ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

Multiclass IoT Intrusion Detection Using a Hybrid CNN–BiLSTM–Attention Architecture

Umaru Mustapha Audu[1]; Zayyanu Umar[2]; Karatu Musa Tanimu[3]; Sirajo Abdullahi Bakura[4]

[1,2,3,4]Department of Computer Science, Federal University Birnin Kebbi, 860222, Nigeria

Publication Date: 2026/09/21

Abstract: The rapid expansion of Internet of Things (IoT) technologies has increased the complexity of network environments and exposed connected devices to diverse cybersecurity threats. Conventional intrusion detection approaches often face limitations in identifying multiple attack categories, particularly in highly imbalanced IoT network traffic. This study proposes a hybrid Convolutional Neural Network–Bidirectional Long Short-Term Memory with Attention (CNN–BiLSTM–Attention) model for multiclass intrusion detection in IoT environments. The study employed the BoT-IoT dataset, which contains 2,934,817 network-flow records representing normal traffic and multiple attack categories, including Denial of Service (DoS), Distributed Denial of Service (DDoS), Reconnaissance, and Theft. Data preprocessing involved data cleaning, categorical encoding, feature normalization, and Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance. The proposed architecture combines CNN for local feature extraction, BiLSTM for learning sequential dependencies, and an attention mechanism for emphasizing informative learned representations. The model was evaluated using accuracy, precision, recall, and F1-score, alongside comparisons with CNN, BiLSTM, and CNN–BiLSTM baseline models. The proposed CNN–BiLSTM–Attention model achieved a test accuracy of 96.27%, with macro-averaged precision, recall, and F1-score of 86.06%, 94.68%, and 89.64%, respectively. The model achieved its strongest class-level performance on Reconnaissance attacks, with an F1-score of 98.95%, while the Theft class recorded the lowest F1-score of 66.67% because of its extremely limited representation in the evaluation data. The results further indicate some classification confusion between DoS and DDoS attacks, reflecting similarities in their network traffic characteristics. Overall, the findings demonstrate that integrating convolutional feature extraction, bidirectional temporal learning, and attention mechanisms can provide an effective approach to multiclass IoT intrusion detection, although rare attack categories remain challenging and require further investigation.

Keywords: *Internet of Things (IoT); Intrusion Detection System; Deep Learning; Convolutional Neural Network; Bidirectional Long Short-Term Memory; Attention Mechanism; Multiclass Classification; BoT-IoT; SMOTE; Cybersecurity.*

How to Cite: Umaru Mustapha Audu; Zayyanu Umar; Karatu Musa Tanimu; Sirajo Abdullahi Bakura (2026) Multiclass IoT Intrusion Detection Using a Hybrid CNN–BiLSTM–Attention Architecture. *International Journal of Innovative Science and Research Technology*, 11(9), 541-550. https://doi.org/10.38124/ijisrt/26sep806

I. INTRODUCTION

The rapid proliferation of the Internet of Things (IoT) has transformed modern computing and communication environments by enabling large numbers of interconnected devices to collect, exchange, and process data with limited human intervention. IoT technologies are increasingly integrated into smart homes, healthcare, transportation, agriculture, industrial systems, and smart-city infrastructures, creating highly distributed and heterogeneous computing environments. This expansion has improved automation, connectivity, and service delivery, but it has simultaneously increased the number of potential entry points available to cyber attackers. The diversity of IoT devices, communication protocols, network configurations, and computational capabilities creates a complex security environment in which conventional security mechanisms may struggle to provide comprehensive protection. Consequently, the security of IoT infrastructures has become an important research concern, particularly because compromised devices can be exploited to launch denial-of-service attacks, reconnaissance activities, botnet operations, data theft, and other forms of malicious behaviour [1]–[3].

Intrusion Detection Systems (IDSs) constitute an important layer of cybersecurity because they are designed to monitor network activities and identify patterns associated with malicious or abnormal behaviour. Traditional IDS approaches have commonly relied on predefined signatures or manually engineered rules, making them effective for previously known attack patterns but less suitable for detecting evolving or previously unseen threats. The dynamic nature of IoT networks further complicates this problem because traffic patterns can vary considerably according to device type, application, communication protocol, network

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

conditions, and user behaviour. Consequently, machine learning and deep learning techniques have increasingly been investigated as alternatives capable of learning discriminative representations directly from network traffic. Deep learning architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Graph Neural Networks (GNNs) have demonstrated considerable potential for automatically extracting complex representations from IoT traffic and improving automated intrusion detection [4]–[6].

The application of deep learning to IoT intrusion detection has produced promising results across different datasets and attack scenarios. For example, Jullian *et al.* [3] developed a distributed deep-learning framework using Feed-Forward Neural Networks and LSTM models and evaluated the framework using the NSL-KDD and BoT-IoT datasets. Their findings demonstrated that deep learning could effectively detect different forms of malicious traffic, with reported accuracy reaching 99.95% across the evaluated configurations. Such findings demonstrate the ability of deep neural networks to learn complex patterns from network traffic and provide a foundation for automated cyber-attack detection. Similarly, Sahu *et al.* [7] investigated a hybrid CNN-LSTM approach for IoT attack detection and reported an accuracy of approximately 96%, demonstrating the potential of combining convolutional feature extraction with sequential modelling. Zhong *et al.* [8] also employed Text-CNN and GRU-based sequential modelling to represent network and application-layer traffic, further demonstrating the usefulness of sequential deep-learning approaches for IoT intrusion detection.

Other studies have explored alternative deep-learning architectures to capture different characteristics of network behaviour. Altaf *et al.* [9], for instance, investigated a Sequential Gated Graph Convolutional Network for IoT botnet detection. The approach was designed to capture spatial and temporal relationships in IoT traffic and demonstrated improved performance over conventional approaches across evaluated datasets. Graph-based approaches are particularly relevant because IoT environments contain relationships among devices, communication endpoints, and traffic flows that may not be adequately represented when network records are considered independently. Similarly, Gelenbe and Nakip [10] proposed an online compromised-device identification system based on traffic-based sequential learning and evaluated it using Mirai botnet traffic. Their approach demonstrated high balanced accuracy while maintaining relatively low online training and execution requirements. However, such approaches primarily concentrate on identifying compromised devices or specific forms of malicious behaviour rather than providing comprehensive multiclass classification of diverse intrusion categories.

The literature also demonstrates the effectiveness of models specifically designed to capture temporal dependencies in network traffic. Ramzan *et al.* [11], for example, evaluated RNN, LSTM, and GRU models for Distributed Denial-of-Service (DDoS) attack detection using contemporary intrusion-detection datasets and reported accuracy values around 0.99. Their findings demonstrate the capacity of recurrent architectures to model sequential characteristics of malicious network traffic, although the investigation was focused primarily on DDoS detection rather than simultaneous classification of multiple intrusion categories. Alasmary *et al.* [12] similarly developed ShieldRNN, a distributed flow-based architecture for DDoS detection in IoT networks, using sequence-based analysis and majority voting. Although the approach achieved very high detection performance, its primary focus remained DDoS detection. These findings indicate that deep learning can achieve strong performance when the detection problem is narrowly defined, but they also highlight the need for models capable of distinguishing among several attack types within a unified IoT intrusion-detection framework.

A related challenge concerns the distinction between binary and multiclass intrusion detection. A considerable proportion of existing IoT IDS research treats the problem as binary classification, where network traffic is assigned to either a benign or malicious category. Although binary detection is useful for determining whether suspicious activity is present, it provides limited information about the specific nature of the detected threat. In operational IoT environments, distinguishing between DDoS, Denial-of-Service (DoS), reconnaissance, data theft, botnet activity, and other attacks can be important because different threats may require different investigation and response strategies. Studies such as Sundar *et al.* [13] demonstrated the effectiveness of supervised machine-learning approaches for identifying malicious versus benign IoT traffic, while Salman *et al.* [14] investigated machine-learning techniques for IoT device identification and abnormal traffic detection. However, these approaches primarily emphasize binary abnormality detection rather than comprehensive multiclass identification of different intrusion categories.

The limitation of binary classification is also evident in more recent deep-learning research. Kandhro *et al.* [15] investigated malicious intrusion and attack detection in IoT-enabled cybersecurity infrastructures, while Mohammad and Bharati [16] developed a hybrid deep-learning approach incorporating CNN and BiLSTM architectures for IoT attack detection. These studies demonstrate the continuing progress of deep learning in cybersecurity and the ability of hybrid architectures to learn complex traffic representations. Nevertheless, the literature reviewed in the present study indicates that a substantial number of existing approaches continue to aggregate malicious activities into broad categories or concentrate on particular attack families. This limits the ability of an IDS to provide fine-grained information about the specific attack responsible for observed network behaviour. Therefore, the transition from binary detection toward robust multiclass intrusion classification remains an important research direction.

Beyond the classification problem itself, class imbalance represents another major challenge in IoT intrusion detection. Network intrusion datasets rarely contain

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

equal numbers of examples for every class. In practice, common attacks may generate millions of observations, whereas legitimate traffic and rare attack categories may contain only a small number of instances. Such imbalance can cause a learning algorithm to become strongly influenced by majority classes and consequently achieve high overall accuracy while performing poorly on minority classes. Sayegh *et al.* [17] specifically investigated the use of LSTM-based intrusion detection together with feature selection and Synthetic Minority Over-sampling Technique (SMOTE) to address imbalanced intrusion-detection data. Widodo *et al.* [18] similarly examined multiclass intrusion detection using SMOTE on an imbalanced dataset, demonstrating the importance of addressing class distribution when evaluating multiclass intrusion-detection systems. These studies reinforce the need to consider class-specific precision, recall, and F1-score rather than relying exclusively on overall accuracy.

Feature representation and temporal dependency modelling constitute another important consideration in IoT intrusion detection. Network traffic contains relationships among individual attributes as well as dependencies that emerge across successive traffic observations. CNNs are effective at learning local patterns and relationships among input features, whereas recurrent architectures such as LSTM and BiLSTM can model sequential dependencies. The combination of these complementary capabilities has therefore attracted increasing research attention. Guan *et al.* [19], for example, developed a two-tiered IoT anomaly-classification framework incorporating CNN-BiLSTM modelling, demonstrating the relevance of combining convolutional and recurrent representations for IoT traffic analysis. Omarov *et al.* [20] went further by proposing a one-dimensional Conv-BiLSTM architecture augmented with an Attention mechanism specifically for IoT intrusion detection. Their work demonstrates the relevance of integrating convolutional feature extraction, bidirectional sequential learning, and attention-based representation weighting for distinguishing diverse intrusion patterns. The direct relevance of this architecture to the present study provides an important foundation for investigating the CNN-BiLSTM-Attention combination in a multiclass BoT-IoT setting.

Attention mechanisms provide an additional means of improving the representation learned by deep neural networks by assigning greater weight to informative elements of a learned representation. In sequential intrusion-detection problems, an attention mechanism can assist the network in emphasizing informative temporal representations or feature patterns rather than treating every learned representation with equal importance. The integration of CNN, BiLSTM, and Attention is therefore conceptually attractive because the three components address complementary aspects of the detection problem: CNN can extract local feature-level patterns, BiLSTM can model dependencies in both forward and backward temporal directions, and Attention can emphasize representations that contribute more strongly to the final classification decision. Omarov *et al.* [20] provide direct evidence for the feasibility of this integrated architecture in IoT intrusion detection, while Guan *et al.* [19] further demonstrate the usefulness of CNN-BiLSTM modelling for IoT anomaly classification. However, the existence of related architectures does not eliminate the need to investigate their effectiveness under different datasets, class distributions, preprocessing strategies, and multiclass experimental configurations.

The choice of an appropriate benchmark dataset is consequently important for evaluating multiclass IoT intrusion-detection models. The Botnet Internet of Things (BoT-IoT) dataset was developed as a large-scale benchmark for cybersecurity research and contains network-flow observations representing both legitimate and malicious IoT traffic. The dataset includes several attack categories, including DoS, DDoS, reconnaissance, and data-theft-related activities, making it particularly relevant to studies concerned with multiclass intrusion detection. Its large volume and substantial differences in class representation also make it useful for investigating the practical challenges associated with imbalanced IoT traffic [21]. In the present study, the BoT-IoT dataset is therefore employed as the experimental benchmark for evaluating the proposed multiclass model.

Despite the substantial progress demonstrated by previous studies, several gaps remain. First, many existing studies emphasize binary detection or focus on a single attack category, limiting their ability to provide fine-grained identification of multiple threats. Second, severe class imbalance can obscure the performance of models on rare attack categories when evaluation relies primarily on aggregate accuracy. Third, IoT network traffic exhibits both local feature relationships and sequential dependencies, making a single modelling architecture potentially insufficient for capturing all relevant patterns. Finally, although hybrid architectures have increasingly been investigated, there remains a need for systematic evaluation of CNN-BiLSTM-Attention models for multiclass intrusion detection under highly imbalanced BoT-IoT traffic. The studies by Altaf *et al.* [9], Jullian *et al.* [3], Kandhro *et al.* [15], Mohammad and Bharati [16], Sayegh *et al.* [17], Guan *et al.* [19], and Omarov *et al.* [20] collectively demonstrate substantial progress in deep-learning-based IoT security while also revealing opportunities for further investigation into multiclass discrimination, imbalance handling, and hybrid representation learning.

Against this background, this study proposes a hybrid Convolutional Neural Network–Bidirectional Long Short-Term Memory with Attention (CNN-BiLSTM-Attention) model for multiclass intrusion detection in IoT environments. The proposed architecture is designed to combine local feature extraction through CNN, bidirectional temporal dependency modelling through BiLSTM, and attention-based weighting of informative learned representations. The study evaluates the proposed model using the BoT-IoT dataset and compares its performance with CNN, BiLSTM, and CNN-BiLSTM baseline architectures. Evaluation is conducted using accuracy, precision, recall, and F1-score, with particular emphasis on macro-averaged measures because of the substantial class imbalance within the dataset. Through this approach, the study seeks to provide a more fine-grained

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

assessment of IoT intrusion detection and to determine whether the integration of spatial, temporal, and attention-based representation learning can improve multiclass attack classification.

The main contributions of this study are threefold. First, it develops a hybrid CNN–BiLSTM–Attention architecture for five-class IoT intrusion detection involving Normal, DoS, DDoS, Reconnaissance, and Theft traffic. Second, it incorporates SMOTE-based class balancing and evaluates the model using macro- and weighted classification metrics to provide a more informative assessment under severe class imbalance. Third, it provides an empirical comparison of the proposed architecture with CNN, BiLSTM, and CNN–BiLSTM baseline models on the BoT-IoT dataset, highlighting both the strengths and limitations of hybrid deep-learning approaches for multiclass IoT intrusion detection.

II. METHODOLOGY

➢ *Dataset and Experimental Design*

This study employed a quantitative experimental design using the BoT-IoT dataset, which contains 2,934,817 network-flow records and 19 attributes. The target variable was formulated as a five-class intrusion-detection problem comprising Normal, DoS, DDoS, Reconnaissance, and Theft traffic. The complete dataset exhibited severe class imbalance, with DDoS and DoS accounting for 52.518% and 44.982% of the records, respectively, while Normal and Theft represented only 0.013% and 0.002%.

To construct a computationally manageable multiclass experimental dataset while retaining observations from all target categories, a subset of 60,435 records was used. The subset comprised 20,000 DDoS, 20,000 DoS, 20,000 Reconnaissance, 370 Normal, and 65 Theft observations. This composition preserved the substantial imbalance observed in the original data while ensuring that the minority classes were represented in the experiment.

The resulting dataset was divided into 48,348 training observations and 12,087 test observations using stratified sampling. SMOTE was subsequently applied only to the training data to avoid information leakage between the training and test sets. The resulting balanced training data contained 80,000 observations, with 16,000 observations representing each class. A separate validation set of 12,000 observations was used during model development.

➢ *Preprocessing and Feature Selection*

Data preprocessing involved quality assessment, categorical encoding, and Min–Max normalization. No missing values or duplicate observations were identified. SMOTE was subsequently applied exclusively to the training data, increasing each class to 16,000 observations and producing 80,000 balanced training instances.

Feature analysis identified srate, max, mean, stddev, min, seq, state_number, N_IN_Conn_P_DstIP, and N_IN_Conn_P_SrcIP as informative variables, with srate identified as particularly discriminative. The processed features were transformed into sequential representations for hybrid CNN–BiLSTM learning.

➢ *Proposed CNN–BiLSTM–Attention Model*

The proposed model integrates Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), and an Attention mechanism in a sequential architecture. The CNN layer extracts local patterns from the transformed network-flow features, after which the BiLSTM layer captures bidirectional sequential dependencies. The Attention mechanism assigns greater weights to informative learned representations before the final classification stage. A dense layer with SoftMax activation produces the probability distribution across the five intrusion classes.

The CNN operation is represented as:

$$Z = \sigma(W * X + b) \quad (1)$$

Where

- X is the input features matrix
- W is the convolution filter
- * denotes convolution
- σ denotes convolution

The BiLSTM layer processes the CNN representation in both forward and backward directions:

Operation for a given sequence:

$$h_t^{forward} = LSTM(x_t)$$

$$h_t^{backward} = LSTM(x_t)$$

$$h_t = \left[h_t^{forward}, h_t^{backward}\right]$$

Where h_t is the combined hidden state at time *t*.

The Attention mechanism is then applied to the BiLSTM representation:

$$\alpha_t = \frac{exp(e_t)}{\sum_{k=1}^{T} exp(e_k)} \quad (2)$$

Where:

- e_t represents the relevance score of time step *t*
- α_t is the normalized attention weight

The context vector is then generated as:

$$c = \sum_{t=}^{T} \alpha_t h_t \quad (3)$$

And the SoftMax function produces the probability of each of the five classes:

$$P(y = c|X) = \frac{exp(z_c)}{\sum_{k=1}^{k} exp(z_k)} \quad (4)$$

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

Where:

- *C* is the specific class (Normal, Dos, DDoS, Reconnaissance, Data Theft)
- *K* is the total number of classes

The implemented model contains approximately 76,422 trainable parameters**,** uses Adam optimization**,** categorical cross-entropy loss, SoftMax output activation, and dropout regularization. The reported validation set contained 12,000 observations.

➢ *Model Evaluation*

Performance was assessed using accuracy, precision, recall, F1-score, and confusion-matrix analysis**.** Macro-averaged metrics were emphasized because of the severe class imbalance, while weighted metrics were additionally reported. The proposed model was compared with CNN, BiLSTM, and CNN–BiLSTM baseline architectures under the same experimental setting.

III. RESULTS AND DISCUSSION

➢ *Experimental Results and Model Configuration*

The experimental evaluation was conducted using the sampled BoT-IoT data described in the methodology. The final training set contained 80,000 observations after SMOTE, while 12,000 observations were used for validation and 12,087 for testing. The implemented CNN–BiLSTM–Attention architecture used Adam optimisation, categorical cross-entropy loss, SoftMax output activation, and dropout regularisation, with approximately 76,422 trainable parameters. The best recorded validation accuracy was approximately 97.81%.

Table 1 Model Configuration

Parameter	**Configuration**
Architecture	CNN–BiLSTM–Attention
Output classes	5
Optimiser	Adam
Loss function	Categorical cross-entropy
Output activation	SoftMax
Regularisation	Dropout
Class-balancing method	SMOTE
Training instances	80,000
Validation instances	12,000
Test instances	12,087
Trainable parameters	76,422

➢ *Overall Classification Performance*

The proposed model was evaluated on 12,087 previously unseen test observations. As shown in Table 2 and Figure 1, the model achieved an overall accuracy of 96.27%, with macro precision, macro recall, and macro F1-score of 86.06%, 94.68%, and 89.64%, respectively. The weighted F1-score was 96.28%.

Table 2 Overall Performance of the Proposed Model

Metric	**Result (%)**
Accuracy	96.27
Macro Precision	86.06
Macro Recall	94.68
Macro F1-score	89.64
Weighted Precision	96.31
Weighted Recall	96.27
Weighted F1-score	96.28

The difference between the macro and weighted measures is important because the dataset remains highly imbalanced. The macro metrics give equal consideration to each class, whereas weighted metrics are strongly influenced by the substantially larger DoS and DDoS classes. Consequently, macro recall of 94.68% and macro F1-score of 89.64% provide a more informative assessment of multiclass performance than accuracy alone.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

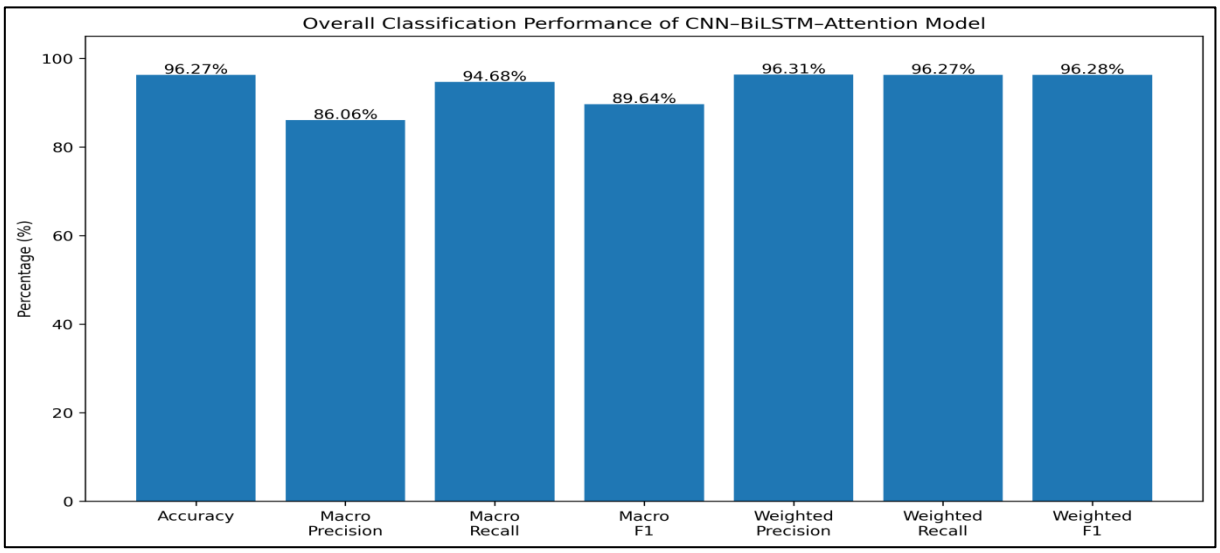

Fig 1 Overall *Classification* Performance of the CNN–BiLSTM–Attention Model

➢ *Class-Wise Classification Performance*

The class-level results are presented in Table 3. The model achieved its strongest performance on Reconnaissance, with precision of 99.20%**,** recall of 98.70%**,** and F1-score of 98.95%**.** DoS and DDoS also achieved high F1-scores of 95.05% and 95.01%**,** respectively.

Table 3 Class-Wise Classification Performance

Class	Precision (%)	Recall (%)	F1-score (%)	Support
Normal	86.05	100.00	92.50	74
DoS	94.51	95.60	95.05	4,000
DDoS	95.55	94.48	95.01	4,000
Reconnaissance	99.20	98.70	98.95	4,000
Theft	55.00	84.62	66.67	13
Macro Average	86.06	94.68	89.64	12,087

The Normal class recorded 100% recall; however, this result should be interpreted cautiously because only 74 Normal observations were present in the test set. Similarly, the Theft class produced the lowest F1-score of 66.67%. Its precision of 55.00% reflects the extremely small number of genuine Theft observations available for evaluation. The original dataset contained only 65 Theft records, making the resulting test estimate substantially less stable than those of the larger classes. The Figure 2 below show the class-level results performance F1-score *of the proposed CNN–BiLSTM–Attention model.*

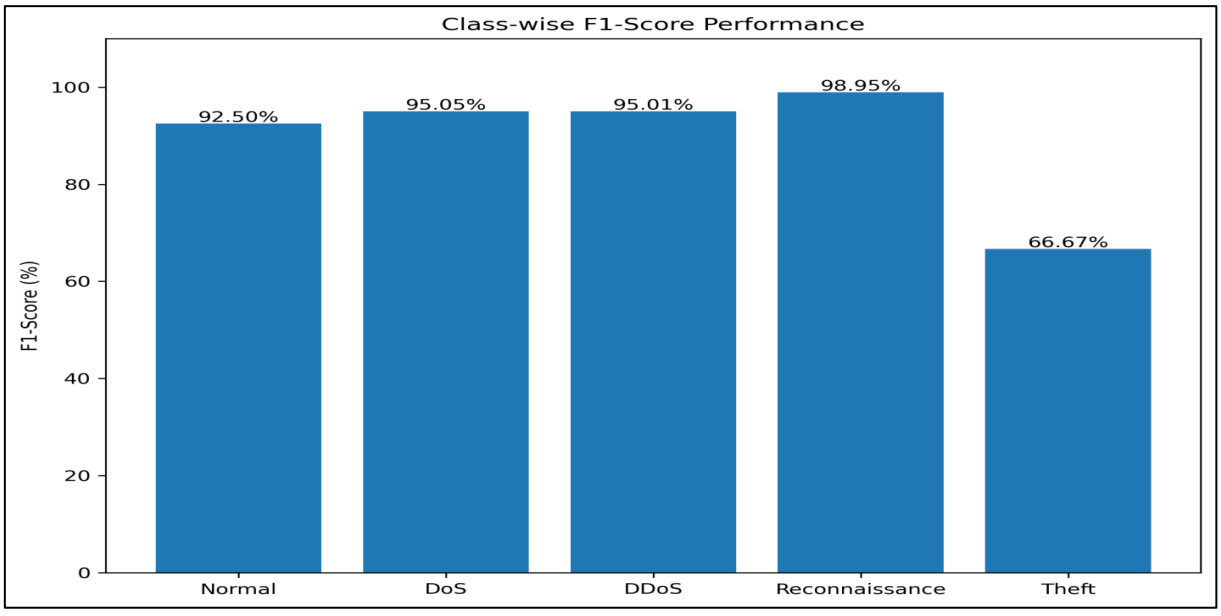

Fig 2 Class-Wise F1-Score of the Proposed CNN–BiLSTM–Attention Model.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

➢ *Confusion Matrix Analysis*

The confusion matrix in Table 4 shows that the majority of test observations were correctly classified. All 74 Normal observations were correctly identified. For DoS, 3,824 of 4,000 observations were correctly classified, while 162 were classified as DDoS. Similarly, 3,779 of 4,000 DDoS observations were correctly classified, with 202 classified as DoS.

Table 4 Confusion Matrix of the Proposed Model

Actual / Predicted	Normal	DoS	DDoS	Reconnaissance	Theft
Normal	74	0	0	0	0
DoS	1	3,824	162	13	0
DDoS	0	202	3,779	19	0
Reconnaissance	9	20	14	3,948	9
Theft	2	0	0	0	11

The reciprocal confusion between DoS and DDoS represents the principal classification error. This indicates that the two attack categories exhibit overlapping network-flow characteristics. Reconnaissance produced the strongest result, with 3,948 of 4,000 observations correctly classified. For Theft, 11 of 13 observations were correctly classified, but the small support limits the reliability of percentage-based estimates. Overall performance of the confusion matrix are shown in Figure 3 below.

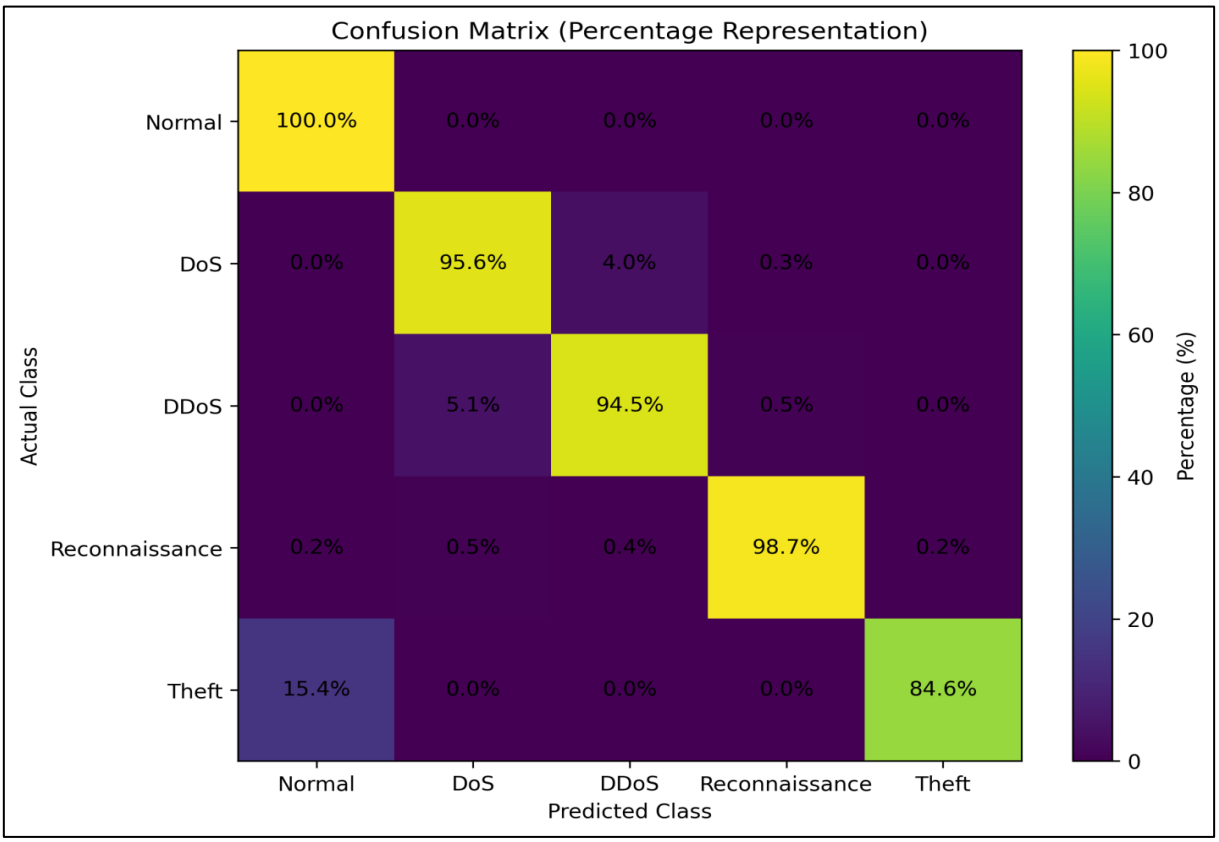

Fig 3 Confusion Matrix of the Proposed Model

➢ *Comparison with Baseline Models*

The proposed model was compared with CNN-only, BiLSTM-only, and CNN–BiLSTM architectures. The results are presented in Table 5 and Figure 4 below. The CNN–BiLSTM–Attention model achieved the highest accuracy (96.27%) and highest macro recall (94.68%) among the evaluated models.

Table 5 Comparison with Baseline Deep Learning Models

Model	Accuracy (%)	Macro Precision (%)	Macro Recall (%)	Macro F1 (%)
CNN	94.80	81.35	91.73	84.95
BiLSTM	92.41	77.82	91.81	81.96
CNN–BiLSTM	95.90	86.89	94.46	90.16
CNN–BiLSTM–Attention	96.27	86.06	94.68	89.64

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

The results indicate that combining CNN and BiLSTM improved overall accuracy compared with either individual architecture. The addition of Attention further increased accuracy and macro recall. However, CNN–BiLSTM achieved a slightly higher macro F1-score (90.16%) than the proposed model (89.64%). Therefore, the results do not support a claim that Attention is superior across all evaluation measures. Rather, the proposed model demonstrated the strongest overall accuracy and macro recall in this experiment.

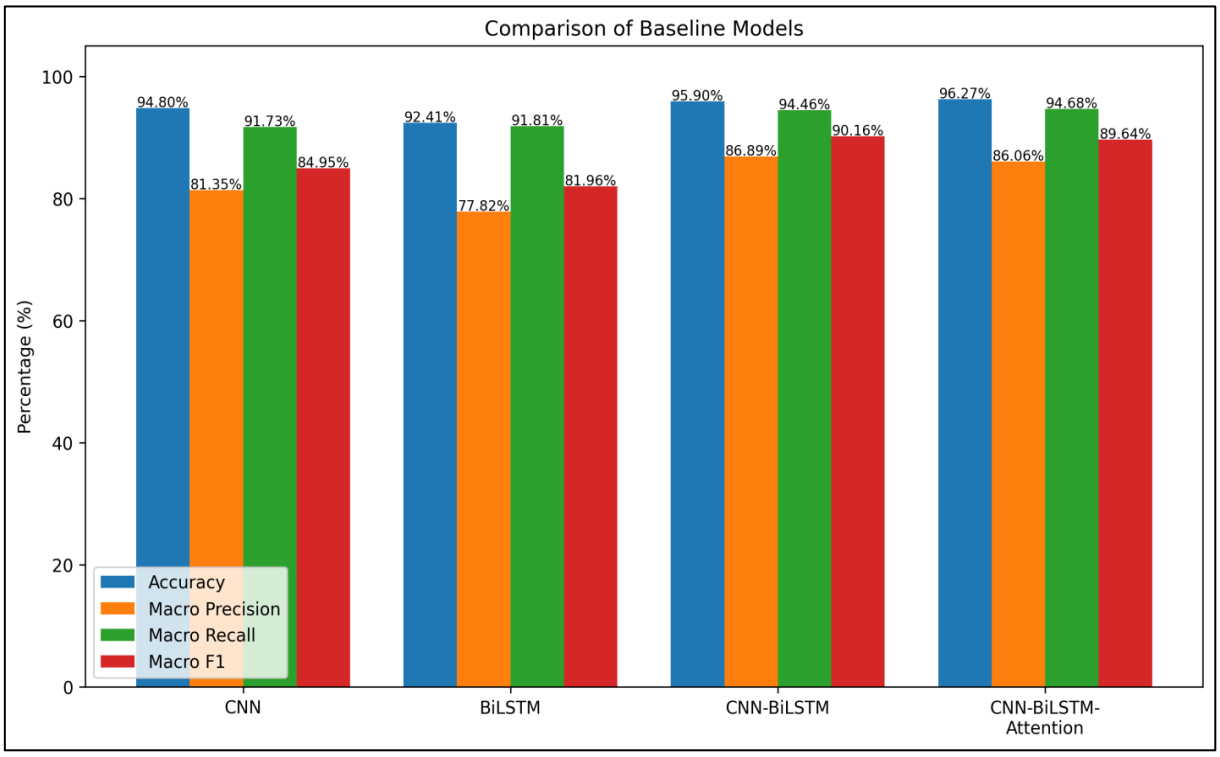

Fig 4 Caption: Comparative Performance of the Proposed Model and Baseline Architectures.

➢ *Baseline Comparison*

Table 5 compares the proposed CNN–BiLSTM–Attention model with CNN, BiLSTM, and CNN–BiLSTM baseline architectures. The CNN model achieved 94.80% accuracy and 84.95% macro F1-score, while the BiLSTM model recorded 92.41% accuracy and 81.96% macro F1-score. Combining CNN and BiLSTM increased accuracy to 95.90% and macro recall to 94.46%.

The proposed CNN–BiLSTM–Attention model achieved the highest accuracy (96.27%) and macro recall (94.68%) among the evaluated architectures. Its macro precision was 86.06%, while its macro F1-score was 89.64%. Although the addition of Attention increased accuracy and macro recall relative to the CNN–BiLSTM model, the CNN–BiLSTM baseline obtained a slightly higher macro F1-score (90.16%).

These results indicate that the contribution of the Attention mechanism is metric-dependent. In this experiment, Attention was associated with improved overall accuracy and macro recall but did not produce the highest macro F1-score. Therefore, the results support the effectiveness of the complete hybrid architecture without implying that Attention universally improves every evaluation measure.

➢ *Discussion of Findings*

The results demonstrate that the proposed CNN–BiLSTM–Attention architecture is capable of performing multiclass intrusion detection on the BoT-IoT dataset. The 96.27% test accuracy and 94.68% macro recall indicate that the model was able to correctly identify a large proportion of observations across the five target classes. The results also show that combining convolutional and recurrent representations provides a useful approach for modelling heterogeneous IoT network traffic.

The strong performance of the Reconnaissance class, which achieved an F1-score of 98.95%, suggests that its learned traffic representation was sufficiently distinguishable from the other evaluated classes. DoS and DDoS also achieved comparable F1-scores of 95.05% and 95.01%, respectively. However, the confusion matrix shows reciprocal misclassification between these two classes, indicating that their traffic characteristics can overlap and remain difficult to distinguish completely.

The results for the minority classes require more cautious interpretation. Although Normal traffic achieved 100% recall, only 74 Normal observations were available in the test set. More importantly, the Theft class achieved an F1-score of 66.67% from only 13 test observations. SMOTE increased the representation of minority classes during training, but synthetic samples cannot fully substitute for the diversity of genuine attack observations. Consequently, the

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

performance of the Theft class should not be interpreted as evidence of poor model capability alone; it also reflects the severe scarcity of genuine examples.

The baseline comparison provides further insight into the architecture. CNN–BiLSTM improved performance over the individual CNN and BiLSTM models, suggesting that combining local feature extraction with bidirectional sequential representation was beneficial. The addition of Attention produced the highest overall accuracy and macro recall, although CNN–BiLSTM retained a slightly higher macro F1-score. This indicates that the hybrid architecture provides competitive multiclass performance, while the precise contribution of Attention may depend on the evaluation metric and class distribution.

Overall, the findings support the feasibility of hybrid deep learning for multiclass IoT intrusion detection while highlighting two important challenges: distinguishing closely related attack categories such as DoS and DDoS and reliably detecting extremely rare attack categories such as Theft. These limitations provide clear directions for further investigation using larger and more diverse IoT intrusion datasets.

IV. CONCLUSION

This study developed and evaluated a hybrid CNN–BiLSTM–Attention model for multiclass intrusion detection in IoT environments using the BoT-IoT dataset. The model achieved 96.27% accuracy, 94.68% macro recall, and 89.64% macro F1-score. Reconnaissance achieved the strongest class-level performance with an F1-score of 98.95%, whereas Theft remained the most challenging class because of its extremely limited representation. Overall, the findings demonstrate the potential of combining convolutional feature extraction, bidirectional sequential learning, and attention-based representation weighting for multiclass IoT intrusion detection.

RECOMMENDATIONS AND FUTURE WORK

Future studies should use larger and more diverse IoT datasets, improve the detection of rare attack classes, and investigate features that better distinguish DoS from DDoS traffic. Further research should also examine model explainability, computational efficiency, and real-time deployment in resource-constrained IoT environments.

REFERENCES

[1]. R. Chataut, A. Phoummalayvane, and R. Akl, “Unleashing the power of IoT: A comprehensive review of IoT applications and future prospects in healthcare, agriculture, smart homes, smart cities, and Industry 4.0,” *Sensors*, vol. 23, no. 16, p. 7194, 2023, doi: 10.3390/s23167194.

[2]. A. Djenna, S. Harous, and D. E. Saidouni, “Internet of Things meet Internet of threats: New concern cyber security issues of critical cyber infrastructure,” *Applied Sciences*, vol. 11, no. 10, p. 4580, 2021, doi: 10.3390/app11104580.

[3]. O. Jullian, B. Otero, E. Rodriguez, N. Gutierrez, H. Antona, and R. Canal, “Deep-learning based detection for cyber-attacks in IoT networks: A distributed attack detection framework,” *Journal of Network and Systems Management*, vol. 31, Art. no. 33, 2023, doi: 10.1007/s10922-023-09722-7.

[4]. R. Atassi, “Anomaly detection in IoT networks: Machine learning approaches for intrusion detection,” *Fusion: Practice and Applications*, vol. 13, no. 1, pp. 126–134, 2023, doi: 10.54216/FPA.130110.

[5]. G. Logeswari, J. D. Roselind, K. Tamilarasi, and V. Nivethitha, “A comprehensive approach to intrusion detection in IoT environments using hybrid feature selection and multi-stage classification techniques,” *IEEE Access*, vol. 13, pp. 24970–24987, 2025, doi: 10.1109/ACCESS.2025.3532895.

[6]. P. Sinha, D. Sahu, S. Prakash, T. Yang, R. S. Rathore, and V. K. Pandey, “A high performance hybrid LSTM CNN secure architecture for IoT environments using deep learning,” *Scientific Reports*, vol. 15, Art. no. 9684, 2025, doi: 10.1038/s41598-025-94500-5.

[7]. A. K. Sahu, S. Sharma, M. Tanveer, and R. Raja, “Internet of Things attack detection using hybrid deep learning model,” *Computer Communications*, vol. 176, pp. 146–154, 2021, doi: 10.1016/j.comcom.2021.05.024.

[8]. M. Zhong, Y. Zhou, and G. Chen, “Sequential model based intrusion detection system for IoT servers using deep learning methods,” *Sensors*, vol. 21, no. 4, p. 1113, 2021, doi: 10.3390/s21041113.

[9]. T. Altaf, X. Wang, W. Ni, G. Yu, R. P. Liu, and R. Braun, “GNN-based network traffic analysis for the detection of sequential attacks in IoT,” *Electronics*, vol. 13, no. 12, p. 2274, 2024, doi: 10.3390/electronics13122274.

[10]. E. Gelenbe and M. Nakip, “Traffic based sequential learning during botnet attacks to identify compromised IoT devices,” *IEEE Access*, vol. 10, pp. 126536–126549, 2022, doi: 10.1109/ACCESS.2022.3226700.

[11]. M. Ramzan, M. Shoaib, A. Altaf, S. Arshad, F. Iqbal, Á. K. Castilla, and I. Ashraf, “Distributed denial of service attack detection in network traffic using deep learning algorithm,” *Sensors*, vol. 23, no. 20, p. 8642, 2023, doi: 10.3390/s23208642.

[12]. F. Alasmary, S. Alraddadi, S. Al-Ahmadi, and J. Al-Muhtadi, “ShieldRNN: A distributed flow-based DDoS detection solution for IoT using sequence majority voting,” *IEEE Access*, vol. 10, pp. 88263–88275, 2022, doi: 10.1109/ACCESS.2022.3200477.

[13]. K. Sundar, A. Neyaz, and Q. Liu, “IoT network attack detection using supervised machine learning,” *International Journal of Artificial Intelligence and Expert Systems*, vol. 10, no. 2, pp. 18–32, 2021.

[14]. O. Salman, I. H. Elhajj, A. Chehab, and A. Kayssi, “A machine learning based framework for IoT device identification and abnormal traffic detection,” *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 3, Art. no. e3743, 2022, doi: 10.1002/ett.3743.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26sep806

[15]. I. A. Kandhro, S. M. Alanazi, F. Ali, A. Kehar, K. Fatima, M. Uddin, and S. Karuppayah, "Detection of real-time malicious intrusions and attacks in IoT empowered cybersecurity infrastructures," *IEEE Access*, vol. 11, pp. 9136–9148, 2023, doi: 10.1109/ACCESS.2023.3238664.

[16]. Z. Mohammad and T. S. Bharati, "Enhancing cybersecurity in IoT systems: A hybrid deep learning approach for real-time attack detection," *Discover Internet of Things*, vol. 5, Art. no. 73, 2025, doi: 10.1007/s43926-025-00156-y.

[17]. H. R. Sayegh, W. Dong, and A. M. Al-madani, "Enhanced intrusion detection with LSTM-based model, feature selection, and SMOTE for imbalanced data," *Applied Sciences*, vol. 14, no. 2, p. 479, 2024, doi: 10.3390/app14020479.

[18]. A. O. Widodo, B. Setiawan, and R. Indraswari, "Machine learning-based intrusion detection on multi-class imbalanced dataset using SMOTE," *Procedia Computer Science*, vol. 234, pp. 578–583, 2024, doi: 10.1016/j.procs.2024.03.042.

[19]. Y. Guan, M. Noferesti, and N. Ezzati-Jivan, "A two-tiered framework for anomaly classification in IoT networks utilizing CNN-BiLSTM model," *Software Impacts*, vol. 20, p. 100646, 2024, doi: 10.1016/j.simpa.2024.100646.

[20]. B. Omarov, Z. Sailaukyzy, A. Bigaliyeva, A. Kereyev, L. Naizabayeva, and A. Dautbayeva, "One dimensional Conv-BiLSTM network with attention mechanism for IoT intrusion detection," *Computers, Materials & Continua*, vol. 77, no. 3, pp. 3765–3781, 2023, doi: 10.32604/cmc.2023.042469.

[21]. N. Koroniotis, N. Moustafa, E. Sitnikova, and B. Turnbull, "Towards the development of realistic botnet dataset in the Internet of Things for network forensic analytics: Bot-IoT dataset," *Future Generation Computer Systems*, vol. 100, pp. 779–796, 2019, doi: 10.1016/j.future.2019.05.041.