Computers, Materials & Continua
DOI: 10.32604/cmc.2022.031303
Article

Tech Science Press

Hyperparameter Tuned Deep Learning Enabled Intrusion Detection on Internet of Everything Environment

Manar Ahmed Hamza[1,2,*], Aisha Hassan Abdalla Hashim[1], Heba G. Mohamed[3], Saud S. Alotaibi[4], Hany Mahgoub[5,6], Amal S. Mehanna[7] and Abdelwahed Motwakel[2]

[1]Department of Electrical and Computer Engineering, International Islamic University Malaysia, Kuala Lumpur, 53100, Malaysia
[2]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia
[3]Department of Electrical Engineering, College of Engineering, Princess Nourah Bint Abdulrahman University, P. O. Box 84428, Riyadh, 11671, Saudi Arabia
[4]Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Saudi Arabia
[5]Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia
[6]Faculty of Computers and Information, Computer Science Department, Menoufia University, Egypt
[7]Department of Digital Media, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo, 11845, Egypt
*Corresponding Author: Manar Ahmed Hamza. Email: ma.hamza@psau.edu.sa
Received: 14 April 2022; Accepted: 10 June 2022

Abstract: Internet of Everything (IoE), the recent technological advancement, represents an interconnected network of people, processes, data, and things. In recent times, IoE gained significant attention among entrepreneurs, individuals, and communities owing to its realization of intense values from the connected entities. On the other hand, the massive increase in data generation from IoE applications enables the transmission of big data, from context-aware machines, into useful data. Security and privacy pose serious challenges in designing IoE environment which can be addressed by developing effective Intrusion Detection Systems (IDS). In this background, the current study develops Intelligent Multiverse Optimization with Deep Learning Enabled Intrusion Detection System (IMVO-DLIDS) for IoT environment. The presented IMVO-DLIDS model focuses on identification and classification of intrusions in IoT environment. The proposed IMVO-DLIDS model follows a three-stage process. At first, data pre-processing is performed to convert the actual data into useful format. In addition, Chaotic Local Search Whale Optimization Algorithm-based Feature Selection (CLSWOA-FS) technique is employed to choose the optimal feature subsets. Finally, MVO algorithm is exploited with Bidirectional Gated Recurrent Unit (BiGRU) model for classification. Here, the novelty of the work is the application of MVO algorithm in fine-turning the hyperparameters involved in BiGRU model. The experimental validation was conducted for the proposed IMVO-DLIDS model on benchmark datasets and the results were assessed under distinct

This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

measures. An extensive comparative study was conducted and the results confirmed the promising outcomes of IMVO-DLIDS approach compared to other approaches.

Keywords: Internet of everything; deep learning; feature selection; classification; intrusion detection; cybersecurity

1 Introduction

Computing and technological advancements have heavily influenced the growth and development of a country's economy as well as its society. Internet of Everything (IoE) is one of the technological developments that constitutes an interlinked network of persons, processes, data, and objects [1]. Big data and network intelligence are two prime components in IoE atmosphere that handle expandability, feasibility, and controllability of mounting advances in network connectivity [2]. IoE makes all the connections (new persons, processes, piece of data, and materials) that come online, more appropriate and beneficial though they have various security and privacy concerns [3]. Moreover, the fast development of big data from IoE applications, in return, adds remarkable value in terms of data transformation from context-aware mechanisms to prosecutable information. This information leads to a reasonable amount of impact on IoE atmosphere in terms of security and privacy perspectives [4]. The past few decades have acted as an evidence for mounting familiarity with Intrusion Detection Systems (IDSs), thanks to its intrinsic capability to find an intrusion on real-time basis [5]. Intrusion detection is referred to as a process that monitors and follows the events in a computer. It is commonly used in the identification of symbols related to security problems, while the activities are observed based on event-based methods and security information.

IDS is considered as a method to track the activities of a network among various bodies, by forecasting its integrity and existing principles [6]. A classic Intrusion Detection System comprises of data source, pre-processing, and decision-making methods to recognize the susceptible elements in a network. The first and the foremost step is the collection of raw form of data from host traces or network trafficking. The second one covers the overall construction of structures that pass on to decision-making methods which is likely to find out the hazards [7]. In favour of adapting the highly advanced network technologies and ensuring network security under distinct scenarios, the generalizing capability of the classifier requires further betterment, specifically in recognizing unknown attacks. But, the generalization capability of a single classifier is restricted while the cost of training the ensemble techniques is high [8]. In order to develop a proficient IDS model, huge volumes of data are required for training and testing purposes. The quality state of the data is highly analytical and influential in nature, which is mainly based on the outcomes of IDS design [9]. The low-quality and inappropriate information, identified in data, could be removed only after the collection of statistical property from its observable attributes and components [10].

The current study develops an Intelligent Multiverse Optimization with Deep Learning Enabled Intrusion Detection System (IMVO-DLIDS) for IoT environment. The presented IMVO-DLIDS model involves data pre-processing to convert the actual data into useful format. In addition, Chaotic Local Search Whale Optimization Algorithm-based Feature Selection (CLSWOA-FS) technique is also employed to choose the optimum feature subsets. At last, MVO algorithm is exploited with Bidirectional Gated Recurrent Unit (BiGRU) model for classification process. The novelty of the work lies in using MVO algorithm for fine-tuning the hyperparameters involved in BiGRU model.

The proposed IMVO-DLIDS model was validated experimentally using benchmark datasets and the results were assessed under distinct measures.

2 Related Works

Ullah et al. [11] presented a hybrid DL method for cyber-attack detection in IoV. The method was presented based on GRU and LSTM. The experimental results demonstrated that the presented method can accomplish high performance in terms of attack detection. Mehmood et al. [12] introduced a new methodology using NB classifier model with Intrusion Detection System (IDS). IDS was deployed as a multi-agent system throughout the network to sense irregular or misbehaving traffic and the activities of the nodes. In literature [13], the vulnerability of external and intra-vehicle networks is deliberated. A multi-tiered hybrid IDS, integrating anomaly-and-signature-based IDS, was presented to identify both known as well as unknown attacks on vehicular network. The experimental results illustrated that the presented technique can identify different kinds of known attacks. Liu et al. [14] developed a PSO-based Gradient Descent (PSO-LightGBM) for intrusion detection. The presented method was utilized to extract the features of the dataset. These features were then fed as input into one-class SVM (OCSVM) to identify and classify the malicious information.

Farzaneh et al. [15] projected an anomaly-based lightweight IDS-based threshold value for identification of attacks on RPL technique. As per the results, the presented method is highly effective in identifying the attacks and is suitable for largescale networks. In literature [16], Naïve Bayes and K-means clustering method were combined and applied to evaluate the unlabelled dataset and explore malicious attacks. Also, the study presented a solution to alleviate IoT attacks with the help of IDS and firewall. The authors [17] designed a two-tier scalable IDS for embedded systems to resolve the problem. The presented method depends on Spark and was deployed in cloud environment. The experiment results inferred that the presented technique can enhance the scalability and detection efficiency. Cheng et al. [18–20] presented a temporal convolution network with global attention model to develop an in-vehicle network IDS named TCAN-IDS. The feature extraction method extracts the spatial-temporal details.

3 The Proposed Model

In this study, an effective IMVO-DLIDS model has been developed for identification and classification of intrusions in IoT environment. At first, the proposed IMVO-DLIDS model pre-processes the actual data to convert it into a useful format. Besides, a novel CLSWOA-FS technique is employed to choose the optimum feature subsets. Finally, MVO-GRU model is applied for recognition and classification of intrusions. Fig. 1 illustrates the overall processes of IMVO-DLIDS technique.

3.1 Data Pre-processing

At primary level, z-score normalization approach is used to convert the original data into useful format. In order to normalize the data by employing z-score, the mean of populations from raw data point can be subtracted and divided by SD which offers an ideal score between -3 and $+3$. Here, x signifies the value of a certain sample, μ represents the mean and σ denotes the SD.

$$z_{-score} = \frac{(x - \mu)}{\sigma} \tag{1}$$

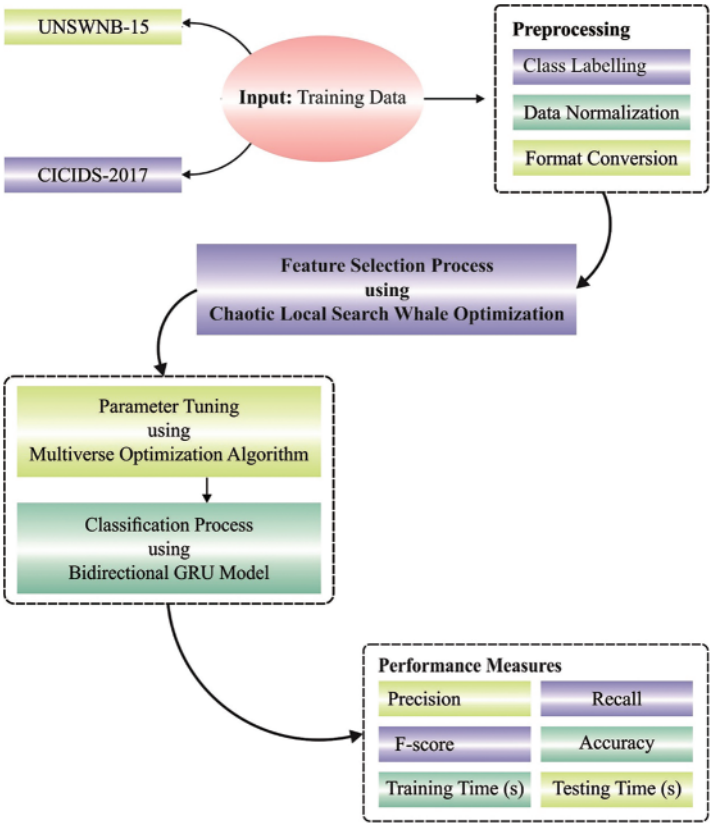

Figure 1: Overall processes of IMVO-DLIDS technique

3.2 Algorithmic Process of CLWOA-FS Technique

After pre-processing the data, CLWOA-FS technique is applied to derive the optimal feature subsets [21–23]. WOA comprises of shrink envelop, exploration, and exploitation. This section has three more sub-sections which detail about the processes involved and equivalent arithmetical model. Consider that the amount of whale populations that contribute to predation is N and the dimension is d. Those whale individuals, in the existing optimum location, are fixed as searching agents whereas other individuals upgrade their position within the region, in which the searching agent is situated. It is mathematically expressed as follows.

$$D = \left| C\vec{X} \ * \ (t) - X(t) \right| \tag{2}$$

$$\vec{X}\,(t+1) = \vec{X} \ * \ (t) - A \cdot \vec{D} \tag{3}$$

whereas t denotes the existing number of iterations and A and C indicate the coefficients. $\vec{X} \ * \ (t)$ denotes the location vector of searching agent, and $\vec{X}(t)$ denotes the location vector of the present whale. Both A and C coefficients are attained using the formulae given below.

$$\vec{A} = 2\vec{a}.\vec{r}_1 - \vec{a} \tag{4}$$

$$C = 2r_2 \tag{5}$$

$$a = 2 - 2t/T_{\max} \tag{6}$$

Here r_1 and r_2 denote the arbitrary values within $(0, 1)$. t indicates the existing number of iterations and $T_{\max}$ indicates the maximal number of iterations. During exploitation stage, the whale groups are defined through the existing objective. The whole population spirally upgrades its location to attack the prey. It is arithmetically expressed as given herewith.

$$\vec{X}(t+1) = \vec{X} \ * \ (t) + D_p e^{bl} \cos(2\pi l) \tag{7}$$

$$D_p = \left| \vec{X} \ * \ (t) - \vec{X}(t) \right| \tag{8}$$

Now D_p indicates the distance between whale and the prey, $\vec{X} \ * \ (t)$ denotes the existing location vector of optimal whale whereas $\vec{X}(t)$ denotes the existing location vector. l denotes an arbitrary value within $(-1, 1)$. In developmental stage, the whale population spirally upgrades its location while enclosing its prey. By following a synchronous behavior method and by mimicking the real-time attack procedure, the approach considers that the possibility of the whale population selecting the spiral upgrade location is similar to the possibility of selecting the shrinking encircling process, viz., $p = 0.5$. It can be mathematically expressed as follows.

$$\vec{X}(t+1) = \begin{cases} \vec{X} \ * \ (t) - A \cdot D & P < 0.5 \\ \vec{X} \ * \ (t) + D_p e^{bl} \cos(2\pi l) & P \geq 0.5 \end{cases} \tag{9}$$

During exploitation stage, when a differs, the value of A lies within [1,1]. Specifically, the whale individuals stay at other locations than the existing location. In order to seek a searching space, the approach arbitrarily choses an individual whale as the searching agent to guide the present whale individual. It assists the whale population to seek distance from the present location. While the above scenario is mathematically expressed as follows.

$$D = \left| C \ * \ \vec{X}_{rand} - \vec{X}(t) \right| \tag{10}$$

$$\vec{X}(t+1) = \vec{X} - A \cdot D \tag{11}$$

Now $\vec{X}_{rand}$ indicates the place vector of the present searching agent. It is significant to observe that the presence of this phase assists the approach in avoiding local optimal efficiency. CLWOA is derived from the application of chaotic concepts to improve the performance of WOA.

Chaos is otherwise called as a random-like phenomenon that is established in non-linear and deterministic schemes. Arithmetically, chaos implements the searching process rapidly than the ergodic searching. Massive amounts of series are attained by altering the primary value. In current study, logistic map is utilized in the generation of chaotic series as given below.

$$o^{s+1} = Co^s(1 - 0^s) \tag{12}$$

The initialized variable are $C = 4, 0^s = rand(0, 1)$, and $C_1 \neq 0.25, 0.5$ and 0.75. Even though the execution time is satisfactory on a small-scale for chaos optimization, once the searching space becomes huge, the execution time remains unsatisfactory. The chaotic scheme features are taken to

attain a searching operator, while the operator is incorporated with metaheuristic algorithm. The solution that generates CLS is attained as follows.

$$Cs = (1 - \mu) \times T + \mu \grave{C}_i, i = 1, 2, \ldots, n \tag{13}$$

whereas Cs denotes the candidate solution, T indicates the targeted location and μ can be attained as given below.

$$\mu = \frac{\text{Max } Iter - currIter + 1}{\text{Max}_i ter} \tag{14}$$

Now, *MaxIter* and *currIter* denote the Max amount of iterations and the existing iteration correspondingly. To map $\grave{C}$ into the field, 0.

$$\grave{C} = LB + C_i \times (UB - LB) \tag{15}$$

Here, UB and LB indicate the upper and lower bounds of the initialized solution.

The aim of CLWOA-FS technique is to identify the optimum set of features for a given data set that has minimal features and high classification accuracy. These two indicators exert a distinct effect upon classification accuracy. So, both are integrated together using a single weighted indicator whereas a similar Fitness Function (FF) is employed as follows.

$$fitness = \omega_1 \times acc(classifier) + \omega_2 \times \left(1 - \frac{s}{p}\right) \tag{16}$$

Now s characterizes the quantity of the selected features, and p represents the entire count of features. Here, ω_1 and ω_2 values are 1 and 0.001, correspondingly. $acc\,(classifer)$ signifies the classification accuracy obtained from BiGRU classification model as shown below.

$$acc(classifier) = \frac{n_c}{n_c + n_i} \times 100\% \tag{17}$$

Now, n_c and n_i denote the number of accurately-and inaccurately-classified instances, correspondingly. The fitness value achieves the objective i.e., the selected feature has maximum classification accuracy along with minimal amount of features.

3.3 BiGRU Based Classification

In this study, intrusions are recognized and classified with the help of BiGRU model [21]. GRU employs two gated components to adjust the cell state. It has various advantages such as less parameters, low computation difficulty and good efficacy compared to LSTM model in NLP. Especially, during time t, embedded vector $w_t \in \mathbb{R}^{d_w}$ is attained for the existing input vector E and the aspect embedded vector $v_a \in \mathbb{R}^{d_a}$ from A, the existing hidden state vector h_t in GRU can be upgraded as follows.

$$\begin{aligned}
z_t &= \sigma(W_z h_{t-1} + U_z[w_t, v_a] + b_z), \\
r_t &= \sigma(W_r h_{t-1} + U_r[w_t, v_a] + b_r) \\
\widetilde{h}_t &= \tanh(W_h(h_{t-1} \odot r_t) + U_h[w_t, v_a] + b), \\
h_t &= h_{t-1} \odot (1 - z_t) + z_t \odot \widetilde{h}_t,
\end{aligned} \tag{18}$$

whereas z and r denote the update and reset gates, correspondingly; the sigmoid function $\sigma(\cdot)$ is utilized to control the preservation of effective data and remove the redundant data; W_z, W_r, $W_h \in \mathbb{R}^{d_h \times d_h}$, U_z, U_r, $U_h \in \mathbb{R}^{d_h \times (d_w + d_a)}$, b_z, b_r, $b \in \mathbb{R}^{d_h}$ represents the weight and bias learned in GRU training method; $\odot$ represents elementwise multiplication; $[w_t, v_a]$ denotes the splicing vector of the embedding word w_t and the embedding aspect v_a. Fig. 2 illustrates the framework of BiGRU.

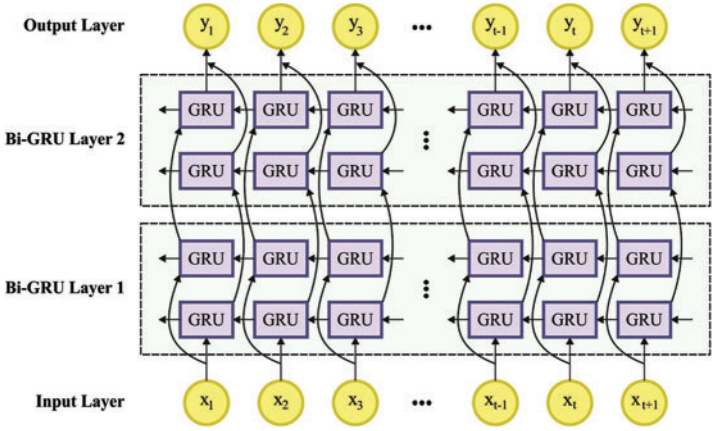

Figure 2: Structure of BiGRU

Later, the hidden layer $[h_1, h_2, \ldots, h_N]$ of the sentence with length N is considered by the last contextual word depiction. Here, BiGRU has been adapted to acquire the context depiction of a sentence. In comparison to one-way GRU, BiGRU involves backward hidden state $\overleftarrow{h_i^t} \in \mathbb{R}^{d_h}$ and forward hidden state $\vec{h_i^t} \in \mathbb{R}^{d_h}$ during time t, whereas d_h characterizes the number of hidden layers. Next, the backward hidden layer $\overleftarrow{h_i^t}$ and forward hidden layer $\vec{h_i^t}$ are interconnected with last contextual hidden depiction $h_i^t = [\vec{h_i^t}; \overleftarrow{h_i^t}] \in \mathbb{R}^{2d_h}$.

3.4 MVO Based Hyperparameter Optimization

In this final stage, MVO algorithm is introduced as a hyperparameter optimizer for BIGRU model [24]. MVO approach is inspired by the concepts that theoretically exist in astronomy. It includes white holes that form a major component in the creation of universe and no one has observed it in the whole universe. The abovementioned process depends on evolving population. Multiple candidate solutions assist one another and share data among themselves to move towards the promising areas. In order to integrate the solution, black and white holes are arbitrarily generated in universe and the movement of objects is created. MVO employs both black as well as white holes to exploit the searching space, whereas it employs wormholes to explore the searching space.

The component of object exchange, throughout the universe, is that the higher-inflation universe often tries to discard the objects and transmits them into receiving universe with lower inflation. Eventually, the inflation rate in every universe is balanced and remains in a stable state. In this method, the universe is initialized and arranged, according to inflation as given in the following equation.

$$X_i^j = \begin{cases} \begin{cases} \chi_j + TDR * r1 & (rd1) < WEP \\ x_j - TDR * r1 & (rd1) \geq WEP \end{cases} & (rd2) < WEP, \\ X_i^j & (rd2) \geq WEP \end{cases} \tag{19}$$

Here x_j represents the *jth* unit in the formed universe; x_i^j specifies the j^{th} unit from i^{th} universe; Worm hole Existence Probability (WEP) and Traveling Distance Rate (TDR) are two major coefficients; the values of $rd1$, $rd2$ and rl denote the random variables within [0, 1]. It is described as follows.

$$TDR = 1 - \left(\frac{l^{\frac{1}{p}}}{L}\frac{1}{p}\right) \tag{20}$$

$$WEP = \min - l * \left(\frac{\min - \max}{L}\right) \tag{21}$$

whereas $p(= 6)$ refers to the accuracy of exploitation over iteration, max indicates the upper limits of WEP and min represents the lower limit of WEP, l refers to the present iteration and L indicates the maximal number of iterations.

MVO technique, for hyper-parameter optimization, calculates FF as the minimized classifier error rate which is demonstrated below. An optimal solution holds lesser error value and conversely.

$$fitness\,(x_i) = Classifier\ Error\ Rate\,(x_i) = \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \tag{22}$$

4 Experimental Validation

In this section, a detailed investigation was conducted to validate the intrusion detection performance of the proposed method using CICIDS-2017 benchmark datasets. At first, IMVO-DLIDS model selected a total of 47 features out of 80 features from CICIDS-2017 dataset.

Fig. 3 illustrates the confusion matrices generated by IMVO-DLIDS model on test CICIDS 2017 dataset. The figure report that the proposed IMVO-DLIDS model classified all the samples under seven class labels effectively.

Tab. 1 and Fig. 4 highlight the overall classification outcomes accomplished by IMVO-DLIDS model on CICIDS-2017 dataset. The experimental values indicate that the proposed IMVO-DLIDS model produced effectual classification outcomes under distinct aspects. For instance, on entire dataset, IMVO-DLIDS model produced $accu_y$, $reca_l$, $prec_n$, and F_{score} values such as 99.71%, 83.01%, 90.08%, and 86.17% respectively. Also, on 70% of TRS dataset, the presented IMVO-DLIDS model yielded $accu_y$, $reca_l$, $prec_n$, and F_{score} values such as 99.71%, 83.01%, 90.08%, and 86.17% respectively. In addition, on 30% of TSS dataset, the proposed IMVO-DLIDS model accomplished $accu_y$, $reca_l$, $prec_n$, and F_{score} values such as 99.71%, 83.01%, 90.08%, and 86.17% respectively.

The results from Training Accuracy (TA) and Validation Accuracy (VA), attained by IMVO-DLIDS model, on CICIDS-2017 dataset are demonstrated in Fig. 5. The experimental outcomes imply that the proposed IMVO-DLIDS model gained maximum TA and VA values. To be specific, VA seemed to be higher than TA.

The results from Training Loss (TL) and Validation Loss (VL), achieved by the proposed IMVO-DLIDS model on CICIDS-2017 dataset, are showcased in Fig. 6. The experimental outcomes infer that the proposed IMVO-DLIDS model accomplished the least TL and VL values. To be specific, VL seemed to be lower than TL.

A brief precision-recall analysis was conducted upon IMVO-DLIDS model using CICIDS-2017 dataset and the results are portrayed in Fig. 7. From the figure, it can be understood that IMVO-DLIDS model accomplished the maximum precision-recall performance under all classes.

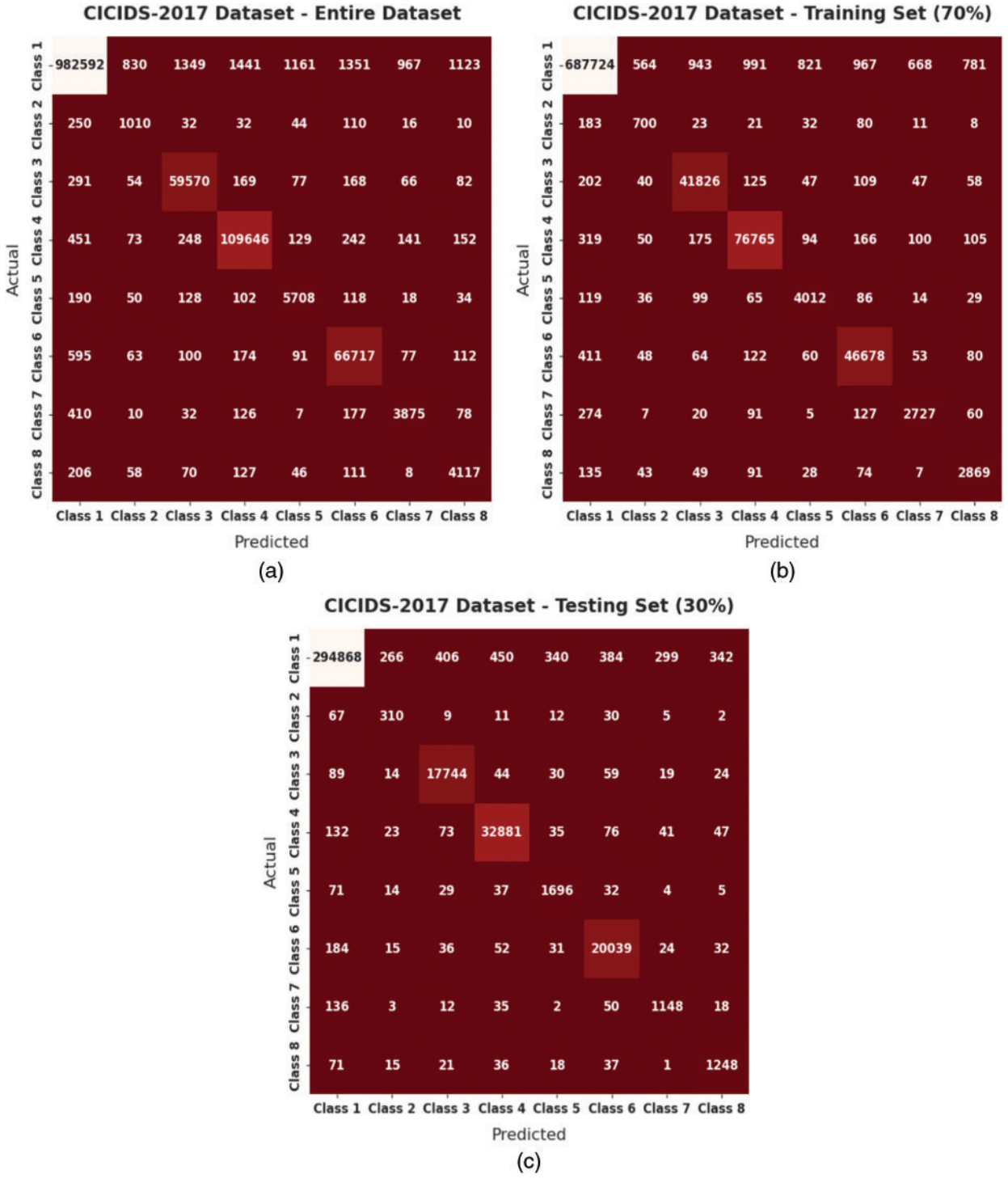

Figure 3: Confusion matrix of IMVO-DLIDS technique on CICIDS 2017 dataset

Table 1: Results of the analysis of IMVO-DLIDS technique under distinct measures on CICIDS-2017 dataset

Class labels	Accuracy	Precision	Recall	F-score
Entire dataset				
Class 1	99.15	99.76	99.17	99.46
Class 2	99.87	47.02	67.15	55.31
Class 3	99.77	96.82	98.5	97.65
Class 4	99.71	98.06	98.71	98.38
Class 5	99.82	78.59	89.92	83.87
Class 6	99.72	96.7	98.22	97.45
Class 7	99.83	74.98	82.18	78.42
Class 8	99.82	72.13	86.8	78.79
Average	99.71	83.01	90.08	86.17
Training set (70%)				
Class 1	99.15	99.76	99.17	99.46
Class 2	99.87	47.02	67.15	55.31
Class 3	99.77	96.82	98.5	97.65
Class 4	99.71	98.06	98.71	98.38
Class 5	99.82	78.59	89.92	83.87
Class 6	99.72	96.7	98.22	97.45
Class 7	99.83	74.98	82.18	78.42
Class 8	99.82	72.13	86.8	78.79
Average	99.71	83.01	90.08	86.17
Testing set (30%)				
Class 1	99.15	99.76	99.17	99.46
Class 2	99.87	47.02	67.15	55.31
Class 3	99.77	96.82	98.5	97.65
Class 4	99.71	98.06	98.71	98.38
Class 5	99.82	78.59	89.92	83.87
Class 6	99.72	96.7	98.22	97.45
Class 7	99.83	74.98	82.18	78.42
Class 8	99.82	72.13	86.8	78.79
Average	99.71	83.01	90.08	86.17

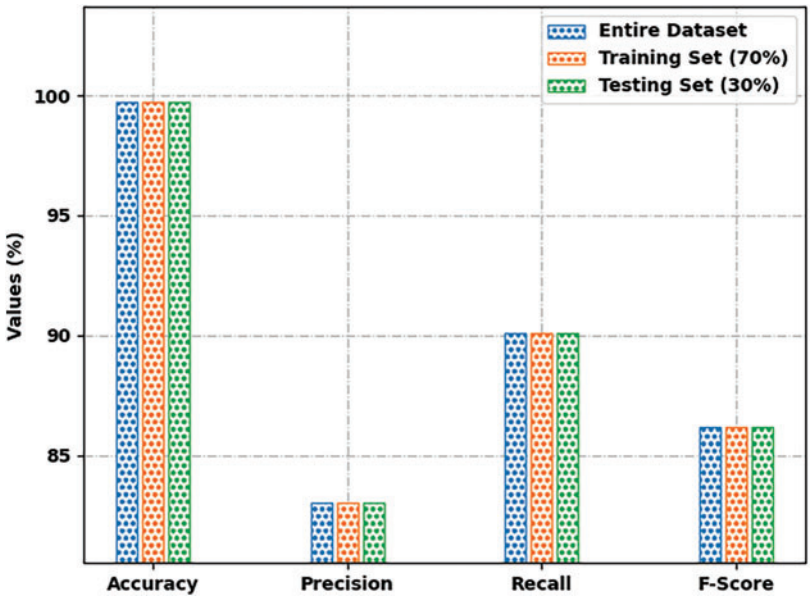

Figure 4: Results of the analysis of IMVO-DLIDS technique on CICIDS-2017 dataset

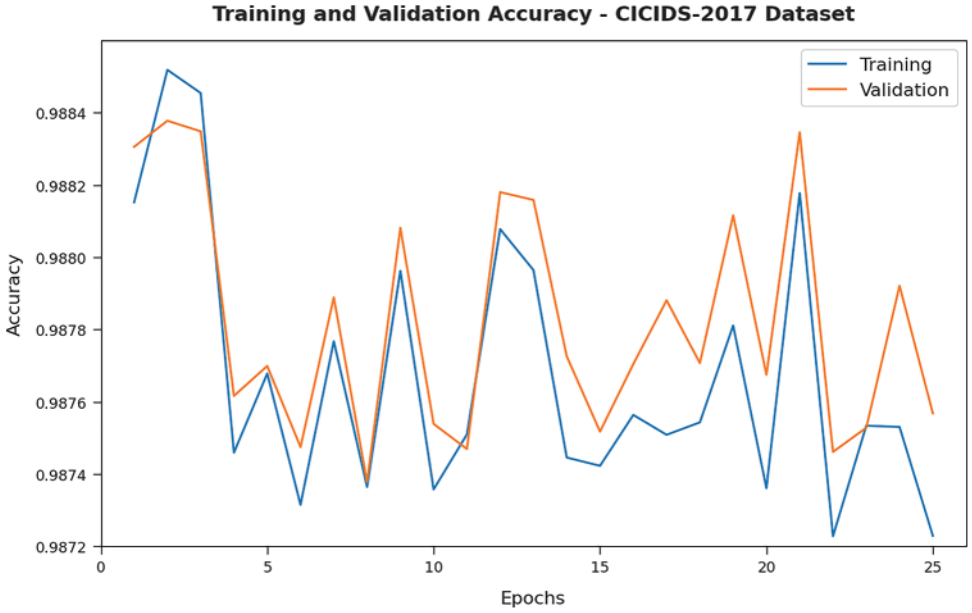

Figure 5: TA and VA analyses results of IMVO-DLIDS technique on CICIDS-2017 dataset

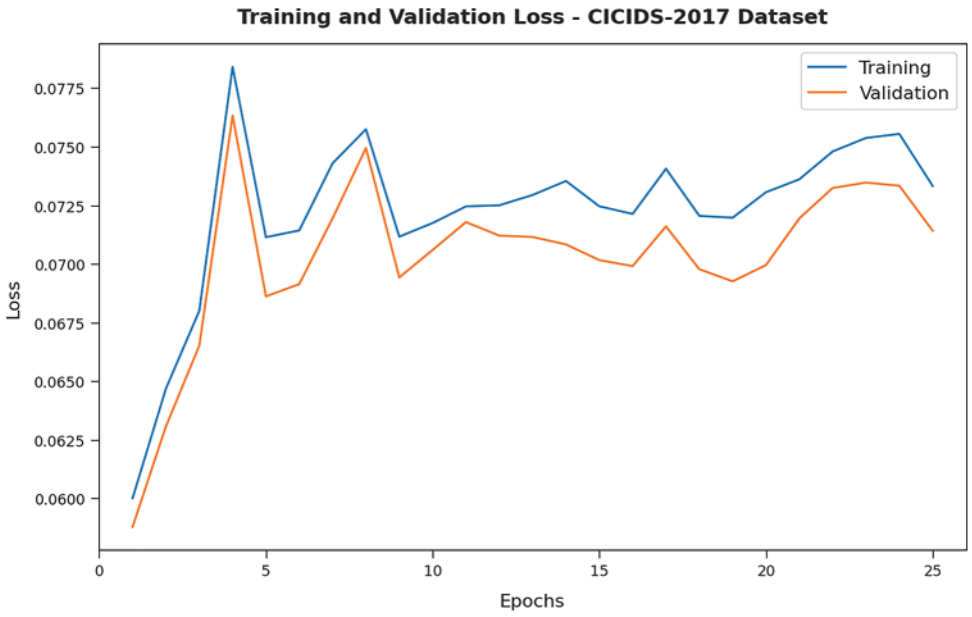

Figure 6: TL and VL analyses results of IMVO-DLIDS technique on CICIDS-2017 dataset

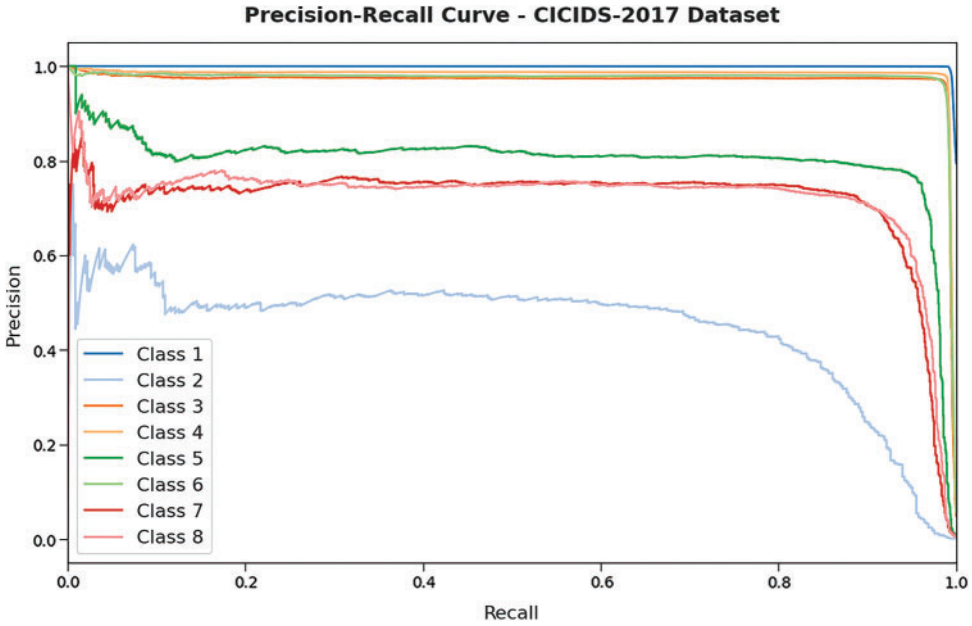

Figure 7: Precision-recall analysis results of IMVO-DLIDS technique on CICIDS-2017 dataset

Tab. 2 and Fig. 8 show the comparative study results achieved by IMVO-DLIDS and other recent models on CICIDS-2017 dataset [25]. The experimental results imply that EM and DBSCAN methods exhibited the least classification performance. KODE system tried to achieve a slightly improved performance with $accu_y$, $reca_l$, $prec_n$, and F_{score} values such as 88.51%, 88.54%, 81.84%, and 88.91% respectively. Though K-means and one-class SVM models accomplished reasonable performances,

the proposed IMVO-DLIDS model achieved the maximum performance with $accu_y$, $reca_l$, $prec_n$, and F_{score} values such as 99.71%, 90.08%, 83.01%, and 86.17% respectively.

Table 2: Comparative analysis results of IMVO-DLIDS and other existing algorithms on CICIDS-2017 dataset

Methods	Accuracy	Recall	Precision	F-score
IMVO-DLIDS	99.71	90.08	83.01	86.17
K-mean algorithm	98.15	88.31	78.49	85.02
One-Class SVM model	95.96	81.09	82.19	84.42
DBSCAN algorithm	81.70	82.72	80.97	81.40
EM algorithm	78.94	77.94	78.30	79.58
KODE model	88.51	88.54	81.84	88.91

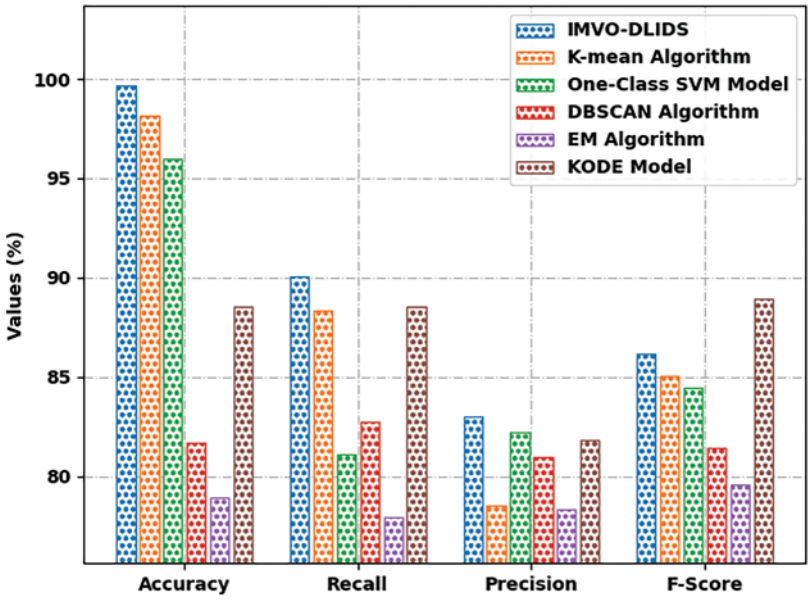

Figure 8: Comparative analysis results of IMVO-DLIDS technique on CICIDS-2017 dataset

Brief Training Time (TRT) and Testing Time (TST) investigation were conducted between IMVO-DLIDS model and the existing models while the results are shown are given in Tab. 3 and Fig. 9. The results infer that K-means, one-class SVM, and KODE model showcased poor outcomes with maximum TRT and TST values. Followed by, DBSCAN approach achieved somewhat higher outcomes with a TRT of 186 s and a TST of 32.10 s. Next, EM model resulted in a considerable TRT of 119 s and a TST of 31.20 s. However, the proposed IMVO-DLIDS model outperformed all other methods with the least TRT of 31.24 s and a TST of 29.38 s respectively.

Table 3: Training and testing time analyses results of IMVO-DLIDS technique and other existing approaches on CICIDS-2017 dataset

Methods	Training time (s)	Testing time(s)
IMVO-DLIDS	31.24	29.38
K-mean algorithm	400.40	3662.90
One-Class SVM model	37.00	4882.80
DBSCAN algorithm	186.00	32.10
EM algorithm	119.00	31.20
KODE model	223.10	4961.00

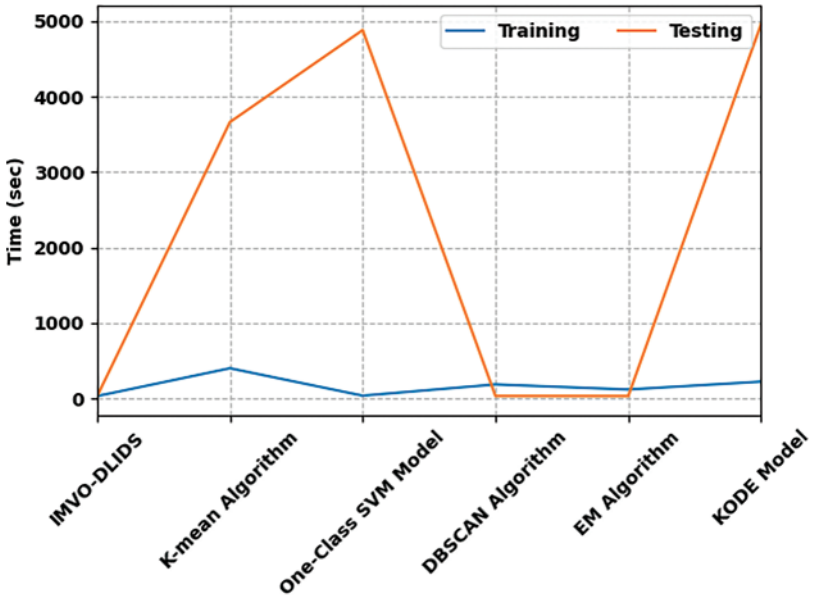

Figure 9: TRT and TST analyses results of IMVO-DLIDS technique on CICIDS-2017 dataset

5 Conclusion

In this study, an effectual IMVO-DLIDS model has been developed for identification and classification of intrusions in IoT environment. The proposed IMVO-DLIDS technique primarily carries out data pre-processing to convert the actual data into useful format. Besides, a novel CLSWOA-FS technique is employed in the selection of optimum feature subsets. Finally, MVO-GRU technique is implemented for recognition and classification of intrusions. The proposed IMVO-DLIDS method was experimentally validated on benchmark datasets and the results were assessed under distinct measures. An extensive comparative study was conducted and the results attained by IMVO-DLIDS method were promising than the compared approaches under different aspects. Thus, IMVO-DLIDS technique can be utilized as an effectual tool for intrusion detection in IoT environment. In future, outlier removal approaches can also be included to improve the detection efficiency of IMVO-DLIDS model.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through Large Groups Project under grant number (46/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R140), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4210118DSR13).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] D. E. Comer, *The Internet book: Everything you need to know about computer networking and how the Internet works*. UK, Chapman and Hall/CRC, 2018.

[2] S. P. Mohanty, V. P. Yanambaka, E. Kougianos and D. Puthal, "PUFchain: A hardware-assisted blockchain for sustainable simultaneous device and data security in the internet of everything (IoE)," *IEEE Consumer Electronics Magazine*, vol. 9, no. 2, pp. 8–16, 2020.

[3] D. Zhang, J. Hu, F. Li, X. Ding, A. K. Sangaiah *et al.,* "Small object detection via precise region-based fully convolutional networks," *Computers, Materials and Continua*, vol. 69, no. 2, pp. 1503–1517, 2021.

[4] J. Wang, Y. Wu, S. He, P. K. Sharma, X. Yu *et al.,* "Lightweight single image super-resolution convolution neural network in portable device," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15, no. 11, pp. 4065–4083, 2021.

[5] J. Wang, Y. Zou, P. Lei, R. S. Sherratt and L. Wang, "Research on recurrent neural network based crack opening prediction of concrete dam," *Journal of Internet Technology*, vol. 21, no. 4, pp. 1161–1169, 2020.

[6] J. Zhang, J. Sun, J. Wang and X. G. Yue, "Visual object tracking based on residual network and cascaded correlation filters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 8, pp. 8427–8440, 2021.

[7] S. He, Z. Li, Y. Tang, Z. Liao, F. Li *et al.,* "Parameters compressing in deep learning," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 321–336, 2020.

[8] S. R. Zhou and B. Tan, "Electrocardiogram soft computing using hybrid deep learning CNN-ELM," *Applied Soft Computing*, vol. 86, no. 4, pp. 105778, 2020.

[9] M. Nasir, A. R. Javed, M. A. Tariq, M. Asim and T. Baker, "Feature engineering and deep learning-based intrusion detection framework for securing edge IoT," *Journal of Supercomputing*, vol. 78, no. 6, pp. 8852–8866, 2022.

[10] A. R. Javed, S. ur Rehman, M. U. Khan, M. Alazab and G. T.R., "CANintelliIDS: Detecting in-vehicle intrusion attacks on a controller area network using CNN and attention-based GRU," *IEEE Transactions on Network Science and Engineering*, vol. 8, no. 2, pp. 1456–1466, 2021.

[11] S. Ullah, M. A. Khan, J. Ahmad, S. S. Jamal, Z. Huma *et al.,* "HDL-IDS: A hybrid deep learning architecture for intrusion detection in the internet of vehicles," *Sensors*, vol. 22, no. 4, pp. 1340, 2022.

[12] A. Mehmood, M. Mukherjee, S. H. Ahmed, H. Song and K. M. Malik, "NBC-MAIDS: Naïve bayesian classification technique in multi-agent system-enriched ids for securing IoT against DDOS attacks," *The Journal of Supercomputing*, vol. 74, no. 10, pp. 5156–5170, 2018.

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

[14] J. Liu, D. Yang, M. Lian and M. Li, "Research on intrusion detection based on particle swarm optimization in IoT," *IEEE Access*, vol. 9, pp. 38254–38268, 2021.

[15] B. Farzaneh, M. A. Montazeri and S. Jamali, "An anomaly-based ids for detecting attacks in RPL-based internet of things," in *2019 5th Int. Conf. on Web Research (ICWR)*, Tehran, Iran, pp. 61–66, 2019.

[16] R. K. Shrivastava, S. Ramakrishna and C. Hota, "Game theory based modified naïve-bayes algorithm to detect DOS attacks using honeypot," in *2019 IEEE 16th India Council Int. Conf. (INDICON)*, Rajkot, India, pp. 1–4, 2019.
[17] M. Liu, Z. Xue and X. He, "Two-tier intrusion detection framework for embedded systems," *IEEE Consumer Electronics Magazine*, vol. 10, no. 5, pp. 102–108, 2021.
[18] P. Cheng, K. Xu, S. Li and M. Han, "TCAN-IDS: Intrusion detection system for internet of vehicle using temporal convolutional attention network," *Symmetry*, vol. 14, no. 2, pp. 310, 2022.
[19] M. A. Alohali, F. N. Al-Wesabi, A. M. Hilal, S. Goel, D. Gupta *et al.,* "Artificial intelligence enabled intrusion detection systems for cognitive cyber-physical systems in industry 4.0 environment," *Cognitive Neurodynamics*, vol. 42, no. 3, pp. 16, 2022.
[20] A. M. Hilal, M. A. Alohali, F. N. Al-Wesabi, N. Nemri, J. Hasan *et al.,* "Enhancing quality of experience in mobile edge computing using deep learning based data offloading and cyberattack detection technique," *Cluster Computing*, vol. 76, no. 4, pp. 2518, 2021.
[21] M. Li, G. Xu, Q. Lai and J. Chen, "A chaotic strategy-based quadratic opposition-based learning adaptive variable-speed whale optimization algorithm," *Mathematics and Computers in Simulation*, vol. 193, no. 8, pp. 71–99, 2022.
[22] A. G. Hussien and M. Amin, "A self-adaptive Harris Hawks optimization algorithm with opposition-based learning and chaotic local search strategy for global optimization and feature selection," *International Journal of Machine Learning and Cybernetics*, vol. 13, no. 2, pp. 309–336, 2022.
[23] X. Li, L. Ding, Y. Du, Y. Fan and F. Shen, "Position-enhanced multi-head self-attention based bidirectional gated recurrent unit for aspect-level sentiment classification," *Frontiers in Psychology*, vol. 12, pp. 1–11, 2022.
[24] I. Benmessahel, K. Xie and M. Chellal, "A new competitive multiverse optimization technique for solving single-objective and multiobjective problems," *Engineering Reports*, vol. 2, no. 3, pp. e12124, 2020.
[25] M. S. A. Daweri, K. A. Z. Ariffin, S. Abdullah and M. F. E. M. Senan, "An analysis of the KDD99 and UNSW-NB15 datasets for the intrusion detection system," *Symmetry*, vol. 12, no. 10, pp. 1666, 2020.