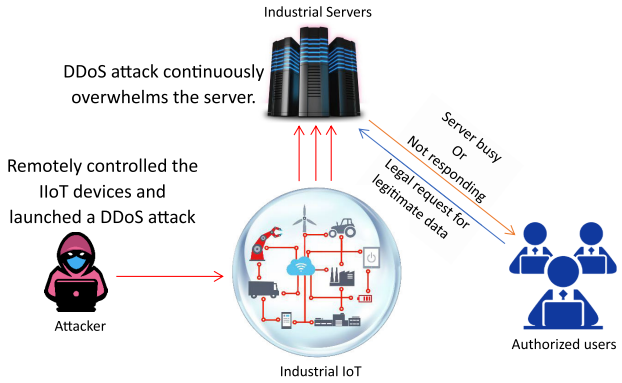


FIGURE 1. A DDoS attack scenario on an industrial IoT network.

or unauthorized behavior occurring within the network [18], [19]. In recent years, machine learning (ML) and deep learning (DL) have emerged as indispensable tools in the field of IDS, significantly enhancing the ability to detect malicious activities within networks [20], [21], [22], [23], [24]. These ML and DL-based IDSs possess generalization capabilities, enabling them to identify previously unseen forms of malicious data present in the network [25], [26]. The existing IDSs demonstrate effective performance when dealing with a limited number of classes and a relatively balanced distribution of training data. However, challenges arise as the number of classes increases and training data becomes imbalanced. The data imbalance poses a problem during training as the model tends to prioritize high-instance classes while paying less attention to low-instance classes. As a consequence, the IDS becomes proficient at detecting attacks that have a higher number of instances in the training data, but it may overlook identifying attacks that were limited instances during the training phase.

To address these challenges, this paper introduces a multi-head attention-based gated recurrent unit (MAGRU) that scrutinizes IIoT network traffic to detect malicious activities. The proposed model utilizes the multi-head attention (MA) layer, which enhances the model's ability to learn limited samples classes by assigning different priorities to various sections of the input. This feature enables the model to effectively handle imbalance issues of training data. Subsequently, the gated recurrent unit (GRU) is employed for the detection of IIoT network behavior. IoT network data often exhibit a sequential nature, capturing events and activities over time. GRUs are specifically designed to handle sequential data, making them a natural choice for processing time-series data in IoT networks. One significant advantage of GRUs is their ability to address the vanishing gradient problem, which is a common issue faced by traditional recurrent neural networks (RNNs) when dealing with long sequences. By doing so, they can effectively capture long-term dependencies in the data, allowing the model to learn patterns and relationships that occur over extended periods. Furthermore, GRUs are computationally less expensive compared to other RNN variants, such as Long Short-Term Memory (LSTM) networks. This efficiency is particularly crucial for IoT devices with limited computational resources and power constraints.

The proposed MAGRU is evaluated using publicly available datasets including the Edge-IIoTset and MQTTset. The evaluation parameters employed for the proposed model include accuracy, macro precision, macro recall, and macro F1-score. Furthermore, to validate the performance of MAGRU, it is compared with several existing ML and DL models. The primary contributions of this paper are outlined as follows:

- This paper introduces a novel deep learning model for intrusion detection in IIoT networks, named MAGRU, comprising multi-head attention and GRU layers. Multi-head attention is employed to address the imbalance issues in the training data, while the GRU is utilized for detecting malicious activities in the network.
- A feature filtering technique is implemented to identify the most pertinent features that enhance the performance of intrusion detection. To achieve this, the extremely gradient boosting (XGBoost) method is employed to effectively filter the relevant features that positively impact intrusion detection.
- To validate the performance of the proposed MAGRU, we implemented various ML and DL models and compared their results against MAGRU using the same datasets.

The rest of the paper is organized as follows: Section II provides a comprehensive review of the related literature and compares them. Section III presents the proposed model in detail. The proposed methodology is explained in Section IV. In Section V, a thorough discussion of the outcomes is provided, along with a comparison of the proposed model against state-of-the-art methods. Finally, Section VI concludes the entire paper.

II. RELATED LITERATURE

The proliferation of smart devices in IIoT networks has led to an increase in potential cyber-attacks, which has caught the attention of researchers. Several experts have been actively working on enhancing the security of IIoT networks and have developed new models for detecting intrusions.

Abdel-Basset et al. [27] presented a forensics-based Deep Learning (DL) method called Deep-IFS for detecting intrusions in a dense network of IIoT. The Deep-IFS model is based on GRU architecture, which was employed as the foundation for their approach. To evaluate the performance of the Deep-IFS model, they used two datasets, namely BoT-IoT and UNSW-NB15. The experimental results demonstrated the effectiveness of their proposed model, achieving accuracy rates of 98.1% and 99.75% for the BoT-IoT and UNSW-NB15 datasets, respectively. Islam et al. [28] proposed an artificial neural network (ANN) for detecting intrusions in IIoT networks. They compared the performance of four different models using the NSL-KDD dataset. To implement their

TABLE 1. Literature overview.

Papers	Year	Model	Dataset	No. of classes	Accuracy (in %)
Abdel et al. [27]	2020	GRU	Bot-IoT, UNSW-NB15	5, 10	98.1, 99.75
Islam et al. [28]	2020	ANN	NSL-KDD	5	96.74
Ferrag et al. [29]	2021	CNN	CIC-DDoS2019, TON_IoT	13, 8	99.95, 99.92
Zhang et al. [29]	2021	GAN	NSL-KDD, CSE-CIC-IDS2018	5, 8	99, 96
Idrissi et al. [30]	2022	GAN	MQTTset	6	99
Le et al. [31]	2022	XGBoost	X-IIoTDS, TON_IoT	10, 8	99.9, 99.87
Liu et al. [32]	2022	VAE	CSE-CIC-IDS2018	8	98.57
Zhang et al. [33]	2022	GNN	Mississippi	8	97.2
Mohy et al. [34]	2023	RF	Bot-IoT, NF-UNSW-NB15-v2	5, 10	99.9, 99.30
El et al. [35]	2023	CNN	NSL-KDD	5	99
Abd et al. [22]	2023	DNN	NSL-KDD, Bot-IoT, KDD99, CIC-IDS2017	5, 5, 5. 7	99.9, 91.6, 92.1, 90.1
Sharma et al. [36]	2023	DNN	UNSW-NB15	10	99
This study	2023	MAGRU	Edge-IIoTset, MQTTset	15, 6	99.94, 99.99

approach, they utilized the R programming environment. Their analysis of the results revealed that the 2-layer ANN achieved a remarkable performance of 96.74%, surpassing other ANNs.

Ferrage et al. [29] adopted three deep learning-based models, which included convolutional neural networks (CNN), deep neural networks (DNN), and recurrent neural networks (RNN), to detect DDoS attacks in the Internet of Agricultural Things (IoAT). These models underwent evaluation using two publicly available datasets: CIC-DDoS2019 and TON_IoT. Notably, the CNN outperformed the other models, achieving accuracies of 99.95% and 99.92%, respectively. Zhang et al. [37] introduced a pretraining Wasserstein Generative Adversarial Network (PWGAN) for the detection of cyber attacks in the network flow of IIoT networks. Their proposed model comprises two components. In the first part, they implemented a pretraining approach, wherein the Wasserstein Generative Adversarial Network (WGAN) with gradient penalty was initially trained using normal data and subsequently retrained using attack classes. In the second part, they employed LightGBM as a classification model. The effectiveness of this system was assessed using the NSL-KDD and CIS-IDS2018 datasets, where they achieved accuracies of 99% and 96%, respectively.

Le et al. [31] adopted the XGBoost classifier for detecting intrusions in IoT networks. The XGBoost method was assessed using two public datasets: X-IIoTDS and TON_IoT. The main objective of the proposed model was to tackle the issue of data imbalance present in these datasets. Remarkably, they achieved accurate detection of 99.9% and 99.87%, respectively, on these datasets. Liu et al. [32] employed a variational autoencoder (VAE) model for the purpose of detecting malicious traffic in Industrial Internet of Things (IIoT) systems. To evaluate their proposed model, they utilized the CSE-CIC-IDS2018 dataset, achieving highly accurate results of 98.57%. Idrissi et al. [30] presented the use of a generative adversarial network (GAN) for detecting anomalies in host-based IoT network devices. The presented model was tested on the MQTTset dataset to validate its performance, and it achieved an impressive 99% accuracy based on the testing results. Zhang et al. [33] adopted a graph neural network (GNN) to effectively detect intrusions in industrial IoT network traffic. To assess the performance of their proposed system, they conducted evaluations using the Mississippi dataset, achieving an accuracy of 97.2%.

Mohy-Eddine et al. [34] introduced an approach that combines Isolation Forest (IF) and Pearson's Correlation Coefficient (PCC) to effectively remove outliers and extract relevant features. In order to accomplish the classification task, they employed a machine learning model called Random Forest (RF). To evaluate the performance of their model, they employed the Bot-IoT and NF-UNSW-NB12-v2 datasets. The experimental results demonstrated accuracy, achieving 99.9% on the BoT-IoT dataset and 99.30% on the NF-UNSW-NB12-v2 dataset. El-Ghamry et al. [35] utilized a convolutional neural network (CNN) for detecting cyber-attacks in agriculture IoT systems. In their approach, they transformed the data into RGB images and employed CNN for the classification task. The researchers evaluated their proposed method using the NSL-KDD dataset, achieving an accuracy of 99%. Elaziz et al. [22] presented a sophisticated hybrid model designed for detecting malicious traffic in cloud-based IoT. Their innovative approach combines swarm intelligence with deep neural networks (DNN). In order to assess the effectiveness of their model, the researchers conducted evaluations using four public datasets, namely NSL-KDD, BoT-IoT, KDD99, and CIC2017. They achieved high levels of accuracy on these datasets, obtaining 99.9%, 91.6%, 92.1%, and 90.1% respectively. Sharma et al. [36] proposed a DNN model for the detection of intrusions in IoT networks. They evaluated their proposed model using the UNSW-NB15 dataset. Initially, without balancing the dataset, their approach yielded an accuracy of 84%. However, to improve the performance, they employed the generative adversarial networks (GANs) method for data balancing. After applying this technique, they achieved an accuracy of 99%.

A comprehensive overview of the literature is provided in Table 1. Through an analysis of the related literature, it becomes evident that many papers have primarily concentrated on a limited number of classes due to imbalances in the datasets. Consequently, when faced with a larger number

of classes, these systems tend to struggle to achieve accurate detection performance. In contrast, this paper introduces a novel approach called the MAGRU paradigm, which enhances the performance of existing systems for both a limited number of classes and a higher number of class detections, bridging the gap in accuracy across different scenarios.

III. THE PROPOSED MAGRU MODEL

The proposed model comprises a multi-head attention module combined with GRU layers, followed by a softmax layer. The proposed MAGRU is presented in Algorithm 1. Initially, the input data undergoes a multi-head attention operation, which endows the model with the ability to concurrently concentrate on various segments of the input sequence, considering diverse positions. This capability enhances the model's capacity to grasp intricate relationships and dependencies within the input data. During this attention process, the input is partitioned into multiple smaller vectors, all of which are simultaneously processed [38]. This partitioning allows the attention mechanism to focus on different aspects (heads) of the input in parallel. The output of the attention layer is passed into gated recurrent unit (GRU) layers, which leverage past timestamps to discern the network's behavior [39].

Algorithm 1 Proposed MAGRU Algorithm

Require: Input data sequence X
Ensure: Predicted output
function MultiHeadAttention(X)
  $Q, K, V \leftarrow$ LinearTransform(X) ▷ Linear transformations
  $H_i \leftarrow$ Attention(Q, K, V) ▷ Multi-head attention
  return H_i
end function
function GRULayers(H_i)
  $H_{\text{GRU}} \leftarrow$ GRU(H_i) ▷ Gated Recurrent Unit layers
  return H_{GRU}
end function
function SoftmaxLayer(H_{GRU})
  $Y \leftarrow$ Softmax(H_{GRU}) ▷ Softmax layer for predictions
  return Y
end function
$H_i \leftarrow$ MultiHeadAttention(X)
$H_{\text{GRU}} \leftarrow$ GRULayers(H_i)
$Y \leftarrow$ SoftmaxLayer(H_{GRU})
return Y

The multi-head attention (MA) layer works on queries (Q), keys (K), and values (V), all of which are sequence vectors, as shown in Fig 2. These vectors are extracted from the same input but are utilized in different ways to capture distinct aspects of the input data. The MA layer processes all heads in parallel and computes the importance (attention) of each head in the input. The proposed model takes input in the shape of

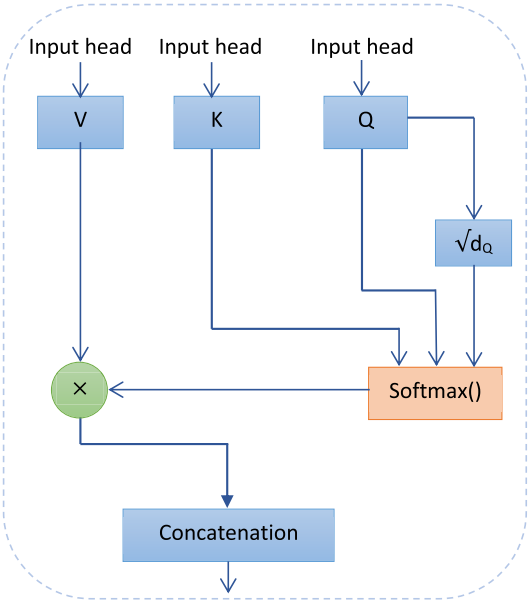


FIGURE 2. Basic architecture of multi-head attention.

(instances, attributes, 1), where '1' represents each instance. The input sequence is divided into eight equal-sized heads [40]. The size of the head is calculated using Eq 1.

$$S_{head} = \lceil \frac{S_{Input}}{N_{heads}} \rceil \tag{1}$$

where S_{head} is the size of the head, S_{Input} is the length of the input sequence, and N_{heads} represents the total number of heads. The Q, K, and V values for each attention head is calculated using Eq 2, Eq 3, and Eq 4, respectively.

$$Q = W_Q \times X \tag{2}$$
$$K = W_K \times X \tag{3}$$
$$V = W_V \times X \tag{4}$$

where X represents the input head, and W_Q, W_K, and W_V are the learned weights. Eq 5 is employed to calculate the attention for each head. where d_q represent the length of Q and A_h is the attention of each head. Finally, concatenate all attention heads to create the output of this layer.

$$A_h = \text{softmax}\left(\frac{Q \times K^T}{\sqrt{d_q}}\right) \times V \tag{5}$$

The output of the MA layer is then passed to the GRU layers. The GRU architecture consists of a total of two gates: the reset gate and the update gate, accompanied by a single hidden state [41]. Within the GRU, each gate is equipped with a sigmoid activation function, while a separate tanh function is employed to generate the output, as depicted in Fig. 3. Eq 6 and 7 provide a mathematical representation of reset and updated gates of GRU, respectively.

$$r_t = \sigma\left((w_{xr}x_t + w_{hr}h_{t-1} + b_r)\right) \tag{6}$$
$$u_t = \sigma\left((w_{xu}x_t + w_{ur}h_{t-1} + b_u)\right) \tag{7}$$

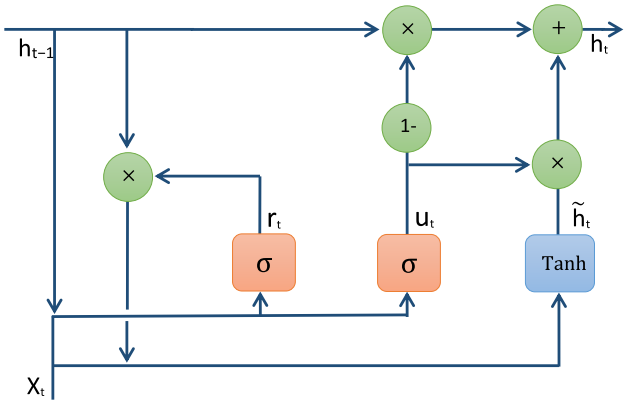


FIGURE 3. Basic architecture of GRU.

where 'rt' represents the reset gate for a time stamp 't' and 'it' represents the update gate. ht-1 represents the previous hidden state of the GRU. w is the weight value and b is the biases of reset and update gates. The hidden state value is computed using Eq 8 and 9.

$$\tilde{h}_t = \tanh\left(w_{hx}x_t + w_h h\left(r_t h_{t-1}\right) + b_u\right)) \tag{8}$$

$$h_t = \left(1 - u_t\right) h_{t-1} + u_t \tilde{h}_t \tag{9}$$

IV. THE PROPOSED METHODOLOGY

This section provides a comprehensive exposition of the proposed methodology, and presents its architecture in Fig 4, illustrating its key stages. The framework commences with a thorough examination of the dataset employed and encompasses the data preprocessing steps. Finally, the data is split into train and test sets using the stratified method. After the preprocessing phase, the model undergoes training and testing, leading to the acquisition of results. A step-by-step workflow of the proposed system is outlined in Algorithm 2.

Algorithm 2 Workflow Algorithm of the Proposed MAGRU

Require: Dataset D with attributes and labels
Ensure: Trained MAGRU model and results

Begin
Load dataset D
function Preprocessing_steps(D)
  Remove null value records from D
  Encode categorical attributes of D via label encoder
  $F, Y \leftarrow D$ ▷ Separate features F and labels Y
  $imp\_v \leftarrow$ XGBoost(F, Y) ▷ compute impact values imp_v of features
  $X \leftarrow \{v \in imp\_v \mid v > 0\}$
  $X_{\text{norm}} \leftarrow$ Normalize(X)
  $Train\_set, Test\_set \leftarrow stratified\_split(X_{\text{norm}}, Y)$
  return $Train\_set, Test\_set$
end function
function MAGRU_training($Train\_set$)
  $X_{train}, Y_{train} \leftarrow$ Split $Train\_set$ in mini_batches
  Repeat step 28 for N epochs
  for each $mini\_batch \in X_{train}$ **do**
    $\hat{y} \leftarrow MAGRU$.predict($mini\_batch$)
    $Loss \leftarrow Compute\_loss(\hat{y}, Y_{train})$
    Update_parameters($Loss$)
  end for
end function
function Evaluation($Test\_set$)
  $X_{train}, Y_{test} \leftarrow Test\_set$ ▷ Separate input X_{test} and labels Y_{test}
  $\hat{y} \leftarrow MAGRU$.predict(X_{test})
  Evaluate($\hat{y}, Y_{test}$) ▷ compute evaluation metrics
end function
$Train\_set, Test\_set \leftarrow$ Preprocessing_steps(D)
MAGRU_training($Train\_set$)
Evaluation($Test\_set$)
End

A. DATASET

The Edge-IIoTset[1] and MQTTset[2] stand as renowned and widely utilized datasets within the research community. Edge-IIoTset encompasses IoT and IIoT traffic instances, derived from a real-world testbed that intricately comprises seven interconnected layers. The testbed consisted of ten smart devices and sensors. The Edge-IIoTset is a compilation of fourteen attacks present, all of which are linked to IoT and IIoT communication protocols [42]. A detailed presentation of the dataset attacks is illustrated in Fig 5. The dataset encompasses a total of 2,219,201 instances, out of which 1,615,643 instances are categorized as normal, while the remaining 603,558 instances correspond to 14 distinct attacks. The MQTTset was generated from a real-time IoT network with eight IoT sensors connected to an MQTT broker, using the IoT-Flock tool [43]. It comprises five distinct attacks, all related to the MQTT communication protocol of IoT networks. The detailed presentation of MQTTset attacks is illustrated in Fig 6. This dataset contains a total of 541,071 instances, with 440,699 instances categorized as normal and the remaining 100,372 instances corresponding to the five distinct attacks.

[1]Access at: https://ieee-dataport.org/documents/edge-iiotset-new-comprehensive-realistic-cyber-security-dataset-iot-and-iiot-applications

[2]Access at: https://www.kaggle.com/datasets/cnrieiit/mqttset

B. DATA PREPROCESSING

Preprocessing steps play a crucial role in preparing the dataset to be well-suited for ML/DL models. In this study, several important steps were undertaken to prepare the dataset for optimal utilization. Firstly, we conducted a thorough check for empty and undefined values in the dataset, and fortunately, no empty value records were found. Next, we proceeded to transform categorical attributes into numerical ones using a label encoder. The label encoder assigned a unique integer value to each attribute record in alphabetic order, ensuring a proper numerical representation of the data.

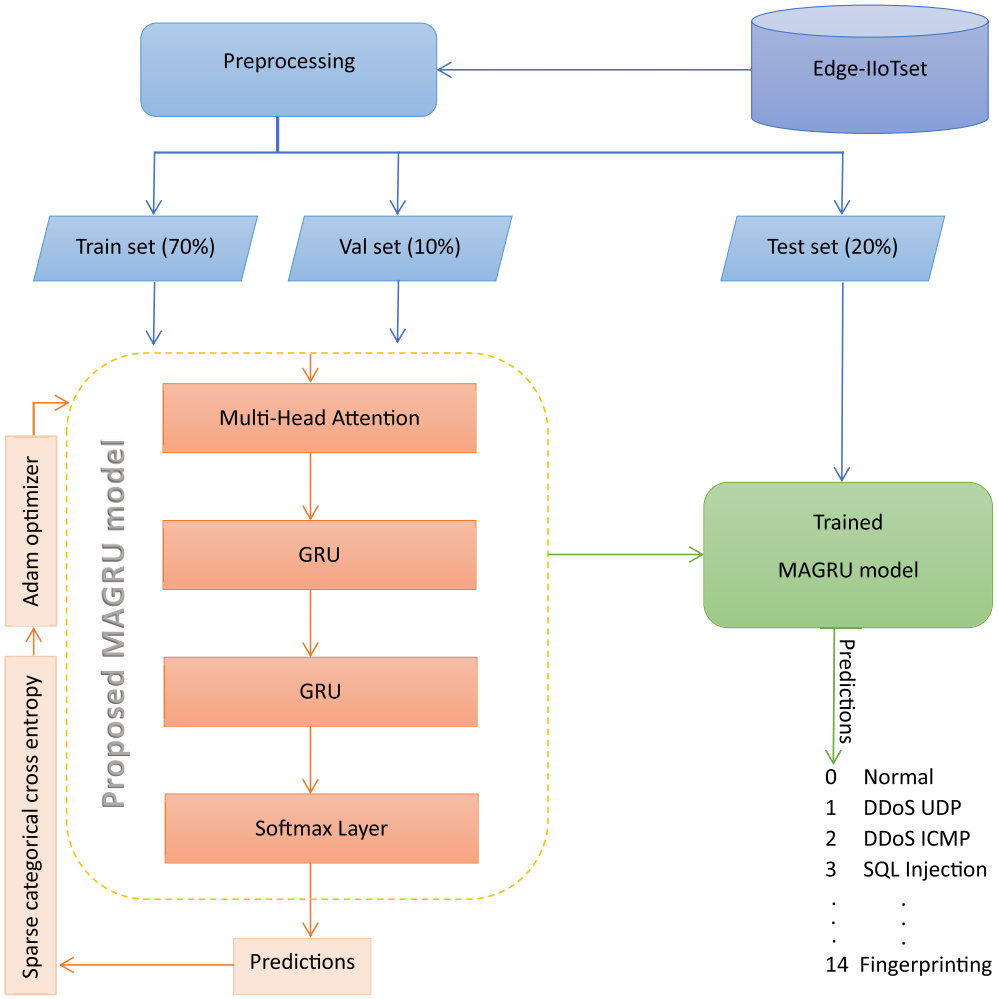


FIGURE 4. The proposed MAGRU architecture.

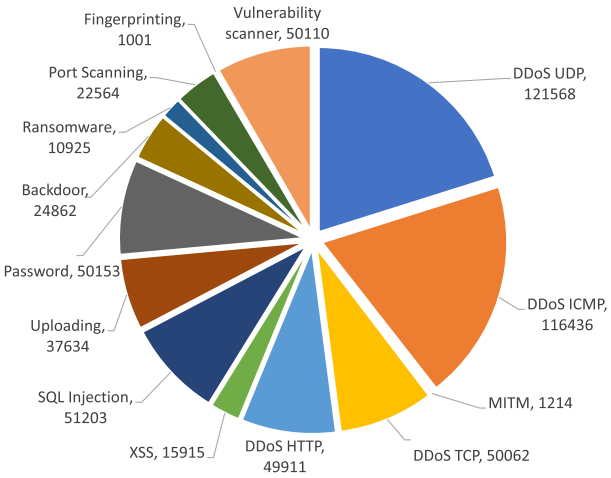


FIGURE 5. Edge-IIoTset attacks distribution.

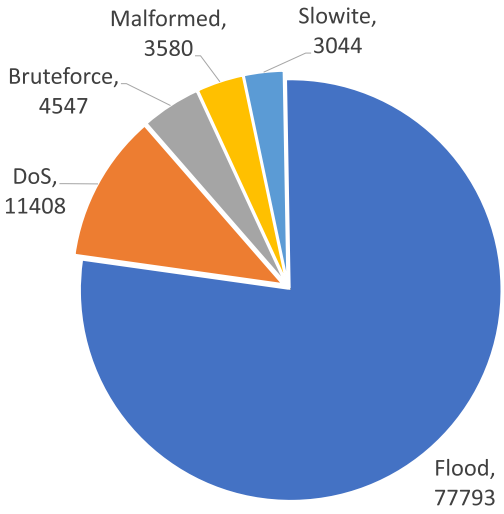


FIGURE 6. MQTTset attacks distribution.

1) FEATURES FILTERING

Feature filtering is a crucial step in the preprocessing of network data to select impactful features for intrusion detection in IIoT networks. This is because IIoT network datasets contain numerous features that can have a detrimental

effect on the detection model, leading to overfitting or underfitting. As a result, the model fails to detect unseen data and is unable to learn effectively from the training data. In this experiment, when we omitted feature filtering, we obtained results for several ML and DL models, including the proposed MAGRU, with less than 1% accuracy. This low accuracy indicates that the models were underfitting due to certain features containing mixed data from all classes, preventing the models from effectively learning. To address this problem and enhance the effectiveness of the proposed model, we performed feature filtering to select the most impactful attributes of the dataset. For this purpose, the extreme gradient boosting (XGBoost) method was employed. XGBoost calculated an importance score for each attribute, sorting them in descending order based on their significance. This score represents the contribution of each feature across all decision trees in the ensemble. Features that resulted in significant reductions of the objective function during splitting were considered more important. We selected features with an importance score greater than '0,' resulting in 31 filtered features from the Edge-IIoTset and 20 features from the MQTTset, while the remaining were neglected for the experiment.

2) NORMALIZATION

After the feature filtering process, we applied the min-max normalization method to scale all attributes within the range of 0 to 1. This normalization was employed because the dataset features were in variant ranges. Finally, to ensure a balanced and representative dataset for training and evaluation, we utilized the stratified method. This method facilitated an equal split of the data into 70% for training, 10% for validation, and 20% for testing sets, enabling a comprehensive assessment of the ML/DL models.

C. THE PROPOSED MAGRU HYPERPARAMETERS

In the implementation of the DL algorithm, we make use of various hyperparameters to achieve optimal performance. In the proposed MAGRU model, for the calculation of loss, we applied the sparse categorical cross-entropy method. To optimize the weights during training, the Adam optimizer was employed. The batch size was set to 32, and the number of epochs was configured to six, ensuring an efficient and effective training process. Within the architecture, both the GRU layers were designed with 31 output units. In the first GRU layer, the parameter "return_sequences" was set to True, while in the second GRU layer, it was set to False.

V. EXPERIMENTATION AND FINDINGS

This section presents a detailed explanation of the experimental environment, showcasing the outcomes of the proposed MAGRU model, and subsequently, comparing its performance with other ML and DL models.

A. PERFORMANCE ASSESSMENT PARAMETERS

To evaluate the performance of the proposed MAGRU, we utilized various parameters, including Accuracy, Precision, Recall, and F1-score [44]. Accuracy is a metric that measures the proportion of data instances that have been correctly classified out of the total number of data instances [45]. Precision is a metric that provides the ratio of true positives to the total number of positives that the model predicts [46]. Recall refers to the model's ability to efficiently classify all positive instances [46]. The F1-score serves as a measure that harmoniously blends both recall and precision [47]. Since there exists a delicate balance between precision and recall, the F1-score becomes a valuable yardstick for evaluating how adeptly our models strike that balance. In this study, we utilized the macro (M) precision, recall, and F1-score, which captured the collective outcomes averaged across all classes. All these parameters were computed using Eq 10-13. Where k is the number of classes, while α, β, γ, and δ represent true positive, true negative, false positive, and false negative respectively.

$$\text{Accuracy} = \frac{\alpha + \beta}{\alpha + \beta + \gamma + \delta} \tag{10}$$

$$\text{M-Precision} = \frac{1}{k}\sum_{i=1}^{k}\frac{\alpha_i}{\alpha_i + \gamma_i} \tag{11}$$

$$\text{M-Recall} = \frac{1}{k}\sum_{i=1}^{k}\frac{\alpha_i}{\alpha_i + \delta_i} \tag{12}$$

$$\text{M-F1 Score} = \frac{2 \times (\text{M-Precision} \times \text{M-Recall})}{\text{M-Precision} + \text{M-Recall}} \tag{13}$$

B. IMPLEMENTATION PLATFORM

All the experiments were implemented on an HP desktop computer equipped with a 9th-generation core-i9 CPU and 64 GB of DDR4 RAM. To achieve high processing performance, a GEFORCE RTX 2080 GPU was installed. All experiments were conducted utilizing the GPU. The classification algorithms were coded in Python 3.9 using Jupyter Notebook. Various libraries, such as Tensorflow, Pandas, scikit-learn, and Numpy, were installed to support the implementations. It is important to note that all of these tasks were performed on a Windows 11 Pro operating system.

C. PERFORMANCE EVALUATIONS OF THE PROPOSED MAGRU

The outcomes of the proposed MAGRU are presented in this subsection. Additionally, to validate the performance of the proposed model, we compared the outcomes with those of other state-of-the-art models. The results are validated using fivefold cross-validation.

1) PERFORMANCE ASSESSMENT WITH DIFFERENT PARAMETERS

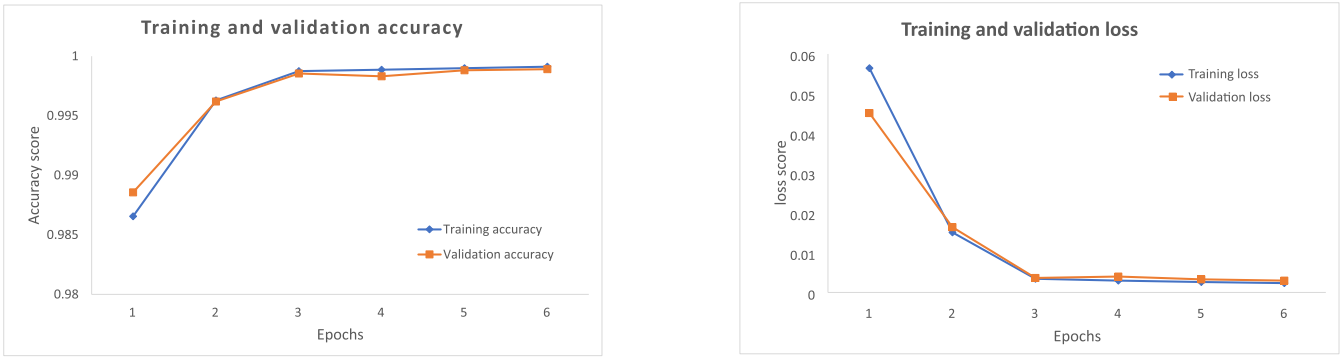
As previously mentioned, the datasets were split into train, validation, and test sets with a ratio of 70%, 10%, and 20% respectively. Subsequently, the models were trained over the train set and validated over the validation set, with different parameter adjustments. Table 2 and Table 3

TABLE 2. Performance assessment with various parameters on Edge-IIoTset.

GRU Layers	Batch Size	Optimizer	M-Precision	M-Recall	M-F1-Score	Accuracy	Training Time (in sec)	Testing Time (in sec)
1	32	Adam	0.9775	0.9509	0.9586	0.9985	254	26
2	32	Adam	0.9933	0.9948	0.9940	0.9994	366	38
3	32	Adam	0.9943	0.9472	0.9577	0.9990	466	59
2	64	Adam	0.9924	0.9909	0.9915	0.9991	190	37
2	128	Adam	0.9929	0.9874	0.9899	0.9991	89	36
2	256	Adam	0.9753	0.9651	0.9696	0.9983	46	35
2	32	Nadam	0.9951	0.9930	0.9940	0.9994	591	39
2	64	Nadam	0.9953	0.9921	0.9936	0.9992	591	46
2	128	Nadam	0.9902	0.9739	0.9801	0.9989	150	45
2	256	Nadam	0.9950	0.9921	0.9937	0.9993	73	35
2	32	Adamax	0.8700	0.8898	0.8772	0.9965	367	38
2	64	Adamax	0.9592	0.9221	0.9231	0.9964	174	37
2	128	Adamax	0.9408	0.8270	0.8369	0.9908	86	36
2	256	Adamax	0.9149	0.8178	0.8210	0.9849	45	35
2	32	SGD	0.7347	0.8302	0.7503	0.8730	347	38
2	64	SGD	0.6327	0.6047	0.5863	0.9622	163	37
2	128	SGD	0.6274	0.6361	0.6244	0.9633	83	36
2	256	SGD	0.6729	0.6098	0.5944	0.9528	44	35

TABLE 3. Performance assessment with various parameters on MQTTset.

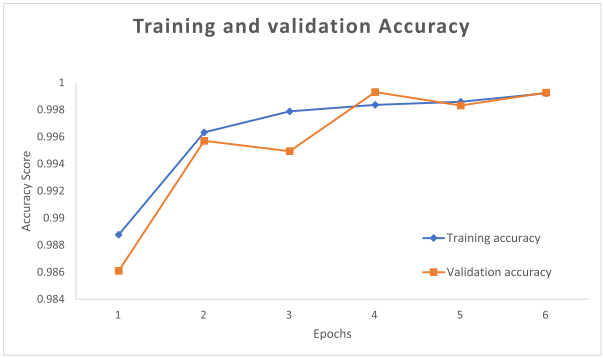
GRU Layers	Batch Size	Optimizer	M-Precision	M-Recall	M-F1-Score	Accuracy	Training Time (in sec)	Testing Time (in sec)
1	32	Adam	0.9869	0.9887	0.9874	0.9994	67	7
2	32	Adam	0.9991	0.9985	0.9988	0.9999	85	10
3	32	Adam	0.9919	0.9889	0.9903	0.9995	111	11
2	64	Adam	0.9957	0.9961	0.9958	0.9997	46	9
2	128	Adam	0.9692	0.957	0.9625	0.998	25	10
2	256	Adam	0.9532	0.9671	0.9592	0.9973	13	9
2	32	Nadam	0.9976	0.9979	0.9978	0.9999	137	10
2	64	Nadam	0.9854	0.9865	0.9856	0.9993	63	9
2	128	Nadam	0.9844	0.9846	0.9840	0.9992	31	8
2	256	Nadam	0.9832	0.9820	0.9823	0.9991	17	8
2	32	Adamax	0.9215	0.9274	0.9237	0.9951	89	10
2	64	Adamax	0.9301	0.9227	0.9267	0.9954	41	9
2	128	Adamax	0.9775	0.9765	0.9768	0.9986	20	8
2	256	Adamax	0.9394	0.9495	0.9432	0.9963	10	8
2	32	SGD	0.6723	0.7503	0.6929	0.9821	84	9
2	64	SGD	0.3903	0.4422	0.4104	0.9705	41	8
2	128	SGD	0.4043	0.4219	0.4128	0.9651	19	8
2	256	SGD	0.4362	0.4138	0.4214	0.9673	10	8



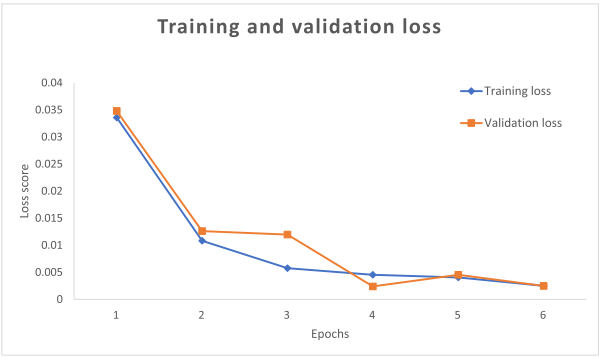
(a) Accuracy (b) Loss

FIGURE 7. Training and validation performance of the proposed MAGRU on Edge-IIoTset.

present a comprehensive comparison of the model testing results achieved through different parameter combinations. The analysis demonstrates that MAGRU attained optimal performance with two layers of GRU, a batch size of 32,



(a) Accuracy



(b) Loss

FIGURE 8. Training and validation performance of the proposed MAGRU on MQTTset.

TABLE 4. Performance comparison with other models on Edge-IIoTset.

Models	M-Precision	M-Recall	M-F1-Score	Accuracy	Training Time (in sec)	Testing Time (in sec)
LSTM	0.9865	0.9824	0.9842	0.9983	740	52
GRU	0.9281	0.9791	0.9338	0.9938	759	64
CNN	0.9905	0.9657	0.9752	0.9985	89	12
DAE	0.9846	0.9778	0.9810	0.9890	84	11
MLP	0.9891	0.9797	0.9837	0.9988	125	.18
LR	0.9288	0.8886	0.8975	0.9924	105	.13
NB	0.9617	0.9352	0.9423	0.9961	.83	2
MAGRU	0.9933	0.9948	0.9940	0.9994	366	38

TABLE 5. Performance comparison with other models on MQTTset.

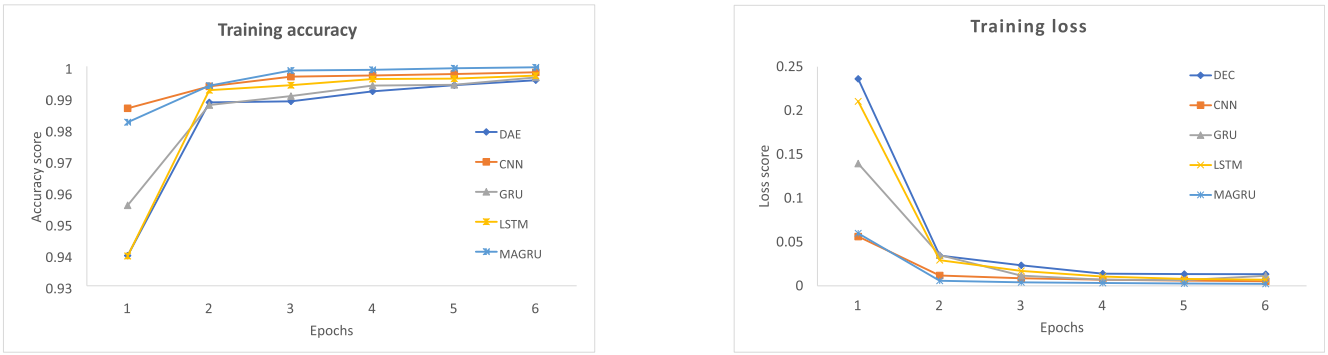
Algorithms	M-Precision	M-Recall	M-F1-Score	Accuracy	Training Time (in sec)	Testing Time (in sec)
LSTM	0.9976	0.9546	0.974	0.9974	453	22
GRU	0.9215	0.9271	0.9229	0.9956	542	23
CNN	0.9711	0.9714	0.9712	0.9981	25	3
DAE	0.9709	0.9736	0.9706	0.9979	25	3
MLP	0.9449	0.9465	0.9448	0.9969	38	.25
LR	0.8957	0.8628	0.8612	0.9932	14	.15
NB	0.7918	0.8113	0.7959	0.9871	.15	.12
MAGRU	0.9991	0.9985	0.9988	0.9999	84	9

and utilizing two optimization functions: Adam and Nadam. These specific parameter combinations outperformed others, establishing their effectiveness for the proposed MAGRU model. The training and validation performance of the proposed MAGRU are depicted in Fig 7 and Fig 8.

2) PERFORMANCE COMPARISON WITH STATE-OF-THE-ART MODELS

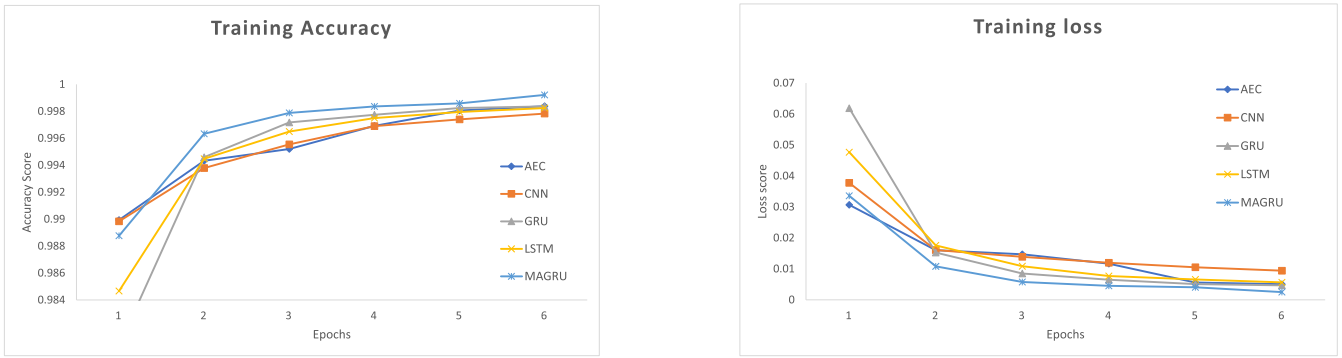
The performance of the proposed MAGRU model has been validated by comparing its results with several state-of-the-art methods. The traditional ML and advanced DL models used for comparison include multi-layer perceptron (MLP), naive bayes (NB), linear regression (LR), deep autoencoder (DAE), long short-term memory (LSTM), gated recurrent units (GRU), and convolutional neural networks (CNN). All these models were implemented in the same environment with the same preprocessing steps as the proposed model, ensuring a fair comparison among them.

For all the models, we employed the sparse categorical cross-entropy loss function, Adam optimizer, and a batch size of 32 during training. The training process was conducted for six epochs. Fig 9 and Fig 10 illustrate the comparison of the training performance between the proposed model and the other models. The results demonstrate that the proposed model exhibits superior learning capability, achieving faster convergence compared to the other models. Furthermore, the comparison of testing performance between these models and the proposed MAGRU model is presented



(a) Accuracy (b) Loss

FIGURE 9. Comparison of training performance with other models on Edge-IIoTset.



(a) Accuracy (b) Loss

FIGURE 10. Comparison of training performance with other models on MQTTset.

TABLE 6. Performance comparison with Edge-IIoTset related articles.

Articles	Precision	Recall	F1-score	Accuracy
[48]	0.9746	0.9657	0.9691	0.9888
[49]	0.9878	0.9722	-	0.9832
[50]	0.9603	-	-	0.9727
[51]	0.885	0.613	0.724	-
[52]	-	-	-	0.9869
[53]	-	-	-	0.9964
Proposed	0.9933	0.9948	0.9940	0.9994

TABLE 7. Performance comparison with MQTTset related articles.

Articles	Precision	Recall	F1-score	Accuracy
[43]	-	-	0.914	0.9159
[54]	-	-	0.9829	0.9697
[55]	0.9738	0.9768	0.9856	0.9835
[56]	0.8283	0.7533	0.8283	0.9076
[57]	0.9829	0.9328	0.9572	0.9584
Proposed	0.9991	0.9985	0.9988	0.9999

in Table 4 and Table 5. In addition, The performance comparison with the other related articles on similar datasets is presented in Table 6 and Table 7. The testing results reveal that the proposed model outperforms the other models, demonstrating optimal performance in the detection of malicious activities in IIoT networks.

VI. CONCLUSION

Existing IDSs face challenges during training when dealing with imbalanced training data and a higher number of classes. These issues can significantly reduce the IDS's performance and may result in missed IIoT network attacks, especially those with fewer training samples. To address the aforementioned challenges, this work introduces a novel DL model called MAGRU for monitoring IIoT network traffic and detecting malicious activities. The proposed model effectively tackles the challenges posed by imbalanced data and a higher number of classes, resulting in improved performance. The imbalanced issue is addressed by utilizing multi-head attention (MA) in the proposed model, which has the ability to assign importance to input features instead of considering the number of samples. GRU is employed for the detection of IIoT network behavior in the context of a higher number of classes. The proposed approach is evaluated using two real-time IIoT network datasets, namely Edge-IIoTset and MQTTset. MAGRU's performance is validated against various ML and DL models, outperforming the others with an average precision, recall, F1-score, and accuracy of 99.62%, 99.67%, 99.64%, and 99.97%, respectively, demonstrating its optimal performance in detecting intrusions in IIoT networks.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Prince Sultan University for paying the Article Processing Charges (APC) of this publication.

REFERENCES

[1] O. Friha, M. A. Ferrag, M. Benbouzid, T. Berghout, B. Kantarci, and K.-K.-R. Choo, "2DF-IDS: Decentralized and differentially private federated learning-based intrusion detection system for industrial IoT," *Comput. Secur.*, vol. 127, Apr. 2023, Art. no. 103097.

[2] P. Peniak, E. Bubeníková, and A. Kanáliková, "Validation of high-availability model for edge devices and IIoT," *Sensors*, vol. 23, no. 10, p. 4871, 2023.

[3] A. Shafique, J. Ahmed, W. Boulila, H. Ghandorh, J. Ahmad, and M. U. Rehman, "Detecting the security level of various cryptosystems using machine learning models," *IEEE Access*, vol. 9, pp. 9383–9393, 2021.

[4] L. A. C. Ahakonye, C. I. Nwakanma, J. M. Lee, and D.-S. Kim, "Agnostic CH-DT technique for SCADA network high-dimensional data-aware intrusion detection system," *IEEE Internet Things J.*, vol. 10, no. 12, pp. 10344–10356, Jun. 2023.

[5] S. Nagarajan, S. Kayalvizhi, R. Subhashini, and V. Anitha, "Hybrid honey badger-world cup algorithm-based deep learning for malicious intrusion detection in industrial control systems," *Comput. Ind. Eng.*, vol. 180, Jun. 2023, Art. no. 109166.

[6] J. Mao, X. Xu, Q. Lin, L. Ma, and J. Liu, "EScope: Effective event validation for IoT systems based on state correlation," *Big Data Mining Anal.*, vol. 6, no. 2, pp. 218–233, Jun. 2023.

[7] S. Das, S. Namasudra, S. Deb, P. M. Ger, and R. G. Crespo, "Securing IoT-based smart healthcare systems by using advanced lightweight privacy-preserving authentication scheme," *IEEE Internet Things J.*, early access, Jun. 12, 2023, doi: 10.1109/JIOT.2023.3283347.

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

[9] I. A. Khan, N. Moustafa, D. Pi, K. M. Sallam, A. Y. Zomaya, and B. Li, "A new explainable deep learning framework for cyber threat discovery in industrial IoT networks," *IEEE Internet Things J.*, vol. 9, no. 13, pp. 11604–11613, Jul. 2022.

[10] A. Noorwali, A. N. Alvi, M. Z. Khan, M. A. Javed, W. Boulila, and P. A. Pattanaik, "A novel QoS-oriented intrusion detection mechanism for IoT applications," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–10, Jun. 2021.

[11] K.-A. Tait, J. S. Khan, F. Alqahtani, A. A. Shah, F. A. Khan, M. U. Rehman, W. Boulila, and J. Ahmad, "Intrusion detection using machine learning techniques: An experimental comparison," in *Proc. Int. Congr. Adv. Technol. Eng. (ICOTEN)*, Jul. 2021, pp. 1–10.

[12] M. A. Khan, M. A. K. Khattk, S. Latif, A. A. Shah, M. Ur Rehman, W. Boulila, M. Driss, and J. Ahmad, "Voting classifier-based intrusion detection for IoT networks," in *Advances on Smart and Soft Computing*. Cham, Switzerland: Springer, 2022, pp. 313–328.

[13] V. Jayagopal, M. Elangovan, S. S. Singaram, K. B. Shanmugam, B. Subramaniam, and S. Bhukya, "Intrusion detection system in industrial cyber-physical system using clustered federated learning," *Social Netw. Comput. Sci.*, vol. 4, no. 5, p. 452, Jun. 2023.

[14] S. Fraihat, S. Makhadmeh, M. Awad, M. A. Al-Betar, and A. Al-Redhaei, "Intrusion detection system for large-scale IoT NetFlow networks using machine learning with modified arithmetic optimization algorithm," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100819.

[15] T. Saranya, S. Sridevi, C. Deisy, T. D. Chung, and M. K. A. A. Khan, "Performance analysis of machine learning algorithms in intrusion detection system: A review," *Proc. Comput. Sci.*, vol. 171, pp. 1251–1260, Jan. 2020.

[16] I. A. Khan, M. Keshk, D. Pi, N. Khan, Y. Hussain, and H. Soliman, "Enhancing IIoT networks protection: A robust security model for attack detection in internet industrial control systems," *Ad Hoc Netw.*, vol. 134, Sep. 2022, Art. no. 102930.

[17] M. Driss, D. Hasan, W. Boulila, and J. Ahmad, "Microservices in IoT security: Current solutions, research challenges, and future directions," *Proc. Comput. Sci.*, vol. 192, pp. 2385–2395, Jan. 2021.

[18] S. Sivamohan, S. S. Sridhar, and S. Krishnaveni, "TEA-EKHO-IDS: An intrusion detection system for industrial CPS with trustworthy explainable AI and enhanced Krill Herd optimization," *Peer Peer Netw. Appl.*, vol. 16, no. 4, pp. 1993–2021, Aug. 2023.

[19] I. Hidayat, M. Z. Ali, and A. Arshad, "Machine learning-based intrusion detection system: An experimental comparison," *J. Comput. Cognit. Eng.*, vol. 2, pp. 88–97, Jul. 2022.

[20] T. Saba, T. Sadad, A. Rehman, Z. Mehmood, and Q. Javaid, "Intrusion detection system through advance machine learning for the Internet of Things networks," *IT Prof.*, vol. 23, no. 2, pp. 58–64, Mar. 2021.

[21] N. Moustafa, N. Koroniotis, M. Keshk, A. Y. Zomaya, and Z. Tari, "Explainable intrusion detection for cyber defences in the Internet of Things: Opportunities and solutions," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 3, pp. 1775–1807, 3rd Quart., 2023.

[22] M. Abd Elaziz, M. A. A. Al-qaness, A. Dahou, R. A. Ibrahim, and A. A. A. El-Latif, "Intrusion detection approach for cloud and IoT environments using deep learning and capuchin search algorithm," *Adv. Eng. Softw.*, vol. 176, Feb. 2023, Art. no. 103402.

[23] I. A. Khan, D. Pi, M. Z. Abbas, U. Zia, Y. Hussain, and H. Soliman, "Federated-SRUs: A federated-simple-recurrent-units-based IDS for accurate detection of cyber attacks against IoT-augmented industrial control systems," *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8467–8476, May 2022.

[24] T. Saba, A. Rehman, T. Sadad, H. Kolivand, and S. A. Bahaj, "Anomaly-based intrusion detection system for IoT networks through deep learning model," *Comput. Electr. Eng.*, vol. 99, Apr. 2022, Art. no. 107810.

[25] N. Bugshan, I. Khalil, A. P. Kalapaaking, and M. Atiquzzaman, "Intrusion detection-based ensemble learning and microservices for zero touch networks," *IEEE Commun. Mag.*, vol. 61, no. 6, pp. 86–92, Jun. 2023.

[26] G. de Carvalho Bertoli, L. Alves Pereira Junior, O. Saotome, and A. L. dos Santos, "Generalizing intrusion detection for heterogeneous networks: A stacked-unsupervised federated learning approach," *Comput. Secur.*, vol. 127, Apr. 2023, Art. no. 103106.

[27] M. Abdel-Basset, V. Chang, H. Hawash, R. K. Chakrabortty, and M. Ryan, "Deep-IFS: Intrusion detection approach for industrial Internet of Things traffic in fog environment," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7704–7715, Nov. 2021.

[28] F. B. Islam, R. Akter, D.-S. Kim, and J.-M. Lee, "Deep learning based network intrusion detection for industrial Internet of Things," in *Proc. Korea Commun. Soc. Conf.*, Aug. 2020, pp. 418–421.

[29] M. A. Ferrag, L. Shu, H. Djallel, and K.-K.-R. Choo, "Deep learning-based intrusion detection for distributed denial of service attack in agriculture 4.0," *Electronics*, vol. 10, no. 11, p. 1257, May 2021.

[30] I. Idrissi, M. Azizi, and O. Moussaoui, "An unsupervised generative adversarial network based-host intrusion detection system for Internet of Things devices," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 25, no. 2, p. 1140, Feb. 2022.

[31] T.-T.-H. Le, Y. E. Oktian, and H. Kim, "XGBoost for imbalanced multiclass classification-based industrial Internet of Things intrusion detection systems," *Sustainability*, vol. 14, no. 14, p. 8707, Jul. 2022.

[32] C. Liu, R. Antypenko, I. Sushko, and O. Zakharchenko, "Intrusion detection system after data augmentation schemes based on the VAE and CVAE," *IEEE Trans. Rel.*, vol. 71, no. 2, pp. 1000–1010, Jun. 2022.

[33] Y. Zhang, C. Yang, K. Huang, and Y. Li, "Intrusion detection of industrial Internet-of-Things based on reconstructed graph neural networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 2894–2905, Sep. 2022.

[34] M. Mohy-Eddine, A. Guezzaz, S. Benkirane, M. Azrour, and Y. Farhaoui, "An ensemble learning based intrusion detection model for industrial IoT security," *Big Data Mining Analytics*, vol. 6, no. 3, pp. 273–287, Sep. 2023.

[35] A. El-Ghamry, A. Darwish, and A. E. Hassanien, "An optimized CNN-based intrusion detection system for reducing risks in smart farming," *Internet Things*, vol. 22, Jul. 2023, Art. no. 100709.

[36] B. Sharma, L. Sharma, C. Lal, and S. Roy, "Anomaly based network intrusion detection for IoT attacks using deep learning technique," *Comput. Electr. Eng.*, vol. 107, Apr. 2023, Art. no. 108626.

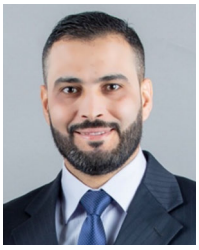
[37] L. Zhang, S. Jiang, X. Shen, B. B. Gupta, and Z. Tian, "PWG-IDS: An intrusion detection model for solving class imbalance in IIoT networks using generative adversarial networks," 2021, *arXiv:2110.03445*.

[38] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021.

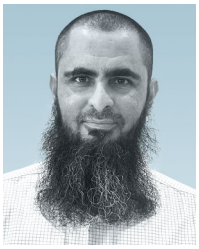
[39] N. B. Singh, M. M. Singh, A. Sarkar, and J. K. Mandal, "A novel wide & deep transfer learning stacked GRU framework for network intrusion detection," *J. Inf. Secur. Appl.*, vol. 61, Sep. 2021, Art. no. 102899.

[40] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008.
[41] R. Koniki, M. D. Ampapurapu, and P. K. Kollu, "An anomaly based network intrusion detection system using LSTM and GRU," in *Proc. Int. Conf. Electron. Syst. Intell. Comput. (ICESIC)*, Apr. 2022, pp. 79–84.
[42] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning," *IEEE Access*, vol. 10, pp. 40281–40306, 2022.
[43] I. Vaccari, G. Chiola, M. Aiello, M. Mongelli, and E. Cambiaso, "MQTTset, a new dataset for machine learning techniques on MQTT," *Sensors*, vol. 20, no. 22, p. 6578, Nov. 2020.
[44] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, *arXiv:2010.16061*.
[45] R. Lohiya and A. Thakkar, "Intrusion detection using deep neural network with antirectifier layer," in *Applied Soft Computing and Communication Networks*. Cham, Switzerland: Springer, 2021, pp. 89–105.
[46] A. A. Salih and A. M. Abdulazeez, "Evaluation of classification algorithms for intrusion detection system: A review," *J. Soft Comput. Data Mining*, vol. 2, no. 1, pp. 31–40, Apr. 2021.
[47] M. Al-Omari, M. Rawashdeh, F. Qutaishat, M. Alshira'H, and N. Ababneh, "An intelligent tree-based intrusion detection model for cyber security," *J. Netw. Syst. Manage.*, vol. 29, no. 2, pp. 1–18, Apr. 2021.
[48] A. A. Alashhab, M. S. M. Zahid, A. Muneer, and M. Abdukkahi, "Low-rate DDoS attack detection using deep learning for SDN-enabled IoT networks," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 11, pp. 1–7, 2022.
[49] D. Javeed, T. Gao, M. S. Saeed, and P. Kumar, "An intrusion detection system for edge-envisioned smart agriculture in extreme environment," *IEEE Internet Things J.*, early access, Jun. 22, 2023, doi: 10.1109/JIOT.2023.3288544.
[50] P. Dini, A. Begni, S. Ciavarella, E. De Paoli, G. Fiorelli, C. Silvestro, and S. Saponara, "Design and testing novel one-class classifier based on polynomial interpolation with application to networking security," *IEEE Access*, vol. 10, pp. 67910–67924, 2022.
[51] T. Shen, L. Ding, J. Sun, C. Jing, F. Guo, and C. Wu, "Edge computing for IoT security: Integrating machine learning with key agreement," in *Proc. 3rd Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Jan. 2023, pp. 474–483.
[52] A. Khacha, R. Saadouni, Y. Harbi, and Z. Aliouat, "Hybrid deep learning-based intrusion detection system for industrial Internet of Things," in *Proc. 5th Int. Symp. Informat. its Appl. (ISIA)*, Nov. 2022, pp. 1–6.
[53] V. Hnamte and J. Hussain, "DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system," *Telematics Informat. Rep.*, vol. 10, Jun. 2023, Art. no. 100053.
[54] M. S. Mahmood and N. B. Al Dabagh, "Improving IoT security using lightweight based deep learning protection model," *Tikrit J. Eng. Sci.*, vol. 30, no. 1, pp. 119–129, Mar. 2023.
[55] C. Prajisha and A. R. Vasudevan, "An efficient intrusion detection system for MQTT-IoT using enhanced chaotic Salp swarm algorithm and LightGBM," *Int. J. Inf. Secur.*, vol. 21, no. 6, pp. 1263–1282, Dec. 2022.
[56] O. Friha, M. A. Ferrag, L. Shu, L. Maglaras, K.-K.-R. Choo, and M. Nafaa, "FELIDS: Federated learning-based intrusion detection system for agricultural Internet of Things," *J. Parallel Distrib. Comput.*, vol. 165, pp. 17–31, Jul. 2022.
[57] S. Rachmadi, S. Mandala, and D. Oktaria, "Detection of DoS attack using AdaBoost algorithm on IoT system," in *Proc. Int. Conf. Data Sci. Its Appl. (ICoDSA)*, Oct. 2021, pp. 28–33.

SAFI ULLAH received the M.Sc. degree in computer science from the University of Malakand, Chakdara, Pakistan, in 2019, and the M.Phil. degree in computer science from Quaid-i-Azam University, Islamabad, Pakistan, in 2022. He is currently a Researcher with the College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia. His research interests include cyber security, the Internet of Things, electric vehicles, and image processing.



WADII BOULILA (Senior Member, IEEE) received the B.Eng. degree (Hons.) in computer science from the Borj El Amri Aviation School, in 2005, the M.Sc. degree in computer science from the National School of Computer Science (ENSI), University of Manouba, Tunisia, in 2007, and the Ph.D. degree in computer science from ENSI and Telecom-Bretagne, University of Rennes 1, France, in 2012. He is currently an Associate Professor of computer science with Prince Sultan University, Saudi Arabia. He is also a Senior Researcher with the RIOTU Laboratory, Prince Sultan University, and the RIADI Laboratory, University of Manouba. He was previously a Senior Research Fellow with the ITI Department, University of Rennes 1. He has participated in numerous research and industrial-funded projects. His research interests include data science, computer vision, big data analytics, deep learning, cybersecurity, artificial intelligence, and uncertainty modeling. He is a member of ACM and a Senior Fellow of the Higher Education Academy (SFHEA), U.K. He served as a TPC member, a reviewer, and the chair, for many leading international conferences and journals. He received the Award of the Young Researcher in computer science from Beit El-Hikma, Tunisia, in 2021, the Award of Best Researcher from the University of Manouba, in 2021, and the Award of Most Cited Researcher from the University of Manouba, in 2022. His work has gained global recognition, and he has been nominated as one of the top 2% of scientists in his field by Stanford University.



ANIS KOUBÂA is currently a Professor of computer science and the Leader of the Robotics and Internet of Things Research Laboratory, Prince Sultan University. He is also a Senior Researcher with CISTER and ISEP-IPP, Porto, Portugal, and a Research and Development Consultant with Gaitech Robotics, China. His current research interests include providing solutions toward the integration of robots and drones into the Internet of Things (IoT) and clouds, in the context of cloud robotics, robot operating systems (ROSs), robotic software engineering, wireless communication for the IoT, real-time communication, safety and security for cloud robotics, intelligent algorithm's design for mobile robots, and multi-robot task allocation. He is a Senior Fellow of the Higher Education Academy (HEA), U.K. He has been the Chair of the ACM Chapter in Saudi Arabia, since 2014.

JAWAD AHMAD (Senior Member, IEEE) is an experienced Researcher with more than ten years of cutting-edge research and teaching experience in prestigious institutes, including Edinburgh Napier University, U.K.; Glasgow Caledonian University, U.K.; Hongik University, South Korea; and HITEC University Taxila, Pakistan. He has taught various courses both at Undergraduate (UG) and Postgraduate (PG) levels during his career. He has coauthored more than 100 research papers in international journals and peer-reviewed international conference proceedings. His research interests include cybersecurity, multimedia encryption, and machine learning. He regularly organizes timely special sessions and workshops for several flagship IEEE conferences. He is an invited reviewer for numerous world-leading high-impact journals (reviewed more than 100 journal articles to date).