Intrusion Detection in Internet of Things Based Smart Farming Using Hybrid Deep Learning Framework

Keerthi Kethineni
Koneru Lakshmaiah Education Foundation
G Pradeepini (✉ Pradeepini_cse@kluniversity.in)
Koneru Lakshmaiah Education Foundation

Research Article

Keywords: Intrusion detection, Label encoding, Smart agriculture, Detection, Attention mechanism, Normalization, Deep learning, Weight optimization

Posted Date: January 25th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2498495/v1

Additional Declarations: No competing interests reported.

Intrusion Detection in Internet of Things Based Smart Farming Using Hybrid Deep Learning Framework

[1,2] Keerthi Kethineni, [*3]G Pradeepini

[1]Research Scholar, Koneru Lakshmaiah Education Foundation, Guntur, India.
[2]Assistant Professor, Department of Computer Science and Engineering,
V R Siddhartha Engineering College, Vijayawada, India.
Email: kkeerthi@vrsiddhartha.ac.in
[*3]Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Guntur, India.
[*]Email: Pradeepini_cse@kluniversity.in

Abstract: Smart agriculture is a popular domain due to its intensified growth in recent times. This domain aggregates the advantages of several computing technologies, where the IoT is the most popular and beneficial. In this work, a novel and effective deep learning based framework is developed to detect intrusions in smart farming systems. The architecture is three-tier, with the first tier being the sensor layer, which involves the placement of sensors in agricultural areas. The second tier is the Fog Computing Layer (FCL), which consists of Fog nodes, and the proposed IDS is implemented in each Fog node. The gathered information is transferred to this fog layer for further analysis of data. The third tier is the cloud computing layer, which provides data storage and end-to-end services. The proposed model includes a fused CNN model with the bidirectional gated recurrent unit (Bi-GRU) model to detect and classify intruders. An attention mechanism is included within the BiGRU model to find the key features responsible for identifying the DDoS attack. In addition, the accuracy of the classification model is improved by using a nature-inspired meta-heuristic optimization algorithm called the Wild Horse Optimization (WHO) algorithm. The last layer is the cloud layer, which collects data from fog nodes and offers storage services. The proposed system will be implemented in the Python platform, using ToN-IoT and APA-DDoS attack datasets for assessment. The proposed system outperforms the existing methods in accuracy (99.35%), detection rate (98.99%), precision (99.9%) and F-Score (99.08%) for the APA DDoS attack dataset and the achieved accuracy of the ToN-IoT dataset (99.71%), detection rate (99.02%), precision (99.89%) and F-score (99.05%).

Keywords: Intrusion detection, Label encoding, Smart agriculture, Detection, Attention mechanism, Normalization, Deep learning, Weight optimization.

1. Introduction

The agricultural region plays a dynamic role in the country's financial growth and is one of the most crucial food providers. In order to meet the demands, the Food and Agriculture Organization (FAO) of the United Nations states that worldwide food productivity must reach 70% by 2050 [1]. Even though the current productivity rate is capable of meeting the demands, a report claimed that about 500 million people across the globe suffer from malnutrition while around 821 million people go hungry [2]. It is also estimated that the global population has increased to over 2 billion, and most of the impact of population growth. It might be seen in countries like India, Nigeria, Ethiopia, Pakistan, Egypt, the Democratic Republic of Congo, the United States, Indonesia and the United Republic of Tanzania [3]. It might also be problematic to meet the water demands of 40% by 2030, and arable land destruction impacts the overall food supply. Therefore, many sustainable systems and resources are required to obtain and maintain higher productivity levels to meet increasing demands worldwide [4, 5].

The increase in productivity requires advancements in cultivation practices and the adaptability of several technologies to provide crucial knowledge about the farm fields to take appropriate actions. Integrating novel and innovative technologies in the fields to gain optimized irrigation process is termed smart farming or precision farming [6, 7]. The technological integration provides more information regarding the fields and plants to enhance the irrigation procedure and to obtain optimal outcomes [8]. This information includes the presence of pests in the crops, the water requirement for the plants, the area needed to achieve higher production, resources required to control pests, the amount of fertilizers needed, etc. All these can be achieved by adapting prediction technologies, measurements of the environment and tools of automation [9, 10]. This combination can escalate the overall agricultural production to several extents without requiring huge amounts of natural resources. Smart farming integrates several technologies, protocols, computing paradigms and devices to empower the farmers to gather and understand most details regarding the farm fields [11, 12].

Since integrating technologies in smart farming offers undeniable benefits to farmers, this integration comes with many difficulties and complexities [13]. Among those challenges, the most terrifying one is intrusions intentionally introduced into embedded technologies to gain access to the accumulated data. Most security challenges are due to the vulnerability of the systems integrated into smart farming and due to their contained power [14, 15]. Technologies like artificial intelligence (AI), the Internet of Things (IoT), deep learning (DL), etc., were used for the effective removal of challenges like resource wastage and shortage of food [16]. Labor cost reduction, water and electricity conservation and the farmers can keep a record of their yield were the main advantages of IoT based smart agriculture [17]. The devices included in the fields collect crucial data regarding the farm fields and transmit them to the servers for storage. During the transfer of the data to the destination, there is a higher possibility of various security attacks that require quick fixes [18]. The most common attack in smart farming is the distributed denial of service (DDoS) attack, which can generate fake traffic on the network. This attack intensifies by compromising multiple devices in the network to generate fake traffic to overwhelm the network [19, 20]. Therefore, this paper focuses on developing an effective intrusion detection system (IDS) that can accurately detect and classify the DDoS attack on the network.

Smart agriculture is a popular domain due to its intensified growth in recent times. This domain aggregates the advantages of several computing technologies, where the IoT is the most popular and beneficial. The IoT system places sensors on agricultural fields to collect important data regarding crops and fields to improve the overall productivity rate. While transmitting the sensed data from the fields to the destination, there is a possibility of the occurrence of cyber-attacks that intruders design to gain access to the contents being transmitted. If the equipment installed in the field increases the production loss, it will be a serious issue. The main objectives of the proposal are as follows:

- Designing a new and effective hybrid deep learning based IDS framework for smart farming applications.
- A 3-tier architecture is considered in the work, where the sensing layer contributes to the agricultural fields on which the sensors are located, fog computing layer (FCL), where the IDS framework is deployed and the cloud computing layer for storage.
- A hybridization of deep learning named Fused and Optimized CNN-BiGRU (FOCB) with the metaheuristic optimization algorithm called Wild Horse Optimization (WHO) is explored to gain significant improvements in classification accuracy.
- The significance of the features representing DDoS attacks is further enhanced by adding an attention mechanism at the end of the BiGRU model.

The following sections are structured as follows: Section 2 explains the related works on IDSs. Section 3 underlines the proposed IDS procedure. Section 4 evaluates the results and discussion. The conclusion of the work is explained in section 5.

2. Related work

Some of the recent and effective IDS frameworks contributing to security in smart farming are discussed below:

Smart farming embeds advanced computing technologies with general farming operations to enhance performance and improve the overall production rate. With the convergence of the IoT with the smart farming scenario, farming operations have achieved considerable improvement. Since IoT devices are placed directly on the fields, many threats are encountered that require prompt solutions to be protected from cyber-attacks. One methodology has been declared by Ferrag et al. [21] depend on the integration of IoT with smart farming, where the IDS was developed to protect the data from DDoS attacks. The system model included a sensing layer, FCL and the cloud layer. The sensors in the sensing layer captured the data from fields and forwarded it to fog nodes where the IDS framework detected the DDoS attacks. The IDS framework was developed using three DL models: Deep Neural Network (DNN), CNN and Recurrent Neural Network (RNN), each trained on the data sets for classification. The results proved the performance of the model compared to previous work.

Another methodology for IDS to secure the data from agricultural fields was introduced by Raghuvanshi et al. [22]. The IoT sensors were placed on the fields in the methodology to collect agricultural data. The NSL-KDD dataset was utilized as the input to the framework, where initially, pre-processing was carried out by converting all the symbolic features into numeric features. Following this step, feature extraction was performed using the principal component analysis (PCA) technique. The classification was performed using machine learning methods such as Random Forest (RF), Support Vector Machine (SVM) and Linear Regression (LR). The methodology was compared to other ML methods to demonstrate the increased performance achieved.

The technologies used in smart agriculture are effective instruments capable of generating temporal, spatial and time-series data streams collected from fields. These generated data must be protected from adversarial attacks to enhance agricultural productivity. Moso et al. [23] introduced an ensemble anomaly detector called Enhanced Local Selective Combination in Parallel Outlier Ensembles (ELSCP) to accomplish the task. A data-driven unsupervised methodology was presented that was applied to two different case studies, where one dealt with global positioning system (GPS) traces and the other dealt with crop data. While dealing with the crop data, the ELSCP framework predicted the crop's state and detected anomalies present in the data. The experimental outcomes of the model proved that the model was capable of accurately identifying the anomalies related to crop damage.

Establishing a smart farm includes several different equipment that is operated to achieve certain functionality. Anomalies in such equipment decrease the reliability of smart farms and cause various troubles. A method to solve the above issue was introduced by Park et al. [24], integrating the deep learning technique. The technique was established to secure pig house equipment and enhance livestock management. The data accumulated in the equipment, such as the environmental factors, were used for training purposes. The RNN model was utilized in the method for training and classification purposes. Environmental factors such as temperature, ventilation, CO2, humidity, external and radiator temperatures are used for training purposes. The method provided better and correct outcomes related to other current mechanisms.

The main intention of developing the precision farming system is to reduce the burden on farmers and increase the overall net productivity. To achieve this, the agricultural sector embeds multiple devices labeled with specific objectives. One such equipment only provides water to the agricultural field when needed. Thakur et al. [25] introduced a methodology for intrusion identification, where several sensors were utilized to obtain the data. The method also focused on detecting intrusions in the fields before the captured data was forwarded to cloud servers. The methodology utilized and monitored soil data, and different forms of intrusions in the fields were detected. The effectiveness of the approach was proved through experiments.

Several IDS methodologies are proposed in the literature to accurately detect and classify the attacks so that the data can be transmitted securely. Among the introduced methodologies, deep learning based techniques are highly flourished due to their excellent learning capacity. The existing techniques, however, suffer from various issues such as poor data quality, training data overfitting, training data underfitting, non-representative and insufficient training data. All these problems are required to be resolved efficiently to improve the overall growth of agriculture.

3. Proposed methodology

The proposed smart farming architecture is three-tier. The first tier is the agricultural sensor layer, which involves the placement of sensors in agricultural areas to gather environmental information data. The gathered agricultural information is transmitted to the fog layer for analysis. The second stage is the FCL, which consists of Fog nodes, and within each Fog node, the proposed IDS is implemented. The third stage is the cloud computing layer for storing the data and providing the end to end services. The proposed model includes label encoding and data normalization for pre-processing and a fused model of bidirectional gated recurrent unit (Bi-GRU) with a CNN model to detect and classify the intrusions. An attention mechanism is included within the BiGRU model to find the key features responsible for identifying the DDoS attack.

Further, the model classification accuracy is polished using a nature inspired meta-heuristic optimization algorithm called the WHO algorithm. The last layer is the cloud layer, which collects data from fog nodes and offers storage services. Figure 1 shows the proposed methodology.

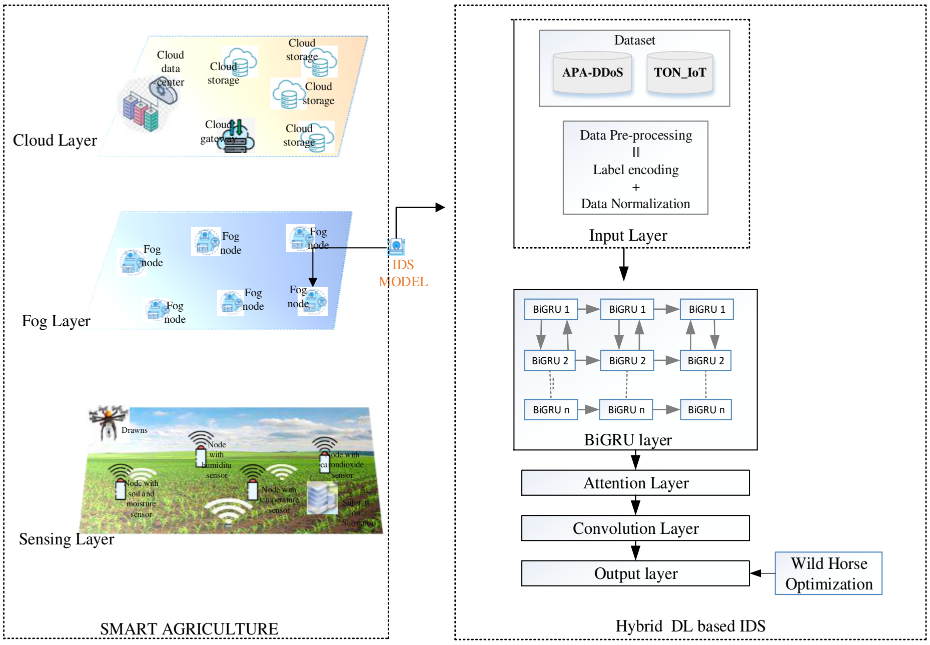

Figure 1: Suggested Methodology

3.1 The BiGRU-CNN based IDS

For the detection of DDoS attacks in the IDS, the ToN-IoT dataset and the APA-DDoS-Attack dataset are used as input data. A number of attacks are included in the dataset used, and the proposed IDS is used to identify the DDoS attack from the dataset. The BiGRU-CNN IDS contains 5 layers: input layer, BiGRU layer, attention layer, convolution layer and output layer.

3.1.1 Input layer

Data from the ToN-IoT dataset and APA-DDoS-Attack dataset are first pre-processed. The pre-processing section contains label encoding and data normalization. In label encoding, non-numerical data are converted to numerical values. The categorical features were converted to numerical values using the label encoder. Each and every categorical value in the dataset was converted to a number using the label encoding technique. The min-max method is the next data normalization process that allows the values to fall within the same range. To normalize data in the range 0 and 1, the expression used is,

$$\overline{Z} = \frac{Z - \min}{\max - \min} \tag{1}$$

Where Z and $\overline{Z}$ are real and normalized data, respectively, $\min$ is the minimum and $\max$ is maximum values.

3.1.2 BiGRU layer

A sequence-processing model called BiGRU consists of two GRUs. One GRU receives input in the forward direction, the other in the reverse direction. The GRU is a modified version of the RNN. Reduced computational cost, training efficiency and simpler structure are the main advantages of GRU. The design of the BiGRU model is depicted in figure 2.

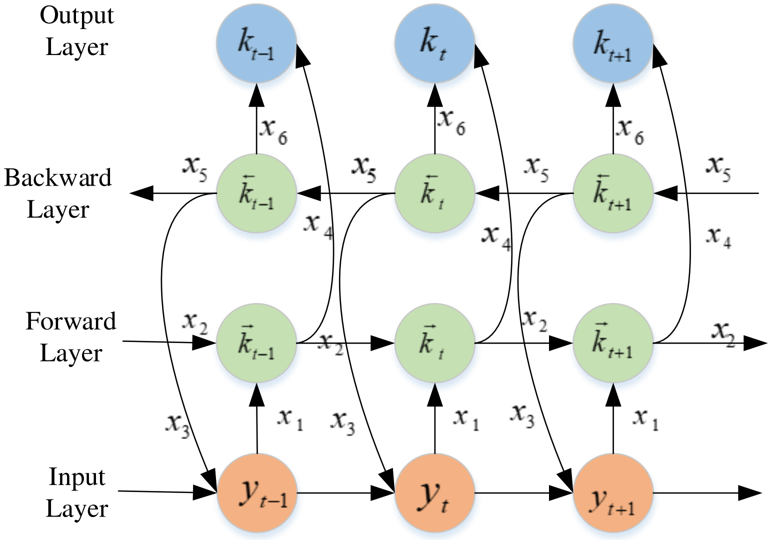

Figure 2: Structure of BiGRU

Equation (2) defines the GRU's underlying computation method,

$$\begin{cases} r_t = \sigma\left(X_r y_t + V_r k_{t-1}\right) \\ z_t = \sigma\left(X_z y_t + V_z k_{t-1}\right) \\ \breve{k}_t = \tanh(X_{\breve{h}} y_t + V_{\breve{h}}(r_t \Theta k_{t-1})) \\ k_t = (1 - z_t)\Theta k_{t-1} + z_t \Theta k_t \end{cases} \tag{2}$$

The sigmoid activation function σ is used to convert intermediate states to the range [0,1], k_{t-1} and k_t are the outputs at the time $t-1$ and t, respectively. y_t is the input arrangement value at the time t. The output state is $\breve{k}_t$; r_t and z_t are the reset and update gates;

X_r, $X_z, X_{\tilde{h}}, V_r, V_Z$ and $V_{\tilde{h}}$ are the coefficient matrices of the weight in every part; *tanh* is a hyperbolic tangent function and Θ is the element wise multiplication._The output k_t of each time step t contains two vectors from forward propagation $\vec{k}_t$ and backward propagation $\overleftarrow{k}_t$, $H_t = [\vec{k}_t, \overleftarrow{k}_t]$.

3.1.3 Attention layer

To strengthen the performance of the IDS, an attention mechanism is added to the output of the BiGRU layer. The most crucial features responsible for IDS are selected using the attention mechanism to identify the DDoS attack. The attention mechanism introduces a weight coefficient according to the importance of identifying the important features for selecting a DDoS attack. The calculation process is shown in equation (3),

$$\alpha_1 = soft\max(w_3 \tanh(W_3 H^T + b_3)) \tag{3}$$

$$f_t = \sum_{i=1}^{T} \alpha_1^t . H_t \tag{4}$$

The output from the BiGRU model after the attention mechanism is given to the convolution layer.

3.1.4 Convolution layer

Multiple convolution kernels are used in the convolution layer to extract deeper features in the intrusion detection system. This layer includes a convolution layer and a pooling layer which is different from other neural networks. For the BiGRU-CNN based IDS, 3 different convolution kernels are used, using maximum pooling to extract the important features. The convolution kernel is related to weight and bias vectors. The *n*th kernel at position *(i,j)* in the *m*th layer calculates the feature value $x_{i,j,n}^m$.

$$x_{i,j,n}^m = f\left(W_n^{m^T} x_{i,j}^{m-1} + b_n^m\right) \tag{5}$$

Where, $x_{i,j}^{m-1}$ is the input patch for the $(m-1)^{th}$ layer, centered at position (i, j). The nth kernel filter›s weight (W) and bias (b) terms, respectively in m^{th} layer. The pooling layer is used to minimize the dimensions and increase robustness. The pooling method that determines the maximum value in the pooling windows uses Max Pooling in the pooling layer.

3.1.5 Output layer

The output of the convolution layer is sent to the fully linked layer of the output layer. The outcome of the convolution layer is given to the softmax classifier, where the softmax function is used for identifying the DDoS attack. The mathematical expression for the softmax function is,

$$z_l = soft\max(vZ + c) \tag{6}$$

Where c is the bias, v is the weight coefficient matrix, and z_l is the attained output that is the DDoS attack.

3.1.6 Parameter tuning using wild horse optimization

In order to optimally tune the bias and weight parameter and thereby reduce the classification error rate, the WHO algorithm is aimed, which improves the accuracy of detecting DDoS attacks.

In WHO, the social behavior of the wild horse is considered for the optimal selection of the parameter, which improves the detection accuracy. Initially, the population is divided into groups, where H symbolizes the number of different attacks and M represents the total number of attacks in the system. The number of stallions equals H, since each group has a leader, and MH represents the population of foals and mares scattered in this group.

$$Y^{i}_{j,H} = 2X\cos(2\pi SX) \times (stlan^{i} - Y^{i}_{j,H}) + stlan^{i} \tag{7}$$

$Y^{i}_{j,H}$ denotes the current position of the mare or foal group member, $stlan$ gives the location of the stallion, S has a stochastic value in the interval -2 to 2 and X denotes the adaptive model projected as,

$$\begin{aligned} R &= \vec{T}_1 < TDR; \quad IDX = (p == 0) \\ X &= T_2 \Theta IDX + \vec{T}_3 \Theta(\approx IDX) \end{aligned} \tag{8}$$

Where, $R, \vec{T}_1, \vec{T}_3\, and T_2$ has an arbitrary value in the range of [0, 1]. The value of TDR reduces to 0 at the end of iteration starting from 1 given as,

$$TDR = 1 - it \times \left(\frac{1}{\max it} \right) \tag{9}$$

The maximal iteration is represented as $\max it$. For a foal to move from group j to a transitory group when the foal moves to group i. The crossover operator is used to simulate horse mating behavior:

$$\begin{aligned} &Y^{R}_{H,K} = Crossover\left(Y^{q}_{H,j'}, Y^{X}_{H,i}\right) \quad j \neq i \neq k, q = p = end, \\ &Crossover = Mean \end{aligned} \tag{10}$$

The group leader struggles a lot for the water hole, and the others should wait till the dominant group leaves the water hole:

$$\overline{Stlan_{H_j}} = T \begin{cases} 2X\cos(2\pi SX) \times (WH - Stlan_{H_j}) \\ +WH \;\; if \;\; T_3 > 0.5 \\ 2X\cos(2\pi SX) \times (WH - stlan_{H_j}) \\ -WH \;\; if \;\; T_3 \leq 0.5 \end{cases} \tag{11}$$

WH is the location of the water hole, and the leader who finds the location is represented as $\overline{Stlan_{H_j}}$. Based on the fitness value, the leader is selected in the next phases, and the leader's position and the selected member are indicated as follows:

$$\overline{Stlan_{H_j}} = \begin{cases} Y_{H,j} & if \quad \cos t(Y_{H,j}) < \cos t(Stlan_{H_j}) \\ Stlan_{H_j} & if \quad \cos t(Y_{H,j}) > \cos t(Stlan_{H_j}) \end{cases} \tag{12}$$

The fitness function used in WHO is the classification error rate, and the goal of the WHO algorithm is to minimize the classification error rate. The flow chart of the proposed WHO is shown in Figure 3,

$$\begin{aligned} fitness(y_i) &= classificationerrorrate(y_i) \\ &= \frac{count\ of\ misidentified\ attacks}{total\ number\ of\ attacks} * 100 \end{aligned} \tag{13}$$

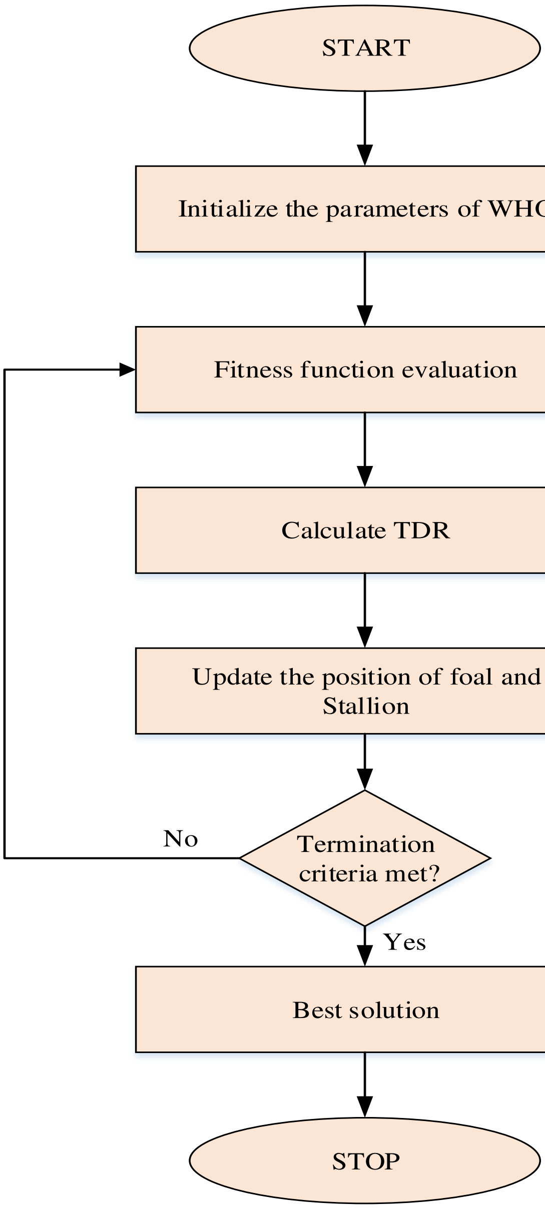
START
Initialize the parameters of WHO
Fitness function evaluation
Calculate TDR
Update the position of foal and Stallion
Termination criteria met?
No
Yes
Best solution
STOP

Figure 3: Flowchart of WHO

The suggested system helps to detect intrusions in smart farming systems with better performance. It mainly detects DDoS and DoS attacks.

4. Results and Discussion

This section analyses the effectiveness of the suggested IDS using different performance metrics. The evaluation metrics used for the analysis are accuracy, precision, detection rate, F-measure, ROC and confusion matrix.

4.1 APA-DDoS Attack Dataset

As the connection between devices increased, the major challenge was the detection of attacks, for an intrusion detection system (IDS) was developed. Machine learning techniques have been developed to find the attacks which need access to attack patterns. The APA-DDoS attack contains different types of DDoS attacks. The dataset holds mainly ACK and PUSH-ACK DDoS attacks [26]. To improve the detection rate, the APA-DDoS attack dataset was used. 70% of the dataset is used for training, while 30% is used for testing.

4.2 ToN-IoT Dataset

Telemetry datasets of Internet of things (IoT) and Industrial IoT (IIoT) sensors are included in the ToN-IoT dataset and split into 70% and 30% for training and testing sets. The dataset is allocated in the ratio of 70:30 for training and testing. The dataset contains 43 features with 9 types of attacks and a normal vector. Attacks in IoT environments like Backdoor, MITM, DDoS, DoS, Injection, Password, Scanning, XSS and Ransomware are included in the dataset.

4.3 Evaluation Metrics

In estimating the suggested IDS, the chosen performance metrics play a vital role. The evaluation metrics selected for the estimation are Accuracy, Precision, Detection rate, F-score, and confusion matrix.

4.3.1 Accuracy

It is an amount of the IDS detection accuracy, which is described as the ratio of correctly detected observations to total observations in the test set.

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

4.3.2 Detection Rate

The detection rate or recall is the number of detected positive cases divided by the entire number of events in the test set.

$$Detection\ Rate/\mathrm{Re}call = \frac{TP}{TP + FN} \tag{15}$$

4.3.3 Precision

Precision gives the ratio of the number of detected attacks divided by the system's total number of attacks.

$$\mathrm{Pr}ecision = \frac{TP}{FP + TP} \tag{16}$$

4.3.4 F-measure

F-measure represents the average precision as well as recall. It measures the accuracy of the suggested IDS system using precision and recall.

$$F - measure = 2 * \frac{\mathrm{Re}call * \mathrm{Pr}ecision}{\mathrm{Pr}ecision + \mathrm{Re}call} \tag{17}$$

4.4 Performance Analysis based on APA-DDoS Attack

This section compares the performance of the proposed IDS to other current approaches [27] in terms of recall, accuracy, F-measure and precision. For the analysis, the APA-DDoS Attack dataset is utilized.

Table 1: Comparison analysis of Accuracy, Precision, F-measure and Detection rate for

APA-DDoS Attack dataset.

Methods	**Accuracy**	**Precision**	**F-measure**	**Detection rate**
SLSTM	98.98	94.11	94.14	94.12
DT	96.92	78.31	77.99	77.85
NB	75.36	69.19	63.47	63.8
RF	97.47	78.4	77.45	76.62
PROPOSED	99.35	99.90	99.08	98.99

Table 1 gives a comparative analysis of the proposed system with current methods. The proposed system's detection rate, precision, accuracy, and F-measure are compared with the existing methods for the APA-DDoS attack dataset.

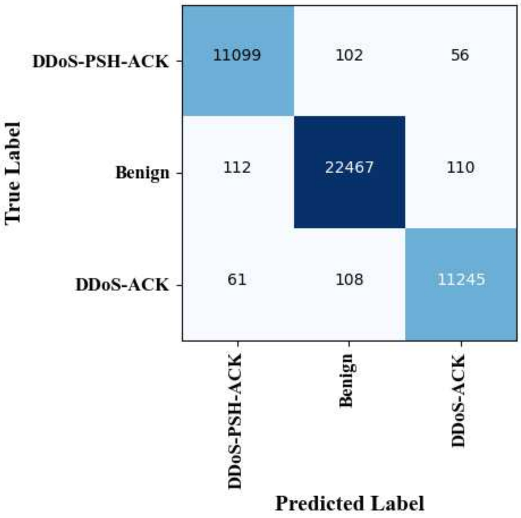

Figure 4: Confusion matrix of APA-DDoS Attack Dataset

The confusion matrix (CM) of the APA-DDoS attack dataset is shown in figure 4. The CM compares the values predicted by the suggested system with the original target values.

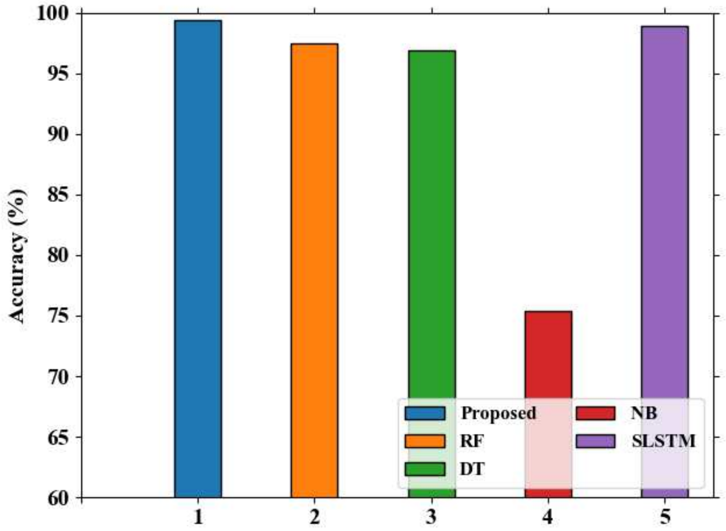

Figure 5: Performance Analysis of Accuracy

The suggested system achieved a good performance in terms of accuracy with 99.35%. Figure 5 gives an accurate performance analysis of the suggested and existing methods.

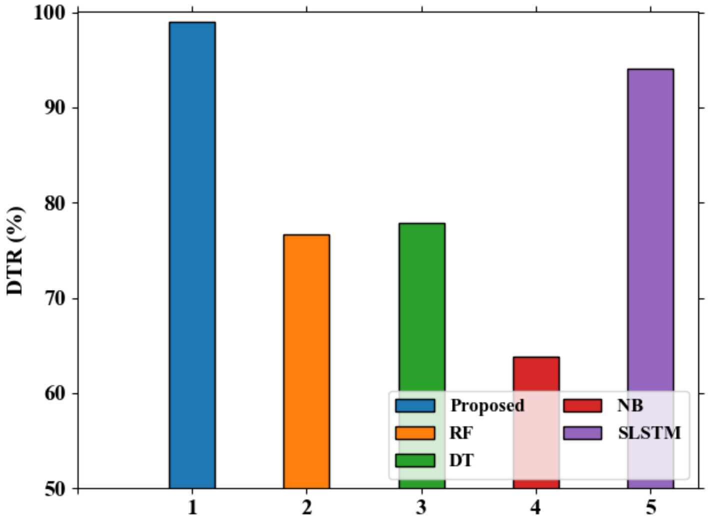

Figure 6: Performance Analysis Detection Rate

Figure 6 gives the performance evaluation in terms of the detection rate of the suggested system with numerous present methods. The suggested system attained a 98.99% detection rate, while the existing methods achieved only 77.85% for DT, 63.8% for NB, 94.12% for SLSTM and RF 76.62%. So the suggested system outperformed in terms of the detection rate.

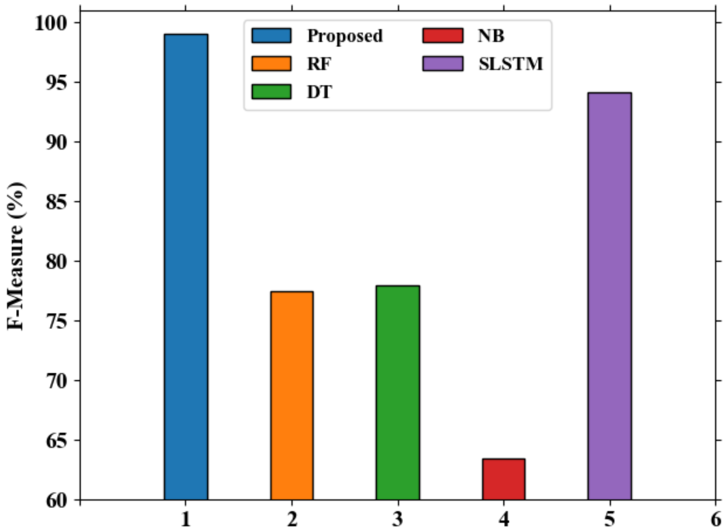

Figure 7: Performance Analysis of F-measure

Figure 7 gives the performance evaluation in terms of the F-measure of the suggested system with various existing methods. The suggested system attained 99.08% F-measure, while the existing methods have achieved only 77.99% for DT, 63.47% for NB, 94.14% for SLSTM and RF with 77.45%. So the suggested system outperformed in terms of F-measure.

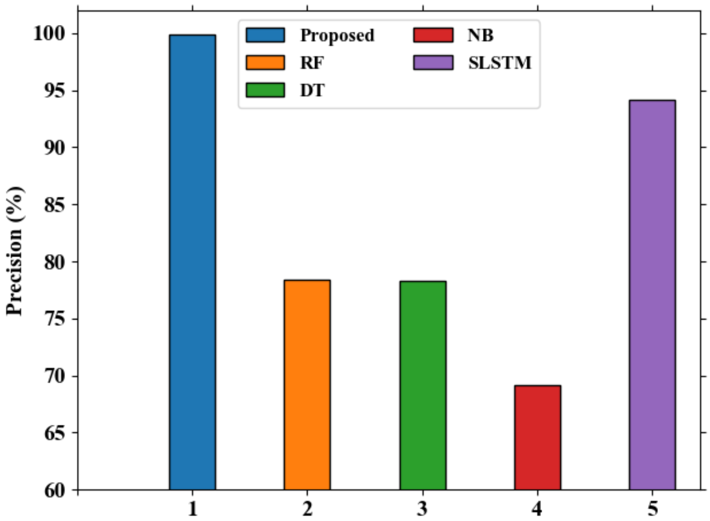

Figure 8: Performance Analysis of Precision

Figure 8 gives the performance evaluation based on the precision of the suggested system with numerous existing methods. The suggested system attained 99.90% precision, while the existing methods have achieved only 78.4% for RF, 78.31% for DT, 94.11% for SLSTM and NB 69.19%. So the suggested system outperformed in terms of precision.

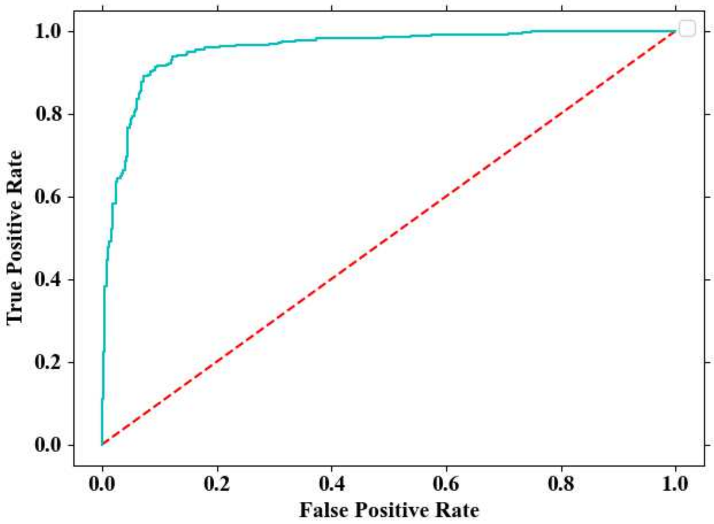

Figure 9: ROC curve for APA-DDoS-Attack dataset

Figure 9 depicts the receiver operating characteristic (ROC) curve, which is a representation of the rate of true positives vs. the rate of false positives. The ROC curve examines the performance of the classifier. Then, the analysis of training accuracy is shown in figure 10.

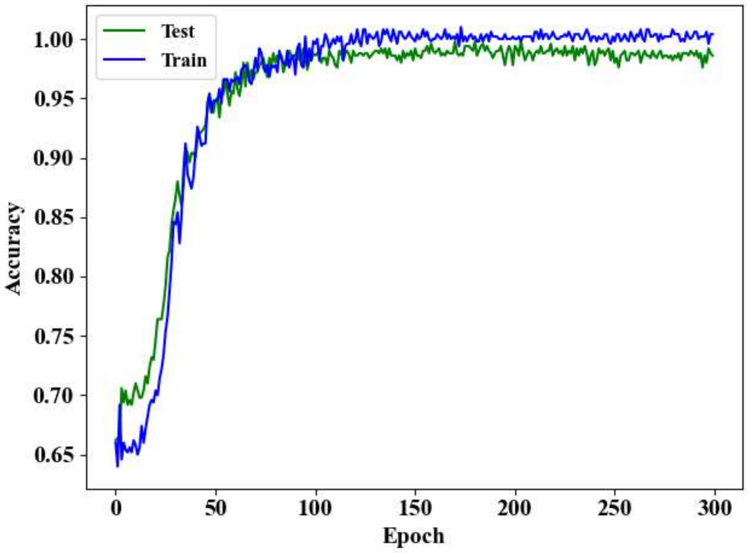

Figure 10: Training accuracy vs testing accuracy for the suggested system with APA-DDoS Attack dataset

Figure 10 shows the training accuracy with the testing for the suggested system for the APA-DDoS attack dataset. The suggested system achieved a better accuracy with 99.35%.

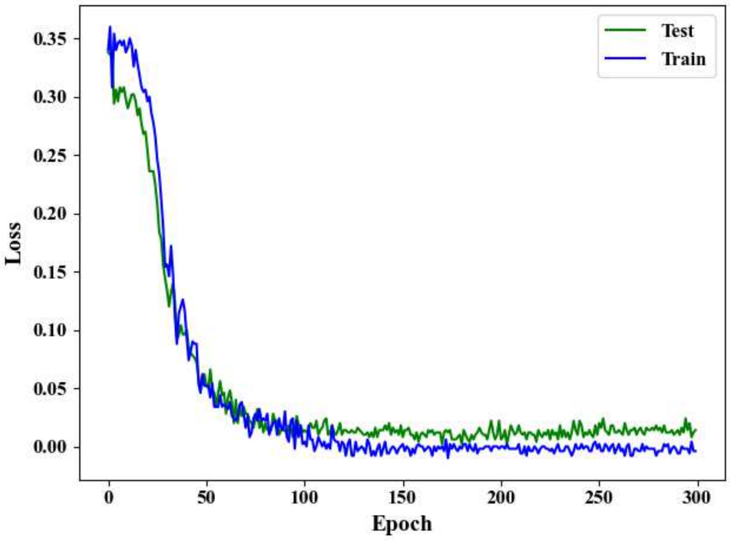

Figure 11: Training loss vs testing loss for the suggested system
With APA-DDoS Attack dataset

The suggested system got low loss in testing and training. The training loss vs. a testing loss of the suggested system for the APA-DDoS Attack dataset is plotted in figure 11.

4.5 Performance Analysis based on ToN-IoT Dataset

This section compares the suggested IDS's performance to other current approaches in terms of recall, accuracy, F-measure and precision. For the analysis, the ToN-IoT dataset is utilized.

Table 2: Comparison analysis of Accuracy, Precision, F-measure and Detection rate for APA-DDoS Attack dataset.

Methods	**Accuracy**	**Precision**	**Detection rate**	**F-measure**
RF	97.81	87.5	85.43	86.41
DT	95.34	74.42	80.0	76.33
NB	90.62	77.68	77.7	72.43
SLSTM	98.64	98.94	98.00	98.87
PROPOSED	99.71	99.89	99.05	99.02

Table 2 gives the comparison analysis of the suggested system with various existing methods like NB, DT, RF and SLSTM.

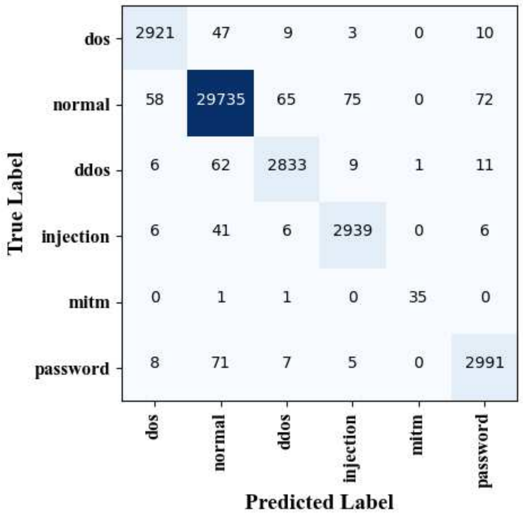

Figure 12: Confusion matrix of ToN-IoT Dataset

The confusion matrix (CM) of the ToN-IoT dataset, which indicates the performance of the classification, is shown in figure 12. The rows in CM symbolize the true label, and the columns symbolize the predicted labels.

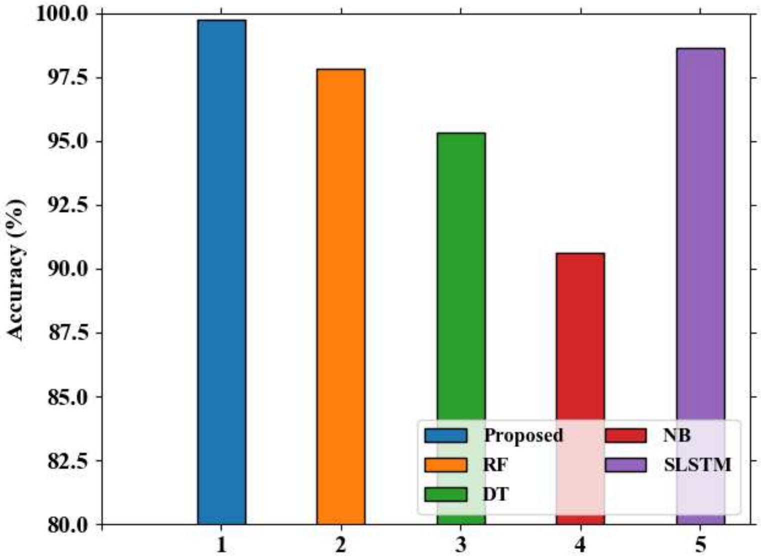

Figure 13: Analysis of Accuracy Performance

Figure 13 gives a performance evaluation in terms of the accuracy of the suggested system with numerous existing methods. The suggested system attained 99.71% accuracy, while the existing methods achieved only 95.34% for DT, 90.62% for NB, 98.64% for SLSTM and RF 97.81%. So the suggested system outperformed in terms of accuracy.

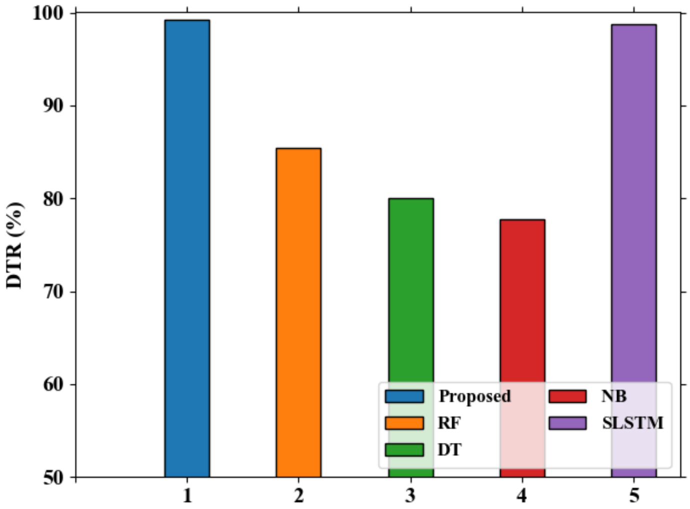

Figure 14: Analysis of Detection rate Performance

The suggested system achieved good performance in terms of a detection rate of 99.02%. Figure 14 gives the detection rate performance analysis of the suggested and present methods.

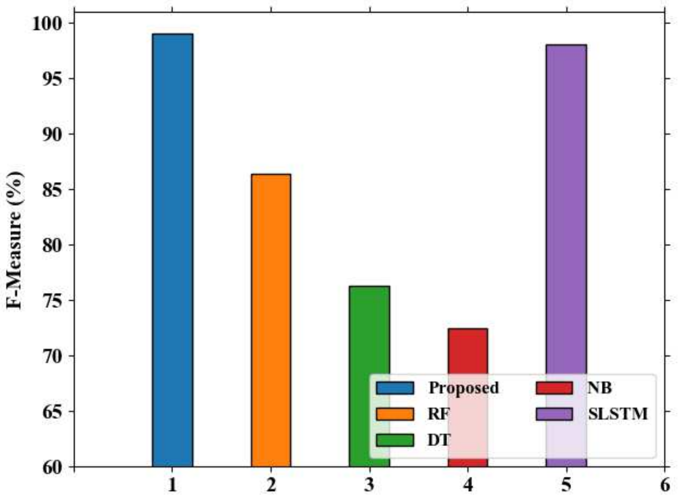

Figure 15: Analysis of F-measure Performance

Figure 15 gives the performance evaluation in terms of the F-measure of the suggested system with numerous present methods. The suggested system attained 99.05% F-measure, while the existing methods have achieved only 76.33% for DT, 72.43% for NB, 98% for SLSTM and RF with 86.41%. So the proposed system outperformed in terms of F-measure.

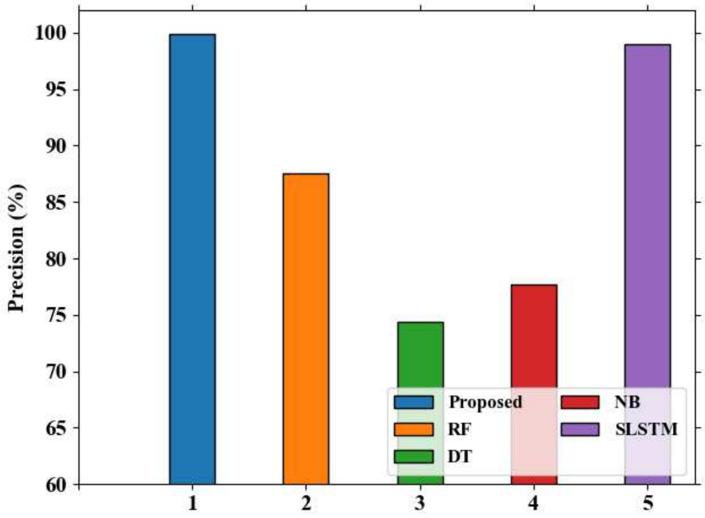

Figure 16: Analysis of Precision Performance

Figure 16 shows the performance evaluation in terms of precision. The various existing methods used are DT, RF, NB and SLSTM, with precisions of 74.42%, 87.5%, 77.68% and 98.94%, respectively. The proposed system attained 99.89% precision.

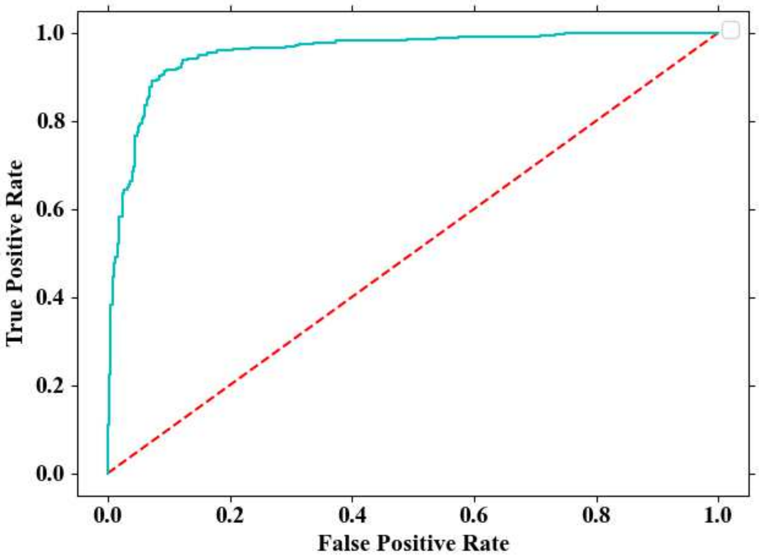

Figure 17: Graph of ROC curve for ToN-IoT dataset

The classifier outcomes were evaluated using the Receiver operating characteristic (ROC), which gives the performance assessment of multiclass vectors in a dataset. To represent the ROC, the TP and the FP rates are considered. The ROC curve of the proposed system is shown in figure 17.

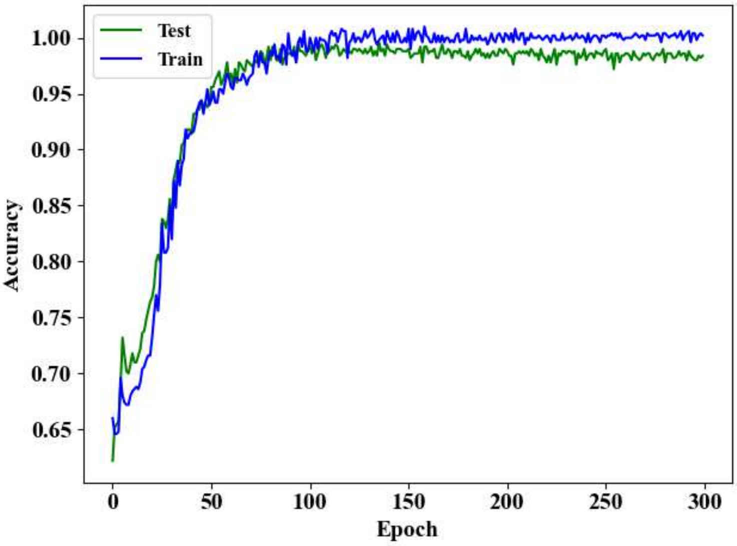

Figure 18: Accuracy of Training vs. testing for the suggested system with APA-DDoS Attack dataset

The accuracy of training and testing of the suggested system for the ToN-IoT dataset is plotted in figure 18. The suggested system with BiGRU-CNN achieved 99.71% accuracy.

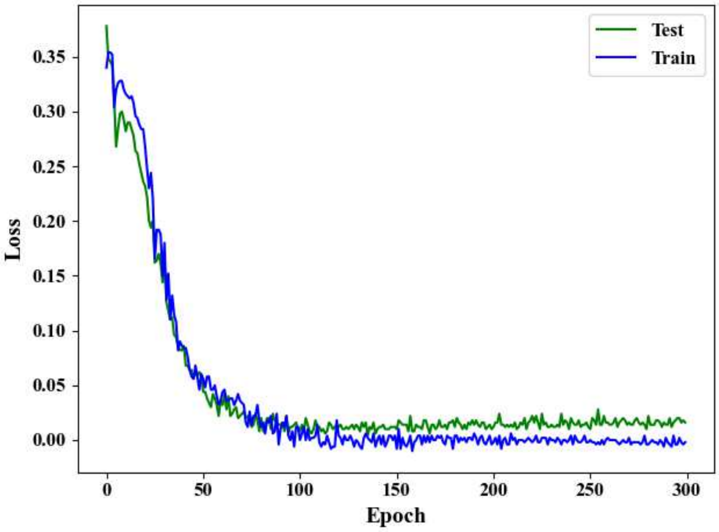

Figure 19: Training loss vs testing loss for the proposed system with APA-DDoS Attack dataset

The proposed system got low loss in testing and training. The training loss vs. a testing loss of the proposed system for the ToN-IoT dataset is plotted in figure 19.

5. Conclusion

In this research, an effective intrusion detection system for DDoS attack detection for smart agriculture was developed. The data collected are pre-processed using data normalization and label encoding. A fused model of CNN with the bidirectional gated recurrent unit (Bi-GRU) model to detect as well as classify the intrusions. The BiGRU model includes an attention mechanism to find the most crucial features responsible for identifying the DDoS attack. Further, the model’s classification accuracy is enhanced with the use of a nature inspired metaheuristic optimization algorithm called the Wild Horse Optimization (WHO) algorithm. Standard performance metrics are used to calculate the intrusion detection system (IDS) performance. The developed IDS model can detect the DDoS attack in the smart agriculture and attained an accuracy of 99.35% for APA-DDoS-Attack and ToN-IoT datasets with 99.71% accuracy. In the future, this technique can be further improved to detect other IoT attacks.

Data availability statement

The entire implementations of the work will be carried out in Python platform. The major performance metrics such as accuracy, precision, recall, f-measure and ROC will be computed and compared with the recent techniques relevant to intrusion detection in IoT enabled smart farming

The data that support this finding of this study are openly available at the following URL/DOI:

https://www.unb.ca/cic/datasets/ddos-2019.html
https://cloudstor.aarnet.edu.au/plus/s/ds5zW91vdgjEj9i.

Code availability

The code that supports this finding of this study are openly available at the following URL/DOI:
https://github.com/keerthikethineni/IDin-IOT-based-Smart-Farming-using-HDLF.git

Author Contributions

Keerthi: Conceptualization, Data Curation, Formal Analysis, Investigation, Resources, Software, Writing original draft. Pradeepini: Methodology, Project administration, Supervision, Validation, Visualization, Writing-Review&editing, Funding acquisition.

Funding

No Funding was received

Acknowledgements

We declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere.

Declarations

Conflict of interest There is no conflict of interest in the present research work.
Informed Consent Author and Co-author are well aware about publication.

References

[1] Yang, X., Shu, L., Chen, J., Ferrag, M.A., Wu, J., Nurellari, E. and Huang, K.: A survey on smart agriculture: Development modes, technologies, and security and privacy challenges. IEEE/CAA Journal of Automatica Sinica. 8(2), 273-302 (2021).

[2] de Araujo Zanella, A.R., da Silva, E. and Albini, L.C.P.: Security challenges to smart agriculture: Current state, key issues, and future directions. Array. 8, 100048 (2020).

[3] Kumar, P., Gupta, G.P. and Tripathi, R.: PEFL: Deep Privacy-Encoding-Based Federated Learning Framework for Smart Agriculture. IEEE Micro. 42(1), 33-40 (2021).

[4] Suhaimi, A.F., Yaakob, N., Saad, S.A., Sidek, K.A., Elshaikh, M.E., Dafhalla, A.K., Lynn, O.B. and Almashor, M.: IoT Based Smart Agriculture Monitoring, Automation and Intrusion Detection System. In Journal of Physics: Conference Series, IOP Publishing. 1962(1), 012016 (2021).

[5] Fróna, D., Szenderák, J. and Harangi-Rákos, M.: The challenge of feeding the world. Sustainability. 11(20), 5816 (2019).

[6] Alsoufi, M.A., Razak, S., Siraj, M.M., Nafea, I., Ghaleb, F.A., Saeed, F. and Nasser, M.: Anomaly-based intrusion detection systems in iot using deep learning: A systematic literature review. Applied Sciences. 11(18), 8383 (2021).

[7] Cicioğlu, M. and Çalhan, A.: Smart agriculture with internet of things in cornfields. Computers & Electrical Engineering. 90, 106982 (2021).

[8] Ferrag, M.A., Shu, L., Friha, O. and Yang, X.: Cyber security intrusion detection for agriculture 4.0: Machine learning-based solutions, datasets, and future directions. IEEE/CAA Journal of Automatica Sinica. 9(3), 407-436 (2021).

[9] Kumar, R., Mishra, R., Gupta, H.P. and Dutta, T.: Smart sensing for agriculture: Applications, advancements, and challenges. IEEE Consumer Electronics Magazine, 10(4), 51-56 (2021).

[10] Kumar, M. and Vikas Reddy, S.: Intrusion Detection and Prevention System for Iot. European Journal of Molecular & Clinical Medicine. 7(8), 2983-2991 (2020).

[11] Bhatt, H., Bhushan, B. and Kumar, N.: IOT: The current scenario and role of sensors involved in smart agriculture. International Journal of Recent Technology and Engineering. 8(4), 12011-12023 (2019).

[12] Riaz, A.R., Gilani, S.M.M., Naseer, S., Alshmrany, S., Shafiq, M. and Choi, J.G.: Applying Adaptive Security Techniques for Risk Analysis of Internet of Things (IoT)-Based Smart Agriculture. Sustainability. 14(17), 10964 (2022).

[13] Bhatnagar, V., Singh, G., Kumar, G. and Gupta, R.: Internet of Thingsin Smart Agriculture: Applications and Open Challenges. (2020).

[14] Yadahalli, S., Parmar, A. and Deshpande, A.: Smart intrusion detection system for crop protection by using Arduino. In 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE. 405-408 (2020).

[15] Murugesan, M.: Smart Agriculture Monitoring System. Turkish Journal of Computer and Mathematics Education (TURCOMAT). 10(3), 1001-1005 (2019).

[16] Tao, W., Zhao, L., Wang, G. and Liang, R.: Review of the internet of things communication technologies in smart agriculture and challenges. Computers and Electronics in Agriculture. 189, 106352 (2021).

[17] Kumar, K.N., Pillai, A.V. and Narayanan, M.B.: Smart agriculture using IoT. Materials Today: Proceedings. (2021).

[18] Eskandari, M., Janjua, Z.H., Vecchio, M. and Antonelli, F.: Passban IDS: An intelligent anomaly-based intrusion detection system for IoT edge devices. IEEE Internet of Things Journal. 7(8), 6882-6897 (2020).

[19] Salim, C. and Mitton, N.: Image similarity based data reduction technique in wireless video sensor networks for smart agriculture. In International Conference on Advanced Information Networking and Applications, Springer, Cham. 448-459 (2021).

[20] Abraham, G., Raksha, R. and Nithya, M.: Smart Agriculture Based on IoT and Machine Learning. In 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), IEEE. 414-419 (2021).

[21] Ferrag, M.A., Shu, L., Djallel, H. and Choo, K.K.R.: Deep learning-based intrusion detection for distributed denial of service attack in Agriculture 4.0. Electronics. 10(11), 1257 (2021).

[22] Raghuvanshi, A., Singh, U.K., Sajja, G.S., Pallathadka, H., Asenso, E., Kamal, M., Singh, A. and Phasinam, K.: Intrusion detection using machine learning for risk mitigation in IoT-enabled smart irrigation in smart farming. Journal of Food Quality. 2022, (2022).

[23] Moso, J.C., Cormier, S., de Runz, C., Fouchal, H. and Wandeto, J.M.: Anomaly Detection on Data Streams for Smart Agriculture. Agriculture. 11(11), 1083 (2021).

[24] Park, H., Park, V. and Kim, S.: Anomaly detection of operating equipment in livestock farms using deep learning techniques. Electronics. 10(16), 1958 (2021).

[25] Thakur, D., Kumar, Y. and Vijendra, S.: Smart irrigation and intrusions detection in agricultural fields using IoT. Procedia Computer Science. 167, 154-162 (2020).

[26] https://www.kaggle.com/datasets/yashwanthkumbam/apaddos-dataset

[27] Kumar, R., Kumar, P., Tripathi, R., Gupta, G.P., Gadekallu, T.R. and Srivastava, G.: SP2F: A secured privacy-preserving framework for smart agricultural Unmanned Aerial Vehicles. Computer Networks. 187: 107819 (2021).