IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 May 2024, accepted 3 June 2024, date of publication 10 June 2024, date of current version 21 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3411612

RESEARCH ARTICLE

Enhancing Cyberattack Detection Using Dimensionality Reduction With Hybrid Deep Learning on Internet of Things Environment

SALAHALDEEN DURAIBI[1] AND ABDULLAH MUJAWIB ALASHJAEE[2], (Member, IEEE)
[1]Department of Electrical and Electronics Engineering, College of Engineering and Computer Science, Jazan University, Jazan 45142, Saudi Arabia
[2]Department of Computer Science, College of Science, Northern Border University, Arar 73213, Saudi Arabia

Corresponding author: Salahaldeen Duraibi (sduraibi@jazanu.edu.sa)

The authors gratefully acknowledge the funding of the Deanship of Graduate Studies and Scientific Research, Jazan University, Saudi Arabia, through Project Number: GSSRD-24.

ABSTRACT Cybersecurity in the Internet of Things (IoT) ecosystem is vital to protect sensitive data, stop unauthorized access, and mitigate the risk of disruptive cyberattacks. Cyberattack recognition utilizing an intrusion detection system (IDS) is a major imperative given the increase in the number of connected devices. Advanced cybersecurity methods deploy machine learning (ML) approaches, anomaly detection, and behavioral analysis to analyze IoT network traffic for irregular patterns indicative of potential cyberattacks. The combination of feature selection (FS) and deep learning (DL) approaches in cyberattack recognition suggests a proactive and sophisticated manner to bolster cybersecurity. Leveraging DL structures like neural networks (NNs) assists the automatic extraction and analysis of intricate patterns in the difficult IoT data landscape. This paper develops an Improved Mayfly Optimization Algorithm with a Hybrid Deep Learning based Intrusion Detection (IMFOHDL-ID) approach in IoT environments. The designed IMFOHDL-ID approach's main goal is to classify intrusions and accomplish security in the IoT environment. The IMFOHDL-ID technique initially follows data normalization as a preprocessing stage. In addition, the IMFOHDL-ID technique makes use of the IMFO-based feature selection (FS) method to elect feature subsets. For IDs, the IMFOHDL-ID technique applies the Long Short Term Memory based Deep Stacked Sequence-to-Sequence Autoencoder (LSTM-DSSAE) model. Finally, the dipper-throated optimization algorithm (DTOA) was utilized for optimal hyperparameter selection of the LSTM-DSSAE method. To highlight better results of the IMFOHDL-ID model, a series of simulation analyses were performed. Extensive comparative results stated the improved outcome of the IMFOHDL-ID technique over existing approaches.

INDEX TERMS Cybersecurity, intrusion detection system, feature selection, hybrid deep learning, Internet of Things.

I. INTRODUCTION

In the current scenario, the Internet of Things (IoT) become peak significant research topic. IoT is a new technical model that is defined as a worldwide web of connected electronic gadgets [1]. Its main goal is to enhance every daily life through programming usual everyday processes on all sides of the lifespan without human involvement. Several devices linked to IoT have been outstretched knowingly and a high increase in attacks against IoT gadgets has complemented this development [2]. Security worries regarding the effect of these attacks on linked devices have improved. Additionally, the sensitivity of data accessible on IoT devices was highly essential to locate solutions for detecting and reacting to these attacks. Due to its weaknesses, IoT is helpless to attacks as well as security threats [3]. Many researchers tried to classify

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Yuan Chen.

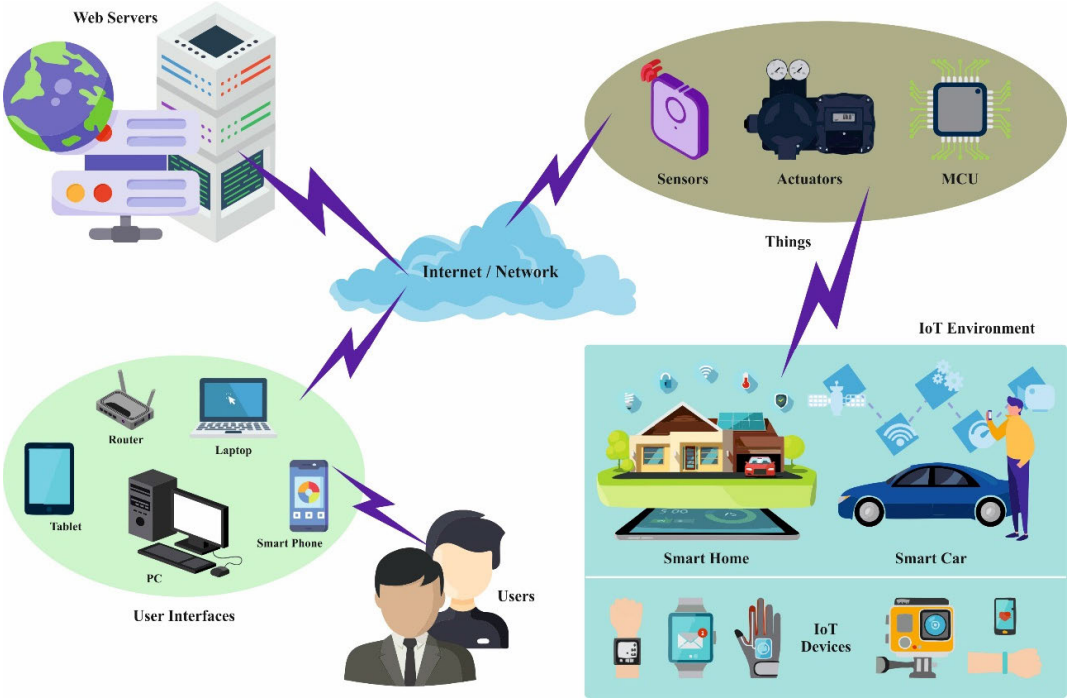

FIGURE 1. Structure of IoT.

assaults, security, vulnerabilities anxieties on IoT so that researchers can able to identify answers easily. For instance, according to IoT architecture layers, the researchers classified the exposures as well as physical safety strengthening as missing [4]. Fig. 1 represents the general structure of IoT.

There are many concerns for IoT devices such as lack of simplicity and device management, uncertain data storage and transfer, insecure passcodes, network interfaces, AI-based attacks, and botnets [5]. Some researchers highlighted IoT's weaknesses and security risks. The researchers mainly pointed out that because IoT uses customary network design, it receives faults. Additionally, growth in terminal devices (i.e. end nodes) with restricted processing abilities is the most important and great vulnerability broken by attackers [6].

Even though there are a few safety measures in location, IoT systems are helpless to many attacks due to their huge attack surface [7]. Therefore, it is essential to plan defense devices to detect attacks. Another defense must be proposed for defending IoT systems besides cyber-attacks. An Intrusion Detection System (IDS) is one of the effective models that will fulfill this purpose. Numerous surveys tried to define Machine Learning (ML)-based IDS for security beside IoT systems [8]. Many studies have done research work on IDSs for cyber-physical systems, Wireless sensor networks, cloud-based IoT systems, and mobile ad hoc networks (MANET). Moreover, customary IDS models are less effective or inadequate for the safety of IoT networks due to their abnormal features, which are revealed above, mainly abundant, heterogeneity, limited energy, restricted bandwidth ability, and overall connectivity [9]. Deep Learning (DL) and ML systems currently gained huge popularity in effective usage for the recognition of system threats containing IoT devices. This is because ML and DL-based techniques can seizure benign as well as irregular behavior in an IoT atmosphere. To learn ordinary patterns, IoT systems, and net traffic can be taken as well as examined. Any deviance from these usual learned designs is utilized to identify abnormal behavior. Moreover, ML/DL approaches have been verified for predicting novel attacks [10]. Therefore, ML/DL methods produce robust safety procedures to design the safety of IoT networks and devices.

This paper presents an Improved Mayfly Optimization Algorithm with a Hybrid Deep Learning based Intrusion Detection (IMFOHDL-ID) method in an IoT environment. IMFOHDL-ID technique initially follows data normalization as a preprocessing stage. In addition, the IMFOHDL-ID technique makes use of the IMFO-based feature selection (FS) model to elect feature subsets. For IDs, the IMFOHDL-ID technique applies the Long Short Term Memory based Deep Stacked Sequence-to-Sequence Autoencoder (LSTM-DSSAE) model. Finally, the dipper-throated optimization algorithm (DTOA) was utilized for optimal hyperparameter selection of the LSTM-DSSAE method. To highlight the better performance of the IMFOHDL-ID model, a series of simulation analyses were executed. Extensive comparative results stated an improved result of the IMFOHDL-ID technique over existing approaches.

II. RELATED WORKS

In [11], a novel DL-based ID is designed in this paper for IoT networks. This quick method employs 4 layer layer-deep fully connected framework for identifying mischievous

traffic that poses threats to associated IoT mechanisms. The designed network is presented as a communication protocol-independent method for decreasing utilization difficulties. Saba et al. [12] designed a model based on a convolutional neural network (CNN) for anomaly-based IDs that profits benefit from IoT's influence, producing the potential to professionally observe entire traffic through IoT. The developed technology shows the capability to distinguish any probable intrusion as well as abnormal traffic actions. In [13], the LSTM-based technique is designed to discover network attacks employing software-defined network (SDN)-supported IDs in IoT systems. The study projects a wide performance estimation of ML and DL methods in a dual SDNIoT-focused dataset. An LSTM-based framework is also developed for the current multi-class detection of network outbreaks in IoT devices.

Ramaiah et al. [14] develop novel IDs for the recognition of malicious threats in a smart atmosphere. The developed Intrusion detection model uses a device of correlation as well as a random forest (RF) model to discover the main independent variables to enhance the neural-based attack detection algorithm. For identifying a mischievous threat, shallow neural networks (SNN) and an improved neural-based detection algorithm were proposed. Ravi et al. [15] develop an end-to-end employing DL-based recurrent methods. The presented method removes features of concealed layers of recurrent methods as well as uses Kernel-based Principal Component Analysis (KPCA) FS approaches. At last, optimum features of the recurrent technique are merged and then detection is completed by employing an ensemble meta-classification algorithm.

In [16], a novel IDS method based on the grouping of DL and optimization techniques is developed. Primary, an extraction of feature models based on CNNs is proposed. Next, a novel FS process is utilized depending on a modified form of Growth Optimizer (GO), which is known as MGO. To enhance the search procedure of GO, the Whale Optimization Algorithm (WOA) is also employed. Wang et al. [17] develop a DL-based bidirectional LSTM (BiLSTM) trivial IoT intrusion detection method. The BiLSTMs and deep neural networks (DNNs) methods are united for feature extraction. The Incremental Principal Component Analysis (IPCA) model employed for feature dimensionality decrease. In addition to that, dynamic quantization is used. The authors in [18] developed a new LSTM based intrusion detection approach with Dynamic Access Control (DAC) model to detect and defends against intrusion. The DAC approach defends further intrusions from the similar source by blocking it for periods related with the number of intrusions.

Liu and Du [19] offer a genetic algorithm (GA)-based FS solution for IDS in an IoT technology that is faced with threats from botnet attacks. Alosaimi and Almutairi [20] present a new technology in combination with DL and three-level approaches to quickly and accurately identify attacks in IoT environments.

In the context of cyberattack detection in IoT using DL, a crucial research gap exists in enhancing FS and parameter tuning to optimize computational efficiency and detection performance. Regardless of the progress achieved in DL models, many researchers often overlook the crucial role of fine-tuning hyperparameters and selecting meaningful features, resulting in diminished efficiency. Incorporating metaheuristic approaches, such as particle swarm optimization (PSO), chaotic mayfly optimization (MFO), and GA, presents an effective strategy for addressing these gaps. This algorithm can effectively navigate the complex search space to detect hyperparameter configurations and optimum feature subsets, thus reducing resource consumption and enhancing model robustness in resource-constrained IoT environments. Exploring the synergy between DL and metaheuristic-based optimization remains an underexplored yet popular research area for progressing cyberattack detection in IoT. Thus, in this study, the IMFO-based FS and DTOA-based hyperparameter selection is involved to increase security in IoT networks.

III. THE PROPOSED MODEL

In this paper, focus is given to the growth of automatic IDS employing the IMFOHDL-ID model in the IoT environment. The developed IMFOHDL-ID method aims to classify intrusions and achieve security in an IoT environment. The IMFOHDL-ID model comprises a sequence of operations namely data normalization, IMFO-based FS, LSTM-DSSAE-based recognition, and DTOA-based hyperparameter tuning. Fig. 2 depicts the entire flow of the IMFOHDL-ID technique.

A. DATA NORMALIZATION

At the primary level, the proposed model follows data normalization as a preprocessing stage. In IoT platforms, especially for IDS, data normalization using the min-max scaling method is crucial for ensuring that inconsistent and diverse data from multiple devices and sensors are brought to a general scale. Min-max normalization facilitates the comparability of the data stream and improves the efficiency of the anomaly detection algorithm by rescaling the data to fall within a certain interval (0 to 1). This allows the IDS to more precisely recognize deviations and potential security threats in the IoT system, making it a vital step in ensuring the reliability and robustness of the security network.

B. FEATURE SELECTION USING IMFO ALGORITHM

At this phase, the IMFOHDL-ID technique makes use of the IMFO-based FS approach to elect feature subsets. The MFO model is reflected as a hybrid technique [21]. Its stimulation derives from the public act of mayflies (MF). It is measured that MFs produced teenagers and rightest ones endure.

At an initial stage, dual groups of populace are produced. They characterize the female and male populace. The aspirants are denoted by d-dimensional vector $= (x_1, \ldots, x_d)$. So, the fitness of applicants is valued by computing the fitness

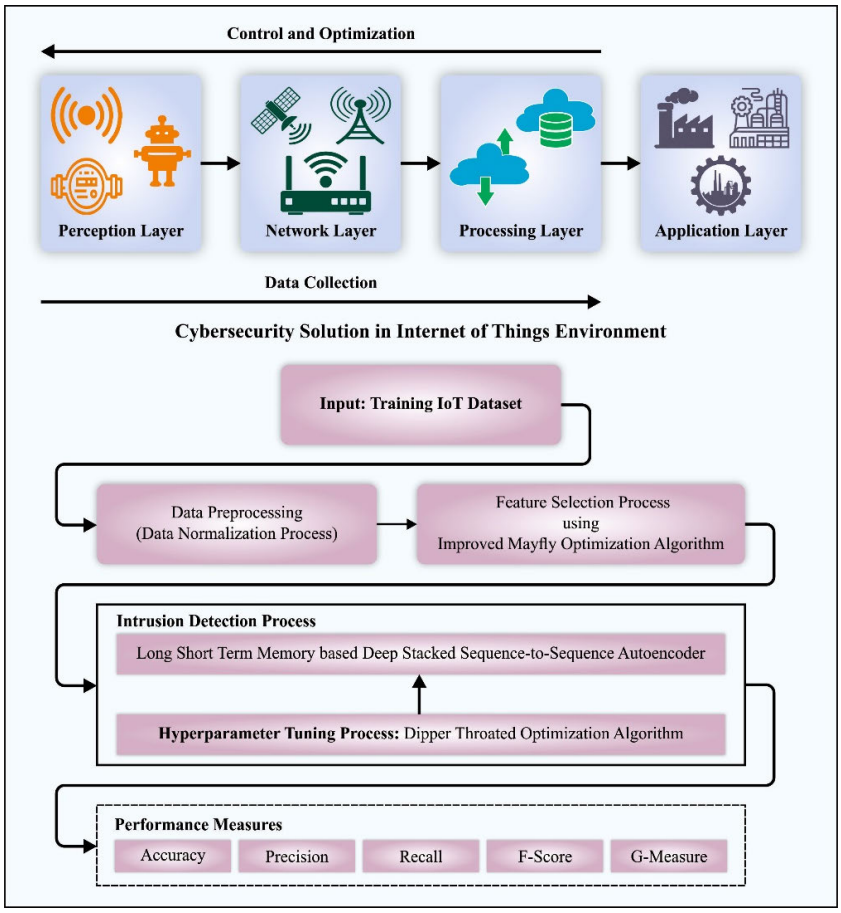

FIGURE 2. Overall flow of IMFOHDL-ID algorithm.

function (FF)(x). Velocity $v = (v_1, \ldots, v_d)$ is alteration in candidate place. Each applicant adjusts its trajectory due to the best location (*pbest*) and optimal place of all MFs (*gbest*). The gathering of male MFs imitates the knowledge of each male in deciding its position with esteem to neighbors' place. Degrading x_i^t as a present place of applicant solution i at t time, the place is adapted by totaling a velocity v_i^{t+1} as follows

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{1}$$

With $x_i^0 U\ (x_{\min}, x_{\max})$.

Assume low velocity of the male populace, the speed is considered as

$$v_{ij}^{t+1} = v_{ij}^t + a_1 e^{-\beta r_p^2}\left(pbest_{ij} - x_{ij}^t\right) + a_2 e^{-\beta r_g^2}\left(gbest_i - x_{ij}^t\right) \tag{2}$$

where v_{ij}^t velocity of MF is i, x_{ij}^t is a place of MF i, and the positive factors a_1 and a_2 represent the attraction. $pbest_i$ refers to an optimal place that an applicant solution i had ever extended, and $pbest_{ij}$ at succeeding step $t+1$ is defined in Eq. (3).

$$pbest_i = \begin{cases} x_i^{t+1}, & if\ f\left(x_i^{t+1}\right) < f\ (pbest_i) \\ same\ as\ before, & otherwise \end{cases} \tag{3}$$

where $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ is the main task to be decreased, *gbest* is the worldwide optimal attained in problematic ever at time t. The factors in Eq. (2) restrict populace's discernibility. r_p represents the space among x_i and $pbest_i$. While r_g determines the space from x_i to *gbest*. r_p and r_g are determined by Eq. (4).

$$\|x_i - X_i\| = \sqrt{\sum_{j=1}^{n}(x_{ij} - X_{ij})^2} \tag{4}$$

where x_{ij} is j^{th} part of i^{th} applicant. X_i is related to *pbest*. The optimal appropriate applicants retain executing up and down motion by changing speeds. The velocities are defined by Eq. (5).

$$v_{ij}^{t+1} = v_{ij}^t + d * r \tag{5}$$

Here d is a constant linked to upper and lower movement and r denotes a randomly generated value amongst [−1,1].

Female MFs will not collect but transport towards males. Consider y_i^t as a present place of female MF i at time t. Alteration in place is considered as follows

$$y_i^{t+1} = y_i^t + v_i^{t+1} \tag{6}$$

with $y_i^0 U(x_{\min}, x_{\max})$. The velocities of female MFs are defined by Eq. (7).

$$v_{ij}^{t+1} = \begin{cases} v_{ij}^t + a_2 e^{-\beta r_{mf}^2} \left(x_{ij}^t - y_{ij}^t\right), & if\ f\ (y_i) > f\ (x_i) \\ v_{ij}^t + fl * r, & if\ f\ (y_i) \leq f\ (x_i) \end{cases} \tag{7}$$

where v_{ij}^t is the speed of i^{th} female at time t, y_{ij}^t is the place of i^{th} female candidate solution at time t, a_2 is a positive constant, β denotes the stable coefficient, r_{mf} signifies space among male candidate and female ones that are calculated utilizing Eq. (4), fl is a number which agrees female is not concerned, and r refers the random value amid -1 and 1. The mating is signified by a crossover operator. A set of female and male parents are chosen. Afterwards, the crossover operator produces dual off-spring as expressed in Eq. (8).

$$\begin{aligned} offspring1 &= L * male + (1 - L) * female \\ offspring2 &= L * female + (1 - L) * male \end{aligned} \tag{8}$$

where L is a random number. At an early stage, offspring velocity is equivalent to zero. The search agent's initialization in optimization issues is executed casually. A random vector is fixed with values extending among predefined minimal as well as maximal boundaries. The growth of an early populace is prepared by hybridization with chaotic maps. This study displays that logistic chaotic maps are enhanced when compared with other existing chaotic maps. This is owing to enhanced computational efficacy and a high probability of pledge random values near 0 and 1. Therefore, the quicker local search produced. The logistic chaotic mapping is stated scientifically in Eq. (9) as follows:

$$y_1 = rand, y_{i+1} = 4 \times y_i \times (1 - y_i), i = 1, 2, \ldots, N \tag{9}$$

where *rand* denotes random vector ranges between [0,1]. The presented IMFO model was firmed by changing a random vector known as $rand_i$ that was calculated by logistic chaotic mapping based on Eq. (9). Therefore, the MFO technique was improved by modification in an early populace by chaotic character.

In the developed IMFO-FS model, the FF planned to have stability among classifier accuracy (maximum) and number of features selected in solution (minimum) achieved by these FS, Eq. (10) signifies the FF to evaluate a solution.

$$Fitness = \alpha \gamma_R (D) + \beta \frac{|R|}{|C|} \tag{10}$$

Here the classifier error value is $\gamma_R(D)$. The cardinality of the particular subset is $|R|$ and overall number of features in the data is$|C|$, and two parameters α and β correspond to the significance of classification algorithm quality and length of the subset. $\in [1, 0]$ and $\beta = 1 - \alpha$.

C. INTRUSION DETECTION USING HDL MODEL

For the intrusion detection procedure, the HDL technique using LSTM-DSSAE was applied. An LSTM-AE is an execution of AE for data series employing an encoder-decoder LSTM model [22]. It is intended to encode as well as decode consecutive information namely text information or time series. In this work, manifold layers of the LSTM unit are fixed and organized in order to create encoder-decoder architecture. LSTM is a kind of recurrent neural network (RNN) that excels at grabbing time-based dependency in a sequence of data. LSTM-based DSSAE marks it suitable for the task including data sequences like time series analysis. Moreover, LSTM has a memory unit that enables to preservation of data over long sequences, utilizing and capturing context data efficiently. Especially, this ability is advantageous while recreating data sequence, where the context and past information are critical. Classical AE does not possess this memory unit, which makes them less effective at encoder and decoder series.

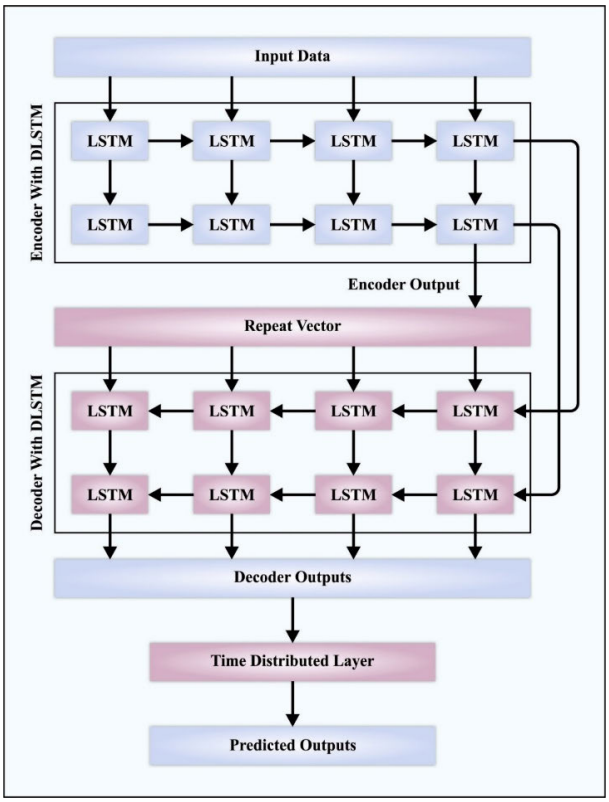

FIGURE 3. Architecture of LSTM-DSSAE.

Consider LSTM-based DSSAE with single-layer LSTM. Input series, represented as X, includes a set of input vectors $X = \left(x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right)$, in which n symbolizes series length. LSTM contains a recurrent layer, and h hidden state in the encoder and decoder phases. The encoding procedures

input series and provide compressed representation z, and the decoder recreates unique series from these representations.

During the encoder, all the vectors of the time-window of length are given to the recurrent unit for performing subsequent calculations:

$$h^{(t)} = LSTM_{Encoder}\left(x^{(t)}, h^{(t-1)}; \omega_e\right) \tag{11}$$

where the hyperparameter of the encoder model is ω_e. The LSTM-based DSSAE is used to generate output series $Y = (y^{(1)}, y^{(2)}, y^{(n)})$ for input sequence X, Generally, $X = Y$ forces AE to learn the semantic meaning of information. Firstly, the encoder LSTM can encode the input series, and later the parameter C is translated by the LSTM decoder.

$$y^{(t)} = LSTM_{Decoder}\left(h^{(t)}, y^{(t-1)}; \omega_d\right) \tag{12}$$

Fig. 3 illustrates the infrastructure of LSTM-DSSAE. The encoder part takes an input series data, processes it through the LSTM layer for extracting useful features, and constructs a compressed representation named hidden space otherwise called hidden space or bottleneck:

LSTM-based DSSAE aims for minimizing the variance between its reconstruction and the input sequence. In such cases, the max MSE can be extensively applied:

$$Loss = \sum \left(x^{(t)} - y^{(t)}\right)^2 \tag{13}$$

Here the input series is $x^{(t)}$, the reconstructed output is $y^{(t)}$, and the sequence length is n. By minimalizing reconstructed loss, LSTM-based DSSAE learns to grab relevant features of input series within the compacted symbol and recreate unique sequences from that symbol.

D. HYPERPARAMETER TUNING USING DTOA

Finally, DTOA optimally elects the hyperparameter values of the HDL approach. DTOA is a novel metaheuristic method inspired by the cooperative nature of the Birds [23]. The DTOA model uses three strategies to enhance exploration: (1) fly towards a novel location, (2) switch to another bird, and (3) proficiently fly over a called region. The utilization technique includes viewing birds and annoying to hunt each other for food.

A flock of bird's swims via space to search food while applying the DTOA technique. The speeds and positions of birds are described as P and V, correspondingly. In this case, DTOA explores the search range for a better solution. The computation of DTOA is given below.

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & P_{1,3} & \ldots & P_{1,d} \\ P_{2,1} & P_{2,2} & P_{2,3} & \ldots & P_{2,d} \\ P_{3,1} & P_{3,2} & P_{3,3} & \ldots & P_{3,d} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ P_{m,1} & P_{m,2} & P_{m,3} & \ldots & P_{m,d} \end{bmatrix} \tag{14}$$

$$V = \begin{bmatrix} V_{1,1} & V_{1,2} & V_{1,3} & \cdots & V_{1,d} \\ V_{2,1} & V_{2,2} & V_{2,3} & \cdots & V_{2,d} \\ V_{3,1} & V_{3,2} & V_{3,3} & \cdots & V_{3,d} \\ V_{m,1} & V_{m,2} & V_{m,3} & \cdots & V_{m,d} \end{bmatrix} \tag{15}$$

The i^{th} bird is represented as $P_{i,j}$, and the bird's speed is represented by $V_{i,j}$ For indexes $\in 1, 2, 3, \ldots, m$ and $j \in 1, 2, 3, \ldots, d$ and j^{th} dimension. Eq. (16) defines the fitness of the bird $= f_1, f_2, f_3, \ldots, f_n$.

$$f = \begin{bmatrix} f_1\left(P_{1,1}, P_{1,2}, P_{1,3}, \cdots, P_{1,d}\right) \\ f_2\left(P_{2,1}, P_{2,2}, P_{2,3}, \cdots, P_{2,d}\right) \\ f_2\left(P_{3,1}, P_{3,2}, P_{3,3}, \cdots, P_{3,d}\right) \\ \cdots \\ f_{\mathrm{m}}\left(P_{\mathrm{m},1}, P_{\mathrm{m},2}, P_{\mathrm{m},3}, \cdots, P_{\mathrm{m},d}\right) \end{bmatrix} \tag{16}$$

Mother birds have maximum fitness among birds while they can produce maximum offspring with the ability to survive and search for food. During the search process, the best location P_{best}, is updated. P_{nd} represents regular birds, which serve as followers of mother birds. During the search process, P_{best} represents the obtained optimum solution. Using the following equations, the optimizer exploits the DTOA algorithm to follow the swimming bird to account for displacement within time and population.

$$X = P_{best}(i) - K_1.|K_2.P_{best}(i) - P(i)| \tag{17}$$

$$Y = P(i) + V(i+1) \tag{18}$$

$$P(i+1) = \begin{cases} X & if\ r_3 < 05 \\ Y & otherwise \end{cases} \tag{19}$$

$$V(i+1) = K_3 V(i) + K_4 r1\left(P_{best}(i) - P(i)\right) + K_5 r_2\left(P_{Gbest} - P(i)\right) \tag{20}$$

Now, $P_{best}(i)$ is the best location of the bird, $P(i)$ is the average location of the bird for i^{th} iterations, and the speed of birds at $i+1$ iteration is $V(i+1)$ .K_1, K_2, and K_3 are weight values dynamically selected between [0,2] and K_4 and K_5 are constants with values 1.7 and 1.8, correspondingly. The random number within [0,1] makes up the values of , r, and $r3$.

The fitness choice is a considerable factor influencing the solution of the DTOA. The hyperparameter choice procedure contains the efficiency encoded method to measure the performance of the applicant solutions. In this case, the DTOA algorithm considers exactness as the main criterion to design the FF which is defined as:

$$Fitness = \max(P) \tag{21}$$

$$P = \frac{TP}{TP + FP} \tag{22}$$

From the above equation, TP signifies true positive and FP represents false positive value.

IV. PERFORMANCE VALIDATION

The proposed model is simulated using the Python 3.8.5 tool. The proposed model is experimented on PC i5-8600k, GeForce 1050Ti 4GB, 16GB RAM, 250GB SSD, and 1TB HDD.

A. DATASET USED

In this study, intrusion detection outcomes of the IMFOHDL-ID system can be examined with the BoT-IoT database [24], as defined in Table 1.

TABLE 1. Dataset specification.

Labels	Classes	Sample Numbers
C1	Service Scanning	1000
C2	OS Fingerprinting	1000
C3	DDoS TCP	1000
C4	DDoS UDP	1000
C5	DDoS HTTP	989
C6	DoS TCP	1000
C7	DoS UDP	1000
C8	DoS HTTP	1000
C9	Normal	477
Overall Samples		**8466**

B. PERFORMANCE MEASURES

An accuracy ($accu_y$), precision ($prec_n$), recall ($reca_l$), F-score (F_{score}), G-measure ($G_{measure}$), and Mathew Correlation Coefficient (MCC) are set of measures used to examine the classification outcomes.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{23}$$

Precision is used to measure the quantity of correctly predicted positive samples against the total samples that have been predicted as positive.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{24}$$

Recall is used to measure the amount of positive instances correctly classified.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{25}$$

Accuracy is used to measure the amount of correctly classified instances (positives and negatives) out of all the instances (amount of instances that were classified).

$$\text{F} - \text{score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{26}$$

F − score is a measure uniting the harmonic mean of precision and recall.

$$MCC = \frac{TN \times TP - FN \times FP}{\sqrt{(TP + FP)\,(TP + FN)\,(TN + FP)\,(TN + FN)}} \tag{27}$$

MCC is a statistical tool utilized for model calculation. *MCC* is the best single-value classification metric that enables the an error matrix or confusion matrix.

$$G - measure = \sqrt{Precision \cdot Recall} \tag{28}$$

The $G - measure$, a.k.a. the geometric mean of precision and recall, is the alternative performance metric for binary classification.

C. RESULTS AND DISCUSSION

Fig. 4 illustrates the analysis of the IMFOHDL-ID model with the test database. Figs. 4a-4b shows the confusion matrix given by the IMFOHDL-ID system at 80:20 of TRAPS/TESPS. This outcome value pointed out that the IMFOHDL-ID model is properly recognized and classified with nine classes. Additionally, Fig. 4c exhibits the PR analysis of the IMFOHDL-ID system. The result pointed out that the IMFOHDL-ID system attains excellent PR performance with each class. Also, Fig. 4d exhibits the ROC curve of the IMFOHDL-ID technique. The outcome described that the IMFOHDL-ID model leads to proficient outcomes with higher ROC values with various classes.

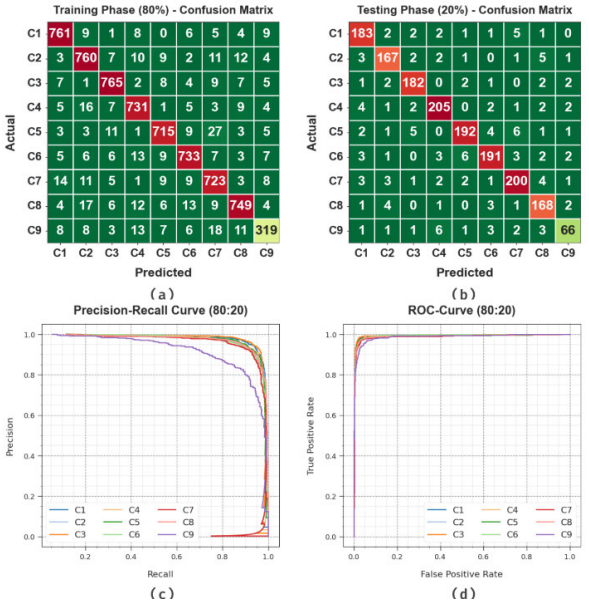

FIGURE 4. Confusion matrices of (a-b) TRAPS of 80% and TESPS of 20% and (c-d) PR curve (80:20) and ROC curve (80:20).

In Table 2 and Fig. 5, the intrusion detection analysis of the IMFOHDL-ID technique can be validated with 80:20 of TRAPS/TESPS. The obtained values indicate that the IMFOHDL-ID technique reaches maximum performance with each class. Additionally, on 80% of TRAPS, the IMFOHDL-ID systems obtain average $accu_y$, $prec_n$, $reca_l$, F_{score}, MCC, and $G_{measure}$values of 98.31%, 92.09%, 91.75%, 91.90%, 90.96%, and 91.91%, respectively. Meanwhile, with 20% of TESPS, the IMFOHDL-ID technique obtains average $accu_y$, $prec_n$, $reca_l$, F_{score}, MCC, and $G_{measure}$ values of 98.16%, 91.30%, 90.91%, 91.08%, 90.06%, and 91.09%, correspondingly.

To evaluate the efficiency of the IMFOHDL-ID algorithm with 80:20 of TRAPS/TESPS, it can be generated $accu_y$ curves for both the TRAPS and TESPS, as exemplified in Fig. 6. These curves offer appreciated insights into the

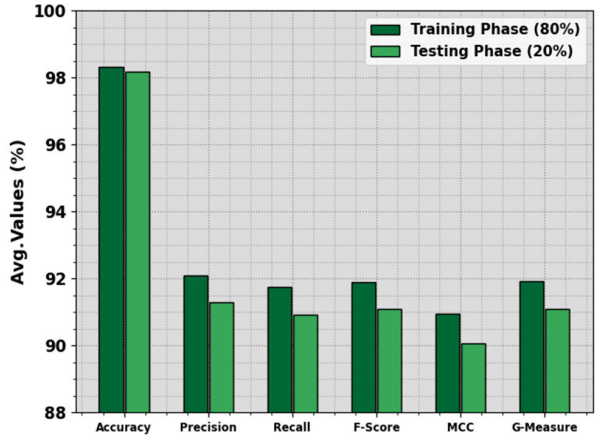

FIGURE 5. Average of IMFOHDL-ID system with 80:20 of TRAPS/TESPS.

TABLE 2. Intrusion detection outcome of IMFOHDL-ID algorithm with 80:20 of TRAPS/TESPS.

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC	$G_{measure}$
TRAPS (80%)						
C1	98.66	93.95	94.77	94.36	93.60	94.36
C2	98.10	91.46	92.91	92.18	91.10	92.18
C3	98.69	94.33	94.68	94.50	93.76	94.50
C4	98.38	92.41	93.60	93.00	92.09	93.00
C5	98.36	93.59	92.02	92.80	91.88	92.80
C6	98.38	93.14	92.90	93.02	92.10	93.02
C7	97.80	89.04	92.34	90.66	89.43	90.67
C8	98.18	93.51	91.34	92.41	91.39	92.42
C9	98.23	87.40	81.17	84.17	83.30	84.23
Average	**98.31**	**92.09**	**91.75**	**91.90**	**90.96**	**91.91**
TESPS (20%)						
C1	98.11	91.04	92.89	91.96	90.90	91.96
C2	98.23	91.76	91.76	91.76	90.77	91.76
C3	98.64	93.33	94.79	94.06	93.29	94.06
C4	98.23	92.76	93.61	93.18	92.17	93.18
C5	98.05	93.66	90.57	92.09	90.99	92.10
C6	97.93	92.72	90.52	91.61	90.44	91.61
C7	97.87	91.32	92.17	91.74	90.52	91.74
C8	98.11	89.36	93.33	91.30	90.27	91.33
C9	98.29	85.71	78.57	81.99	81.18	82.07
Average	**98.16**	**91.30**	**90.91**	**91.08**	**90.06**	**91.09**

method's capability and learning progress to generalize. But if epoch count is enhanced, an observable enrichment in both TRA and TES $accu_y$ curves can be obvious. This improvement represents the method's capacity to distinguish patterns from both TRA and TES data.

Fig. 7 also provides an overview of the IMFOHDL-ID methodology at 80:20 of TRAPS/TESPS loss values at the TRA model. The lessening trend in TRA loss under epochs shows that the model continuously increases its weights to reduce predicted errors on both data. This loss curve

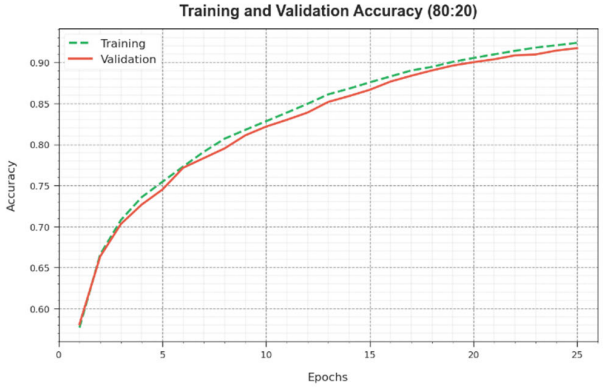

FIGURE 6. $Accu_y$ curve of IMFOHDL-ID model with 80:20 of TRAPS/TESPS.

considers how well the model fits the TRA data. Remarkably, the TRA and TES loss are consistently minimized, representing the model's effectual pattern learning existence in both databases. Then, it exhibits the model's adaptation for decreasing differences among new and predicted TRA labels.

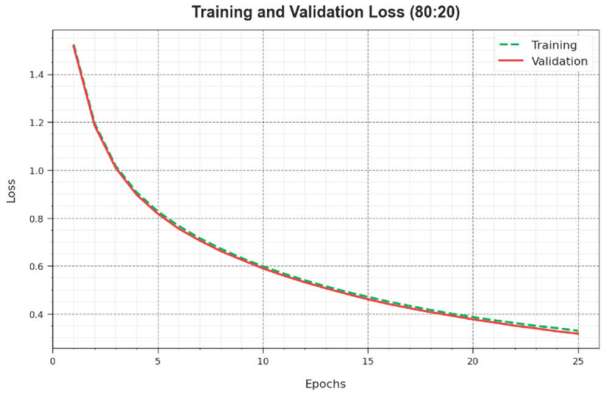

FIGURE 7. Loss curve of IMFOHDL-ID system with 80:20 of TRAPS/TESPS.

Fig. 8 shows the classifier outcome of the IMFOHDL-ID algorithm with the test dataset. Figs. 8a-8b depicts the confusion matrix provided by the IMFOHDL-ID model at 70:30 of TRAPS/TESPS. The outcome pointed out that the IMFOHDL-ID model is appropriately recognized and classified into 9 classes. Moreover, Fig. 8c represents the PR performance of the IMFOHDL-ID methodology. The simulation result shows that the IMFOHDL-ID system achieves exceptional PR performance with each class. Besides, Fig. 8d exhibits the ROC analysis of the IMFOHDL-ID algorithm. The value described that the IMFOHDL-ID system leads to proficient values with maximal ROC values with different classes.

In Table 3 and Fig. 9, the intrusion detection output of the IMFOHDL-ID method can be confirmed with 70:30 of TRAPS/TESPS. The obtained values show that the IMFOHDL-ID method attains better performance with nine classes. Moreover, with 70% of TRAPS, the IMFOHDL-ID approach attain average $accu_y$, $prec_n$, $reca_l$, F_{score}, MCC,

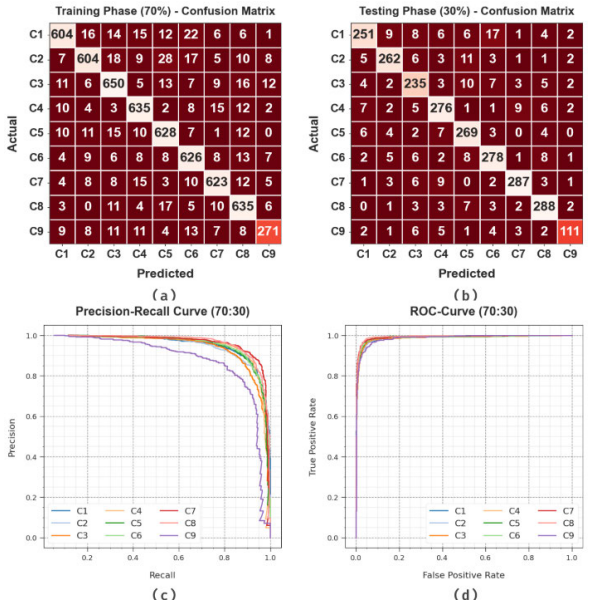

FIGURE 8. Confusion matrices of (a-b) TRAPS of 70% and TESPS of 30% and (c-d) PR curve (70:30) and ROC curve (70:30).

and $G_{measure}$ values of 97.56%, 88.94%, 88.49%, 88.67%, 87.32%, and 88.69%, correspondingly. Moreover, with 30% of TESPS, the IMFOHDL-ID approach acquires average $accu_y$, $prec_n$, $reca_l$, F_{score}, MCC, and $G_{measure}$ values of 97.52%, 88.94%, 88.41%, 88.62%, 87.26%, and 88.65%, respectively.

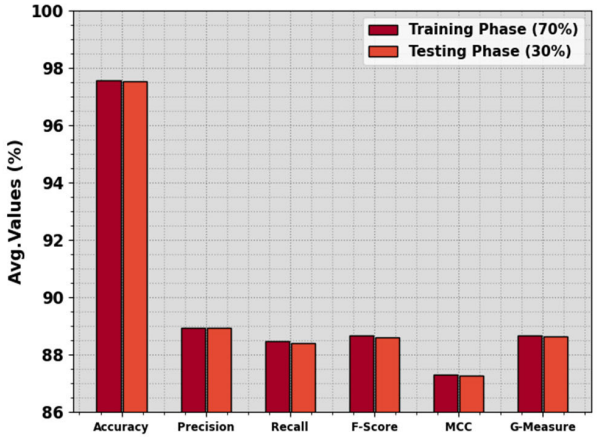

FIGURE 9. Average of IMFOHDL-ID system with 70:30 of TRAPS/TESPS.

To define the efficacy of the IMFOHDL-ID model with 70:30 of TRAPS/TESPS, $accu_y$ curves are made for both the TRAPS and TESPS, as represented in Fig. 10. These curves offer valuable insights into the methodology's ability and learning progress to generalize. As the count of epochs is raised, a noticeable enrichment in both TRA and TES $accu_y$ curves is cleared. This enhancement represents the capability of the technique to detect patterns from both the TRA and TES data.

TABLE 3. Intrusion detection analysis of IMFOHDL-ID algorithm with 70:30 of TRAPS/TESPS.

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC	$G_{measure}$
TRAPS (70%)						
C1	97.47	91.24	86.78	88.95	87.56	88.98
C2	97.23	90.69	85.55	88.05	86.53	88.08
C3	97.22	88.32	89.16	88.74	87.15	88.74
C4	97.76	89.19	91.90	90.52	89.26	90.53
C5	97.42	87.83	90.49	89.14	87.69	89.15
C6	97.44	87.55	90.86	89.17	87.74	89.19
C7	97.87	91.08	90.55	90.82	89.61	90.82
C8	97.55	87.71	91.90	89.75	88.40	89.78
C9	98.11	86.86	79.24	82.87	81.97	82.96
Average	**97.56**	**88.94**	**88.49**	**88.67**	**87.32**	**88.69**
TESPS (30%)						
C1	96.85	90.29	82.57	86.25	84.59	86.34
C2	97.68	90.66	89.12	89.88	88.57	89.88
C3	96.93	84.84	86.72	85.77	84.05	85.77
C4	97.20	87.90	89.32	88.60	87.01	88.61
C5	97.24	85.94	91.19	88.49	86.97	88.53
C6	97.13	87.42	89.39	88.39	86.76	88.40
C7	98.23	93.49	91.99	92.73	91.73	92.73
C8	97.87	89.72	93.20	91.43	90.24	91.45
C9	98.58	90.24	82.22	86.05	85.40	86.14
Average	**97.52**	**88.94**	**88.41**	**88.62**	**87.26**	**88.65**

FIGURE 10. $Accu_y$ curve of IMFOHDL-ID system with 70:30 of TRAPS/TESPS.

Fig. 11 also provides an overview of the IMFOHDL-ID methodology at 70:30 of TRAPS/TESPS loss values at the TRA procedure. The reducing trend in TRA loss over epochs shows that the model continuously raises its weights to decline predictive errors on both TRA and TES datasets. This loss curve considers how well the model fits the TRA data. Considerably, the TRA and TES losses are reliably minimized, representative of the methodology's effectual learning of patterns existing in both databases. Further, it shows the

model's adaptation for minimizing differences among new and predictive TRA labels.

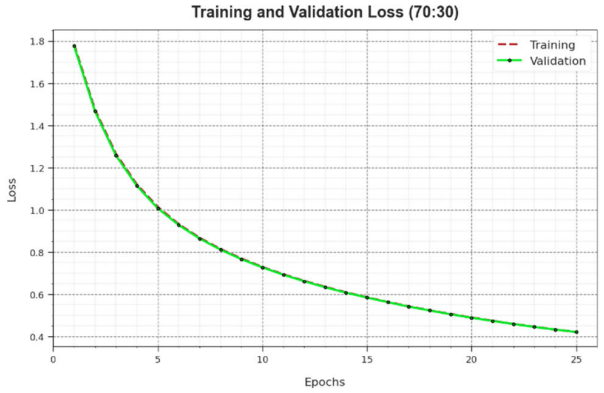

FIGURE 11. Loss curve of IMFOHDL-ID method with 70:30 of TRAPS/TESPS.

Eventually, the comparison investigation of the IMFOHDL-ID system with recent methodologies on the intrusion detection process is stated in Table 4 and Fig. 12 [16], [19], [20]. The achieved outcomes exhibited that the DT and GWO systems obtain poor performance whereas the MGO and LD models have reported certainly increased results. Along with that, the ensemble bag, KNN, and SVM models have accomplished considerable solutions. However, the IMFOHDL-ID approach gains maximal performance over other algorithms with maximum $accu_y$, $prec_n$, $reca_l$, F_{score}, and MCC values of 98.31%, 92.09%, 91.75%, 91.90%, and 90.96%, respectively.

TABLE 4. Comparison analysis of the IMFOHDL-ID algorithm with other models [16], [19], [20].

Classifiers	$Accu_y$	$Prec_n$	$Reca_l$	F_{score}	MCC
DT Model	92.48	85.79	87.68	90.87	87.23
Ensemble Bag	97.81	85.88	90.57	90.84	85.12
KNN Model	97.10	89.64	90.44	88.32	84.24
LD Model	94.77	91.03	85.28	87.12	88.06
SVM Model	97.98	86.51	89.73	87.56	79.95
MGO Model	95.69	90.50	86.94	89.91	80.64
GWO Model	93.51	86.87	88.89	90.48	87.71
IMFOHDL-ID	98.31	92.09	91.75	91.90	90.96

Finally, the comparison computational time (CT) examination of the IMFOHDL-ID method with existing methods on the intrusion detection techniques is illustrated in Table 5 and Fig. 13.

The outcomes achieved portrayed that the DT and GWO methods attain poor performance whereas the MGO and LD techniques have stated certainly inferior results. With that, the ensemble bag, KNN, and SVM approaches have obtained remarkable outcomes. However, the IMFOHDL-ID technique obtains higher performance over other approaches

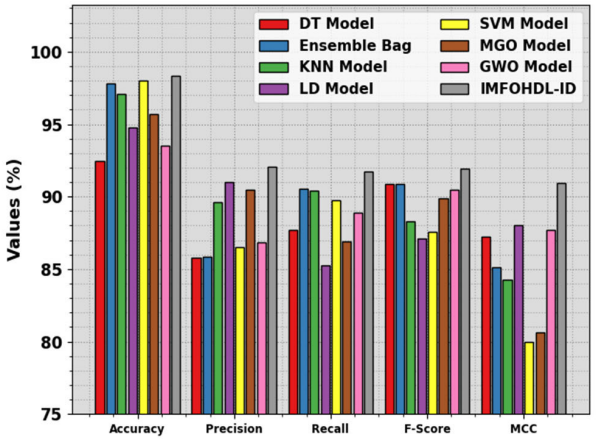

FIGURE 12. Comparison analysis of the IMFOHDL-ID system with other models.

TABLE 5. CT analysis of the IMFOHDL-ID algorithm with other models.

Classifiers	Computational Time (sec)
DT Model	4.07
Ensemble Bag	5.48
KNN Model	3.07
LD Model	4.79
SVM Model	3.91
MGO Model	3.85
GWO Model	3.67
IMFOHDL-ID	1.93

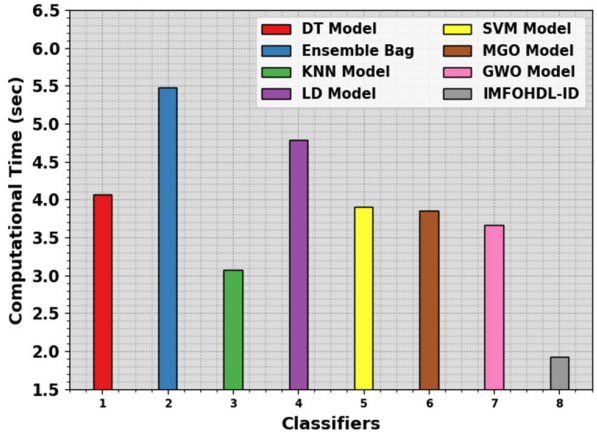

FIGURE 13. CT analysis of the IMFOHDL-ID system with other models.

with a minimum CT of 1.93s. These obtained outcomes confirmed that the IMFOHDL-ID technique accomplishes improved detection results.

V. CONCLUSION

In this study, the developed IMFOHDL-ID method aims to classify intrusions and accomplish safety in the IoT environment. IMFOHDL-ID model encompasses a series of

operations like data normalization, IMFO-based FS, LSTM DSSAE-based detection, and DTOA-based hyperparameter tuning. Primarily, the proposed model follows data normalization as a pre-processing stage. In addition, the IMFOHDL-ID model makes use of the IMFO-based FS approach to elect feature subsets. For intrusion detection, the IMFOHDL-ID technique utilized the LSTM-DSSAE approach. At last, the DTOA was utilized for the optimum hyperparameter choice of the LSTM-DSSAE model. To highlight the optimum solution of the IMFOHDL-ID method, a series of simulation studies were executed. Extensive comparative results stated an enhanced performance of the IMFOHDL-ID model over existing models. Upcoming work can concentrate on handling outlier removal procedures to improve the performance of the IMFOHDL-ID technique.

Future work can focus on the evaluation of the proposed model by testing its performance across different IoT networks, including different types and scales. This involves deploying the model in various IoT setups, ranging from small-scale home automation systems to large-scale industrial IoT environments. Thereby, we aimed to evaluate the model's capability to generalize and maintain higher accuracy in identifying cyberattacks under varying conditions and data distributions. Furthermore, this broader testing will help us detect any environment-specific problems and refine the model to ensure its robustness and practical utility in real-time applications.

REFERENCES

[1] M. A. Ferrag, O. Friha, L. Maglaras, H. Janicke, and L. Shu, "Federated deep learning for cyber security in the Internet of Things: Concepts, applications, and experimental analysis," *IEEE Access*, vol. 9, pp. 138509–138542, 2021.

[2] K. Bresniker, A. Gavrilovska, J. Holt, D. Milojicic, and T. Tran, "Grand challenge: Applying artificial intelligence and machine learning to cybersecurity," *Computer*, vol. 52, no. 12, pp. 45–52, Dec. 2019.

[3] M. L, M. E, and M. A, "Cybersecurity management for (industrial) Internet of Things: Challenges and opportunities," *J. Inf. Technol. Softw. Eng.*, vol. 8, no. 5, pp. 1–9, 2018.

[4] A. Corallo, M. Lazoi, M. Lezzi, and A. Luperto, "Cybersecurity awareness in the context of the industrial Internet of Things: A systematic literature review," *Comput. Ind.*, vol. 137, May 2022, Art. no. 103614.

[5] R. J. Raimundo and A. T. Rosário, "Cybersecurity in the Internet of Things in industrial management," *Appl. Sci.*, vol. 12, no. 3, p. 1598, Feb. 2022.

[6] M. Abdullahi, Y. Baashar, H. Alhussian, A. Alwadain, N. Aziz, L. F. Capretz, and S. J. Abdulkadir, "Detecting cybersecurity attacks in Internet of Things using artificial intelligence methods: A systematic literature review," *Electronics*, vol. 11, no. 2, p. 198, Jan. 2022.

[7] R. Rudenko, I. M. Pires, P. Oliveira, J. Barroso, and A. Reis, "A brief review on Internet of Things, Industry 4.0 and cybersecurity," *Electronics*, vol. 11, no. 11, p. 1742, May 2022.

[8] B. Ghimire and D. B. Rawat, "Recent advances on federated learning for cybersecurity and cybersecurity for federated learning for Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 11, pp. 8229–8249, 2022.

[9] W. Ahmad, A. Rasool, A. R. Javed, T. Baker, and Z. Jalil, "Cyber security in IoT-based cloud computing: A comprehensive survey," *Electronics*, vol. 11, no. 1, p. 16, Dec. 2021.

[10] G. Gardašević, L. Berbakov, and A. Mastilović, "Cybersecurity of industrial Internet of Things," in *Cyber Security of Industrial Control Systems in the Future Internet Environment*. Hershey, PA, USA: IGI Global, 2020, pp. 47–68.

[11] A. Awajan, "A novel deep learning-based intrusion detection system for IoT networks," *Computers*, vol. 12, no. 2, p. 34, Feb. 2023.

[12] 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.

[13] R. Chaganti, W. Suliman, V. Ravi, and A. Dua, "Deep learning approach for SDN-enabled intrusion detection system in IoT networks," *Information*, vol. 14, no. 1, p. 41, Jan. 2023.

[14] M. Ramaiah, V. Chandrasekaran, V. Ravi, and N. Kumar, "An intrusion detection system using optimized deep neural network architecture," *Trans. Emerg. Telecommun. Technol.*, vol. 32, no. 4, Apr. 2021, Art. no. e4221.

[15] V. Ravi, R. Chaganti, and M. Alazab, "Recurrent deep learning-based feature fusion ensemble meta-classifier approach for intelligent network intrusion detection system," *Comput. Electr. Eng.*, vol. 102, Sep. 2022, Art. no. 108156.

[16] A. Fatani, A. Dahou, M. Abd Elaziz, M. A. A. Al-Qaness, S. Lu, S. A. Alfadhli, and S. S. Alresheedi, "Enhancing intrusion detection systems for IoT and cloud environments using a growth optimizer algorithm and conventional neural networks," *Sensors*, vol. 23, no. 9, p. 4430, Apr. 2023.

[17] Z. Wang, H. Chen, S. Yang, X. Luo, D. Li, and J. Wang, "A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization," *PeerJ Comput. Sci.*, vol. 9, Sep. 2023, Art. no. e1569.

[18] M. Alazab, A. Awajan, H. Alazzam, M. Wedyan, B. Alshawi, and R. Alturki, "A novel IDS with a dynamic access control algorithm to detect and defend intrusion at IoT nodes," *Sensors*, vol. 24, no. 7, p. 2188, 2024.

[19] X. Liu and Y. Du, "Towards effective feature selection for IoT botnet attack detection using a genetic algorithm," *Electronics*, vol. 12, no. 5, p. 1260, Mar. 2023.

[20] S. Alosaimi and S. M. Almutairi, "An intrusion detection system using BoT-IoT," *Appl. Sci.*, vol. 13, no. 9, p. 5427, Apr. 2023.

[21] M. A. M. Shaheen, H. M. Hasanien, M. S. El Moursi, and A. A. El-Fergany, "Precise modeling of PEM fuel cell using improved chaotic MayFly optimization algorithm," *Int. J. Energy Res.*, vol. 45, no. 13, pp. 18754–18769, Oct. 2021.

[22] W. Xu, J. He, W. Li, Y. He, H. Wan, W. Qin, and Z. Chen, "Long-short-term-memory-based deep stacked sequence-to-sequence autoencoder for health prediction of industrial workers in closed environments based on wearable devices," *Sensors*, vol. 23, no. 18, p. 7874, Sep. 2023.

[23] A. A. Abdelhamid, E.-S. M. El-Kenawy, A. Ibrahim, M. M. Eid, D. S. Khafaga, A. A. Alhussan, S. Mirjalili, N. Khodadadi, W. H. Lim, and M. Y. Shams, "Innovative feature selection method based on hybrid sine cosine and dipper throated optimization algorithms," *IEEE Access*, vol. 11, pp. 79750–79776, 2023.

[24] *The Bot-IoT Dataset*. Accessed: Jun. 15, 2024. [Online]. Available: https://research.unsw.edu.au/projects/bot-iot-dataset

SALAHALDEEN DURAIBI received the B.S. degree in computer science from Jazan University, Saudi Arabia, the M.S. degree in computer science from Kentucky State University, USA, and the Ph.D. degree in computer science from the University of Idaho, USA. Currently, he is an Assistant Professor with the College of Engineering and Computer Science, Jazan University. His research interests include computer and network security, intrusion detection, AI, and ML.

ABDULLAH MUJAWIB ALASHJAEE (Member, IEEE) received the Ph.D. degree in computer science from the University of Idaho, USA, in 2021. He is currently an Assistant Professor with the Department of Computer Science, Northern Border University, Saudi Arabia. He has published many articles in top academic journals and conferences. His current research interests include specifically on the fields of mobile malware forensics, cybersecurity, digital forensics, and intrusion detection systems.