INTERNATIONAL JOURNAL OF Telecommunications
AIR DEFENSE COLLEGE

Robust BiLSTM-Based Multi-Class Intrusion Detection for IoT Networks Using ToN-IoT Dataset

Citation: Adel, M.; Adel, A.; abo-Taleb, A.; Mohamed, M.
Inter. Jour. of Telecommunications, ***IJT'2025****, Vol. 05, Issue 02, pp. 1-14, 2025.*
Doi: 10.21608/ijt.2025.426708.1134
Editor-in-Chief: Youssef Fayed.
Received: 10/10/2025.
Accepted date: 20/11/2025.
Published date: 20/11/2025.

Mohammed Adel [*1], Ahmed M. C [2], Ahmad M.abo-Taleb [1] And M.A. Mohamed [1]
1 Department of Electronics and Communications Engineering, Faculty of Engineering, Mansoura University
2 Department of Computers & Artificial Intelligence, Military Technical College
Emails: mohammedadelali@std.mans.edu.eg , a.mattar@ieee.org , aptlpp@gmail.com , Mazim12@gmail.com
Corresponding author(s): Mohammed Adel (mohammedadelali@std.mans.edu.eg)

Abstract: The Internet of Things (IoT) is reshaping smart cities, healthcare, and industrial automation, but its devices' insufficient computing power and weak security make them main targets for advanced cyberattacks, such as DDoS and zero-day exploits. Conventional Intrusion Detection Systems (IDSs) often fail to detect evolving threats, necessitating smarter, deep learning-based solutions. We devise a Bidirectional Long Short-Term Memory (BiLSTM) based IDS trained on the ToN-IoT dataset's network traffic. Our preprocessing pipeline featuring label encoding, feature hashing, SMOTE for class balancing, and chi-squared feature selection creates a robust, compact feature space. With class weights to address residual imbalance, the model achieves 99.94% accuracy, with high precision, recall, and F1-scores across nine attack types. Its performance remains stable across reduced feature sets, preserving adaptability. Outperforming state-of-the-art methods, our IDS incorporate exceptional detection accuracy with low inference latency, enabling real-time deployment in resource-constrained IoT environments. This framework safely strengthens security for smart cities, critical infrastructure, and edge computing applications.

Keywords: Internet of Things (IoT); Intrusion Detection System (IDS); BiLSTM; ToN-IoT Dataset; Multi-class Classification; Cybersecurity

1. Introduction

The rapid development of the Internet of Things (IoT) has transformed areas such as smart urban environments, healthcare, and Industrial IoT Operations. By 2025, more than 75 billion IoT devices are expected to find their place inside global networks, extending the landscape where threats like Distributed Denial of Service (DDoS), botnets, and ransomware can emerge. Such attacks leave deeply weak IoT devices insecured, as many of them lack the ability to defend themselves, making dependable protection an urgent necessity. Intrusion Detection Systems (IDS) have come to stand at the front line of IoT security, yet conventional techniques often falter when facing sophisticated or unfamiliar attacks in this constantly shifting and varied ecosystem. To address this, deep

learning (DL) approaches have gained traction for their ability to extract complex patterns from raw network traffic, delivering robust detection performance [1]. Lightweight BiLSTM models have also been proposed to balance detection effectiveness with the limited computational resources of IoT devices [2]. Surveys further highlight the rise of adaptive IDS designs, such as those leveraging reinforcement learning, though these often face challenges like high computational costs and poor generalization to novel threats [3]. Despite this progress, many IDS models are restricted to binary classification, only distinguishing between benign and malicious traffic. This approach is devoid of the granularity required to identify specific attack types, which is required for deploying targeted countermeasures, such as mitigating ransomware versus botnet propagation [4]. also, some studies fail to spot critical preprocessing steps, including class balancing and robust feature encoding, which are needed for reliable functional proficiency in real-world scenarios. This research introduces a deep learning-based IDS framework for multi-class classification of IoT attacks. Our scheme deploys a Bidirectional Long Short-Term Memory (BiLSTM) model enhanced by means of Batch Normalization, integrated with a comprehensive preprocessing pipeline that includes label encoding, feature hashing, and SMOTE to address class imbalance. Trained and evaluated on the network traffic subset of the ToN-IoT dataset, our model demonstrates superior classification precision across diverse attack categories while maintaining efficiency for resource-constrained IoT environments.

The distinctive contributions of this approach can be highlighted as:

1. A BiLSTM-based intrusion detection framework designed for IoT environments, enhanced with Batch Normalization to ensure stable and efficient training.
2. A comprehensive preprocessing pipeline that integrates label encoding, feature hashing, SMOTE oversampling, and chi-squared feature selection to construct a compact yet discriminative feature space.
3. Robust handling of class imbalance by combining SMOTE with class weighting, enabling balanced detection performance across nine different attack categories.
4. Superior multi-class classification performance, achieving 99.94% overall accuracy with consistently high precision, recall, and F1-scores, even when tested with reduced feature sets.
5. Real-time readiness and scalability, demonstrating low inference latency that makes the proposed IDS intended for resource-constrained IoT devices, smart cities, and edge computing infrastructures.

The remainder of this study is further organized as delineated below:
Section 2 investigates related work on intrusion detection systems for IoT networks. Section 3 presents the introduced model and its framework. Section 4 describe the experimental results, while Section 5 delivers an elaborate discussion .As a final point, Section 6 encapsulates the study and points to prospective research avenues.

2. Related Work

The expansion of Internet of Things (IoT) systems has led to an upsurge in security threats, prompting extensive research into intelligent intrusion detection systems (IDS). Existing works have studied heterogeneous machine learning (ML) and deep learning (DL) paradigms to handle the complexity of multi-class classification in IoT-based cyber threats. This section categorizes recent studies using the ToN-IoT dataset into four thematic groups: deep learning models, traditional ML with feature selection, ensemble frameworks, and federated learning.

2.1 Deep Learning-Based Intrusion Detection Models

Explainable deep learning is increasingly prioritized in intrusion detection to overcome the opacity of neural models. A combination of CNN and recurrent layers was trained on ToN-IoT with SHAP interpretation, achieving 97.8% accuracy using only 15 features [5]. A vision-based IDS was also proposed, where PCAP data were converted into images; their custom CNN reached 99.1% in binary and 89.9% in multi-class settings, while ResNet50 and VGG16 also exceeded 94% accuracy [6]. In another study, InceptionTime achieved 100% accuracy on the Win10–Network subset, outperforming conventional LSTM and RF models [7]. Finally, a deep IDS for industrial IoT was introduced with sine-cosine encoding and SMOTE+Tomek balancing, achieving 98.87% accuracy and 99.18% F1-score [8].

2.2 Traditional Machine Learning with Feature Selection Techniques

Traditional ML approaches, when paired with advanced feature engineering, have demonstrated competitive performance in intrusion detection. Statistical filters and evolutionary optimization enabled SVM to reach 99.48% accuracy on ToN-IoT [9]. LightGBM was also applied for multiclass classification, achieving 78% accuracy and 0.82 precision, surpassing logistic regression, SVM, and naïve Bayes [10]. These results emphasize the role of dimensionality reduction and algorithmic tuning in constrained environments.

2.3 Ensemble and Hybrid Learning-Based Intrusion Detection Frameworks

Ensemble learning has proven effective in boosting IDS performance by combining diverse models. Ensemble methods such as RF, XGBoost, and CatBoost were applied to ToN-IoT, achieving up to 98.66% accuracy and 98.61% F1-score [11]. Earlier attempts using standalone classifiers or simple voting ensembles struggled to exceed 86%, underscoring the importance of stacking and model synergy.

2.4 Federated and Distributed Learning-Based IDS Approaches

With privacy becoming a critical concern, federated learning (FL) has gained traction in IDS development. A distributed IDS evaluated on multiple ToN-IoT subsets (network, Windows, Linux) used SMOTE and Chi^2 selection, where XGBoost achieved 99.1% binary and 98.3% multi-class accuracy, reaching 100% on the Windows 10 subset [12]. A centralized DBN reached 94% F1-score, while a FedProx-based FL-DBN achieved 72%, outperforming other FL baselines by over 30% [13]. Finally, a structure-preserving FL approach using Grassmann manifolds was introduced, with the FEDPG framework reaching 93.84% F1-score and 96.97% AUC on ToN-IoT, outperforming deep baselines without compromising data privacy [14].

Table 1. provides a clear comparison of recent intrusion detection approaches utilizing the ToN-IoT dataset. A notable observation is that the majority of existing studies employed binary classification, which simplifies the detection task to merely distinguishing between "attack" and "normal" traffic. While such an approach can yield high accuracy, it fails to capture the real-world complexity of IoT environments. Only a limited number of works—such as [6], [7], [10], and [12]—explored multiclass classification, which poses a greater challenge due to the need to distinguish between various attack types. Among these, the highest multiclass accuracy reported was 98.3% [12], while others remained below 90% [6][10]. Although [7] reported 100% accuracy using InceptionTime, the experiment was conducted on a narrow device-specific subset of the dataset, rather than the more dynamic and diverse network traffic subset. In contrast, the proposed model in this study operates on subset of network traffic data and achieves an outstanding 99.94% multiclass accuracy, surpassing all previous

works. This result underscores the effectiveness and generalizability of the proposed method in accurately detecting a wide range of cyber threats within complex IoT ecosystems.

Table 1. Summary of Related Works in Cybersecurity and IoT

Ref	Technique / Model	Dataset	Classification Type	Best Result
[5]	CNN + RNN + SHAP (15 features)	ToN-IoT	Binary	Accuracy: 97.8%
[6]	Custom CNN, ResNet50, VGG16 (vision-based)	ToN-IoT	Binary & Multiclass	Binary: 99.1%, Multiclass: 89.9%
[7]	InceptionTime	ToN-IoT	Multiclass	Accuracy: 100%
[8]	Deep NN + Sine-Cosine encoding + SMOTE + Tomek	ToN-IoT	Binary	Accuracy: 98.87%, F1-score: 99.18%
[9]	SVM + Statistical Filters + Evolutionary Optimization	ToN-IoT	Binary	Accuracy: 99.48%
[10]	LightGBM vs LR, SVM, NB	ToN-IoT	Multiclass	Accuracy: 78%, Precision: 0.82
[11]	RF, XGBoost, CatBoost (stacked ensemble)	ToN-IoT	Binary	Accuracy: 98.66%, F1-score: 98.61%
[12]	XGBoost + SMOTE + Chi^2	ToN-IoT	Binary & Multiclass	Binary: 99.1%, Multiclass: 98.3%, Win10: 100%
[13]	Centralized DBN, FedProx-based FL-DBN	ToN-IoT	Binary	F1-score: Centralized: 94%, FedProx: 72%
[14]	FEDPG (Grassmann manifold)	ToN-IoT	Binary	F1-score: 93.84%, AUC: 96.97%

3. Methodology

3.1 ToN-IoT Dataset Benchmark

The ToN-IoT dataset is a recent benchmark designed to rate cybersecurity solutions, especially Intrusion Detection Systems (IDS), for Internet of Things (IoT) and Industrial IoT (IIoT) environments. Developed by researchers at UNSW Canberra Cyber, it contains network traffic from a medium-scale testbed that reflect the complexity of today's IoT ecosystems. The dataset combines diverse data sources, including telemetry from IoT/IIoT sensors (such as weather and Modbus sensors) and operating system logs from Windows 7 and Windows 10, providing a comprehensive sight of IoT network behavior. The dataset covers both benign and malicious traffic, featuring nine distinct attack types, such as Distributed Denial of Service (DDoS), ransomware,

and backdoors, making it perfect for testing multi-class classification models. These attacks were carried out across a sundry testbed of virtual machines, physical devices, IoT gateways, and cloud/fog platforms, ensuring a broad range of threat scenarios. This diversity is very well-suited for deep learning models like our proposed Bidirectional Long Short-Term Memory (BiLSTM) framework, thanks to its detailed attack category labels, which allow precise threat detection. A key strength of ToN-IoT is its ability to address IoT security research challenges, such as data imbalance and complex feature spaces. Compared to other benchmarks like BoT-IoT or CICIoT2023, which may focus on specific attack types, ToN-IoT's varied attack scenarios provide a stronger foundation for evaluating advanced IDS models.

3.2 Data Preprocessing and Model Training

In this study, we worked with a reduced version of the ToN-IoT network traffic dataset rather than the full-scale dataset. The subset was constructed to maintain a manageable size for deep learning experiments while still being sufficiently large and diverse to train a robust model. Importantly, the reduced dataset was created while protecting the original proportional distribution of attack categories and normal traffic, ensuring that the statistical characteristics remained consistent with the full dataset. In total, the dataset used for training and evaluation comprised 1,259,340 samples for training and 314,835 samples for testing.

During the merging process of the multiple raw CSV files corresponding to different attack types and normal activity, the samples were selected in a non-sequential manner to enhance diversity and avoid ordering bias. This approach ensured broad coverage of all classes—including rare attacks—while reducing computational overhead. As the dataset remained imbalanced, particularly for minority attack classes, the Synthetic Minority Oversampling Technique (SMOTE) was utilized solely on the training set to equalize class distributions. The test set was preserved in its original imbalanced state to ensure that the evaluation reflects realistic deployment conditions and the model's ability to handle naturally skewed data. This method synthetically generated new samples for underrepresented classes, yielding a balanced training distribution and improving the model's capability to pinpoint less frequent attacks.

Feature preprocessing included encoding categorical variables (protocol types, service names, connection states) using label encoding, and transforming high-cardinality features such as IP addresses via feature hashing to reduce dimensionality. Boolean features were mapped to binary values (0 or 1), while missing values were imputed with zeros. Numerical attributes (e.g., packet sizes, transaction depth) were normalized to the [0,1] range using Min-Max scaling to ensure uniform feature contribution. Finally, the Chi-Squared test was employed to isolate the top 10 highly contributive features, reducing dimensionality, mitigating overfitting, and improving training efficiency. Figure 1 represents the end-to-end data preprocessing and training pipeline adopted in this study.

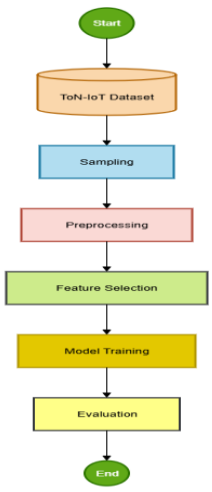

Figure 1. End-to-end pipeline for data preparation and BiLSTM model training on the ToN-IoT dataset.

processed data was then reshaped to fit the sequential input requirements of the BiLSTM model. The architecture includes two stacked Bidirectional Long Short-Term Memory layers with 128 and 64 units, respectively. Each layer is followed by Batch Normalization to stabilize and accelerate convergence, and Dropout (rate = 0.3) to reduce overfitting. A fully connected output layer with SoftMax activation produces probabilities across ten categories (nine attack types and normal traffic). Category weights were also computed and applied during training to handle any residual imbalance and ensure that minority classes contributed proportionately to the cost function. The BiLSTM model was trained for 20 epochs with a batch size of 128 utilizing the Adam optimizer (learning rate = 0.0003) and sparse categorical cross-entropy designated as the loss function. An early stopping technique featuring a patience of five epochs prevented overfitting. The test set was used to evaluate proficiency using accuracy, precision, recall, F1-score, and confusion matrix, providing insights into the detection capability for each class separately. To clarify how sequence modeling was applied in the proposed architecture, each network flow in the ToN-IoT dataset was originally a flat feature vector. For sequence modeling, we reshaped it into a short temporal sequence with one time step and multiple features. Although the dataset doesn't include true temporal links between flows, this setup lets the BiLSTM learn relationships and dependencies among the features within each flow, helping it better capture complex patterns in IoT traffic.

Experimental Infrastructure**:**
The experiments were executed using Google Colab, a cloud-based platform with GPU acceleration. The implementation was executed in Python 3.12 with TensorFlow 2.19 and supporting libraries.

3.3 Evaluation Metrics

To check out the performance of our BiLSTM-based IDS on the ToN-IoT dataset, to assess our model's performance, we used evaluation metrics specifically designed for multi-class classification in IoT security. We used accuracy, precision, recall, and F1-score to measure the model's effectiveness across diverse attack types, such as DDoS, ransomware, and backdoors.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$
$$\text{Precision} = TP / (TP + FP)$$
$$\text{Recall} = TP / (TP + FN)$$
$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

A confusion matrix was generated to analyze per-class performance, providing an assessment of the model's proficiency to differentiate between normal traffic and specific attack categories. Class weights, computed during training, supported balanced evaluation by prioritizing underrepresented attack classes.

4. Results

To investigate the capability of the introduced BiLSTM-based Intrusion Detection System (IDS), we experimented on a reduced yet large-scale version of the ToN-IoT dataset. This subset was carefully constructed to maximize the number of records while preserving the diversity of all traffic categories. After stratified sampling, the dataset contained 1,259,340 training records and 314,835 testing records, each with 44 features. To address the strong class imbalance, SMOTE was applied on the training set, balancing all ten classes to 360,000 records each.

4.1 Class Distribution

Before balancing, the dataset was highly skewed, with scanning, DDoS, and DoS dominating the traffic, while attacks such as ransomware and MITM appeared rarely. A pie chart (Figure 2) illustrates the distribution of classes in terms of counts and percentages, while Figure 3 contrasts the dataset composition before and after applying SMOTE. To facilitate efficient training while preserving proportional representation of all classes, we constructed a reduced version of the ToN-IoT dataset using stratified sampling. As shown in Figures 2 and 3, this subset reflects the original class imbalance while also demonstrating the effect of SMOTE in producing a balanced training distribution.

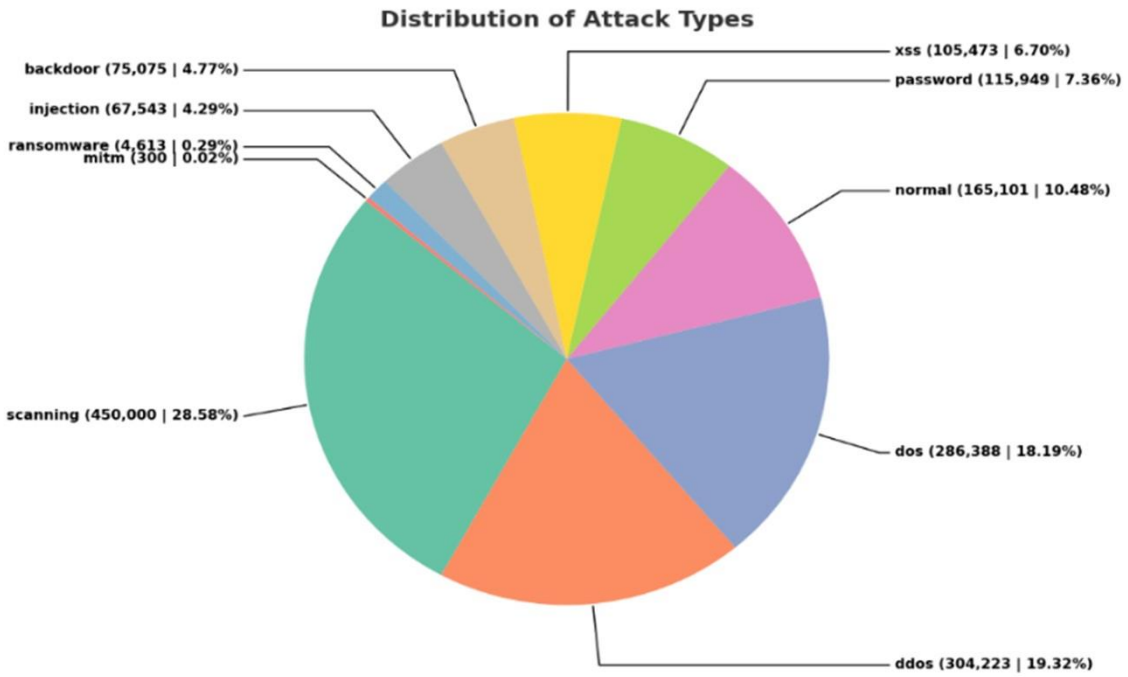

Figure 2. Distribution of attack types

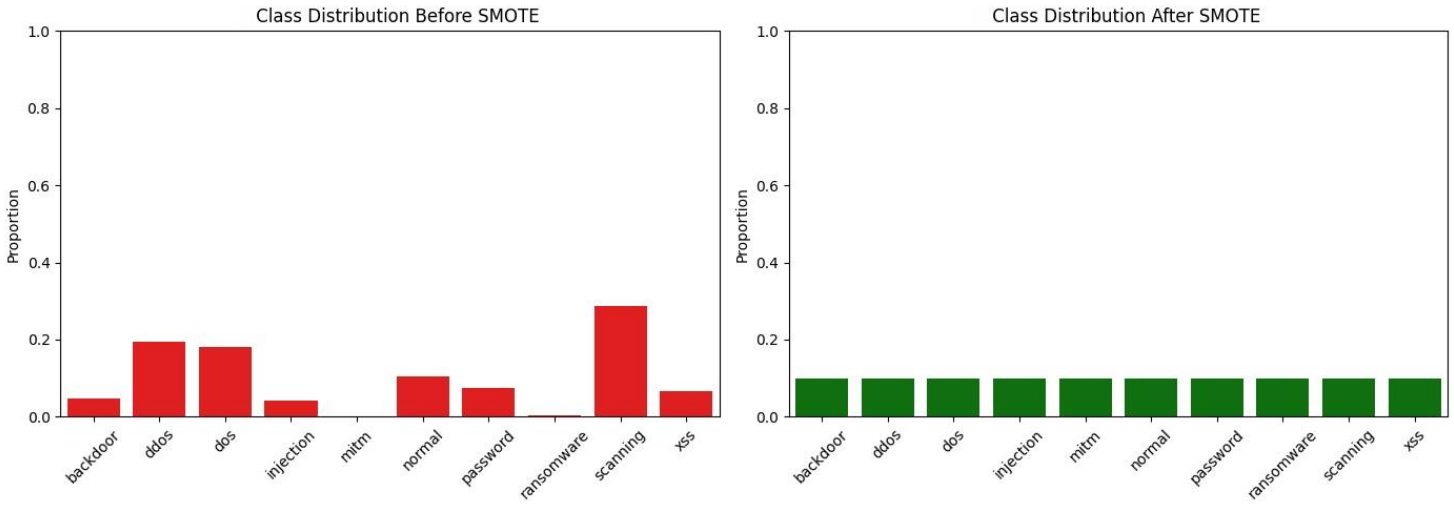

Figure 3. Before and after applying SMOTE to balance the dataset classes

4.2 Overall Performance

We next evaluate the overall classification performance of the proposed BiLSTM model on the test set. The proposed BiLSTM achieved 99.94% test accuracy on the 314,835 unseen samples, showing outstanding ability to discriminate normal traffic from nine different attack categories. Precision, Recall, and F1-score reached values close to 1.00 for both macro and weighted averages, manifest highly consistent performance across all classes. Table 2 summarizes the main evaluation metrics.

Table 2. Summary of Performance metrics of the IDS model across different attack classes in the ToN-IoT dataset.

Class	Precision	Recall	F1-score	Support
Backdoor	1.00	1.00	1.00	15,015
DDoS	1.00	1.00	1.00	60,844
DoS	1.00	1.00	1.00	57,187
Injection	1.00	1.00	1.00	13,509
MITM	0.98	1.00	0.99	60
Normal	1.00	1.00	1.00	33,014
Password	1.00	1.00	1.00	23,190
Ransomware	1.00	1.00	1.00	921
Scanning	1.00	1.00	1.00	90,000
XSS	1.00	1.00	1.00	21,095

Overall Accuracy			99.94%	314,835
Macro Avg	1.00	1.00	1.00	314,835
Weighted Avg	1.00	1.00	1.00	314,835

To assess the role of SMOTE, we retrained the model using class weights only. This resulted in 99.73% accuracy but revealed misclassifications in rare attack classes, such as Injection (often confused with DoS/DDoS) and Ransomware (a few instances mislabeled as Normal). With SMOTE, we achieve perfect classification and 99.94% accuracy, eliminating all residual errors. SMOTE is essential for robust minority class detection in high-stakes intrusion detection; class weights alone are insufficient — though viable in real-time or resource-limited settings.

4.3 Per-Class Evaluation

Beyond aggregate accuracy, it is crucial to examine class-wise detection capability, which we present using the confusion matrix. The confusion matrix (Figure 4) highlights the model's robust generalization, with almost all records assigned to their correct categories. Notably, even the minority classes (e.g., MITM with 300 samples, ransomware with 4,613 samples) achieved nearly perfect classification, thanks to the class-balancing strategy. Larger classes such as scanning and DoS also showed minimal misclassifications, confirming the scalability of the approach.

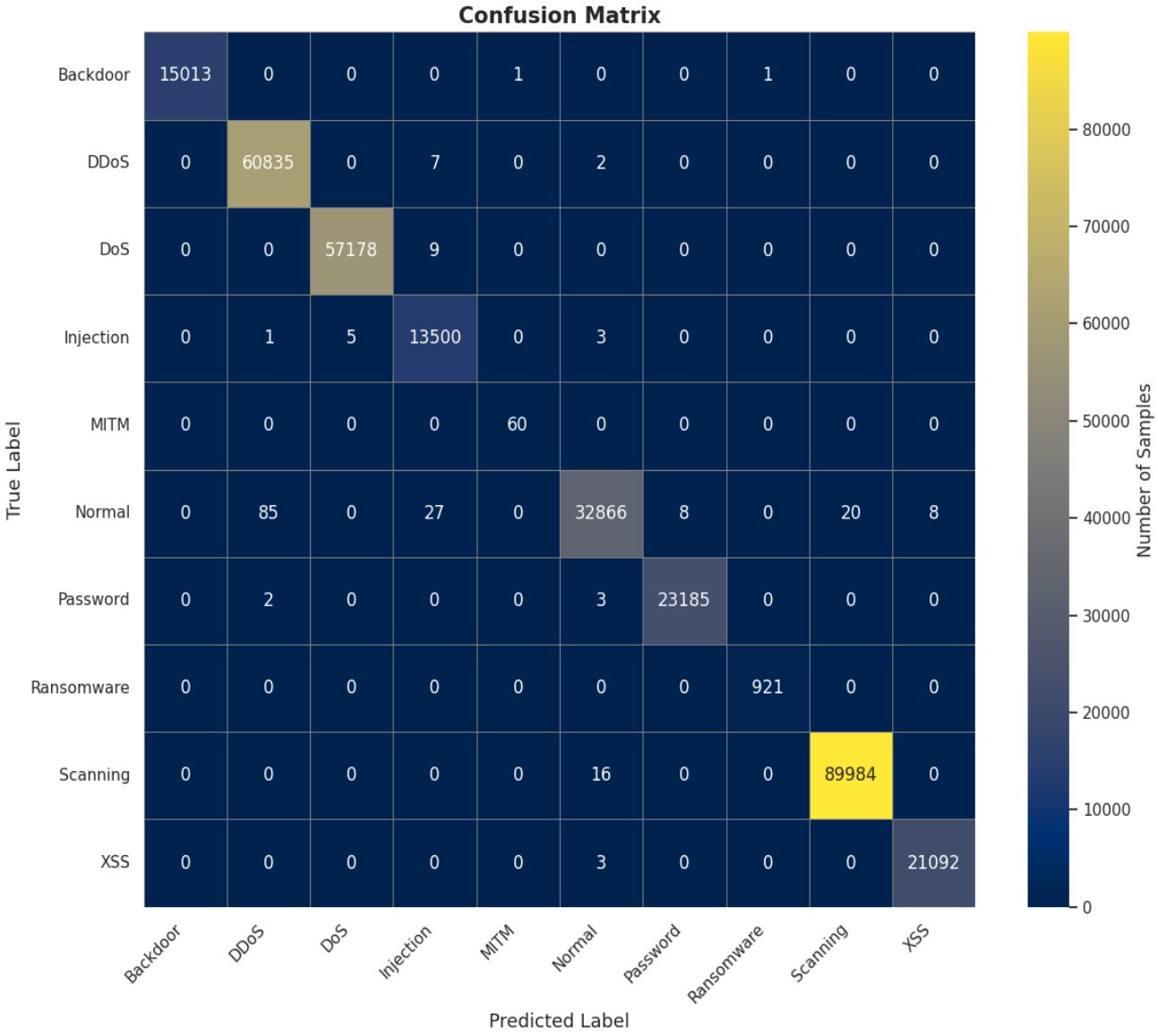

Figure 4. Confusion matrix depicting the distribution of true and false predictions across classes in the ToN-IoT dataset.

4.4 Feature Reduction Experiments

To further assess the model's adaptability to constrained environments, we conducted experiments with reduced feature sets. To test the model under resource constraints, we conducted experiments using 10, 12, and 14 features selected via Chi-square ranking. The BiLSTM maintained high accuracy in all cases, with the best result (99.94%) obtained with 10 features. This shows that the model can operate effectively even with a compact feature space, reducing computational overhead without sacrificing detection capability. Table 3 presents the accuracy obtained under different feature sets.

Table 3. Accuracy of the proposed model under different feature sets

Feature Set	Number of Features	Test Accuracy
Set 1	10	99.94%
Set 2	12	99.91%
Set 3	14	99.93%

Table 4 presents the top 10 features identified by the Chi-square test as the most relevant for distinguishing attack and benign traffic. These features include protocol- and flow-level attributes such as dns_query, dst_port, conn_state, service, and src_port, which capture critical aspects of communication patterns and service interactions in IoT networks. The prominence of DNS-related and port-based features suggests that abnormal connection states and irregular service requests are strong indicators of intrusion attempts. Highlighting these features enhances the interpretability and transparency of the proposed BiLSTM-based IDS.

Table 4. Top 10 Features Selected by Chi-Squared Test

Rank	Feature	Chi-Squared Score
1	dns_query	681,202.2325
2	dst_port	628,826.0225
3	conn_state	563,284.8708
4	service	541,997.6100
5	src_port	220,629.1367
6	src_ip	126,833.2974
7	dns_rcode	90,594.5328
8	dst_ip	68,125.7615
9	ts	51,607.9986
10	proto	48,545.2273

4.5 Training and Convergence

Finally, we analyze the training behavior to verify model stability and generalization. The training dynamics further demonstrate the stability and reliability of the proposed BiLSTM approach. For the best configuration (10 features), the training accuracy increased regularly from 0.9217 at epoch 1 to 0.9975 at epoch 20, while the validation accuracy remained permanently high (≈0.9887–0.9994) and the validation loss stayed very low (≈0.0037–0.059). This indicates that the model converged easily without overfitting around the training process. Figure 5 presents the training and validation accuracy and loss curves, visually supported the steady improvement and low validation loss over the 20 epochs. The final test evaluation on 314,835 unseen samples

yielded an accuracy of 0.9994, with evaluation completed in approximately 35 seconds (≈4 ms per step), highlighting the feasibility of real-time deployment in IoT networks.

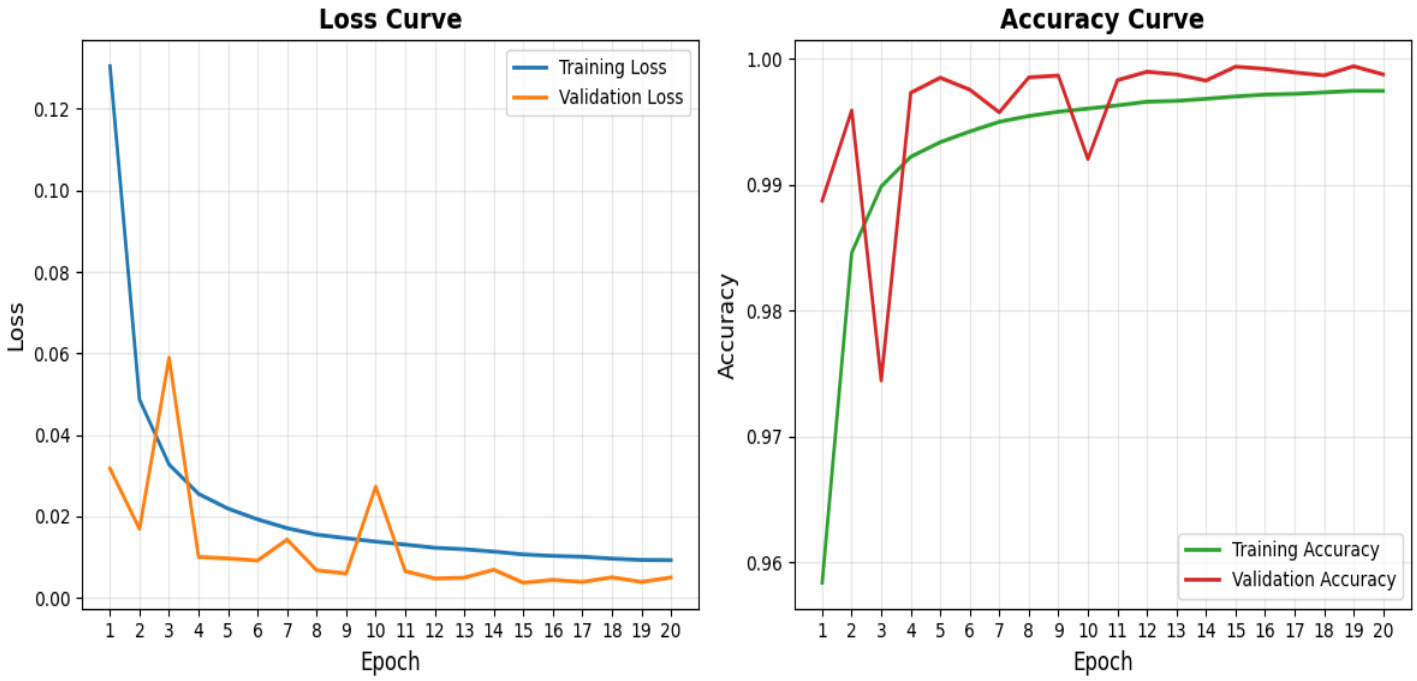

Figure 5. Training and validation accuracy and loss curves

5. Discussion

The proposed BiLSTM-based Intrusion Detection System (IDS) demonstrated remarkable performance on the ToN-IoT dataset, achieving 99.94% test accuracy across ten traffic categories. This result exceeds or matches the efficacy of previous deep learning models such as CNNs, InceptionTime [5–7], and hybrid ensemble approaches [11], while requiring fewer input features and preserving computational efficiency suitable for real-time IoT deployment. The model effectively handled class imbalance, exactly classifying both majority classes (e.g., scanning, DoS) and minority classes (e.g., MITM, ransomware) due to the SMOTE-based training strategy. Our BiLSTM-based IDS can be deployed on various IoT platforms, from small edge devices to centralized servers. It offers fast and reliable intrusion detection, even on hardware with limited resources, and can be updated over time as new data becomes available. Future work will explore a federated learning setup, allowing IoT nodes to collaboratively improve the model without sharing raw data, thereby maintaining both privacy and scalability. The ToN-IoT dataset was split using stratified random sampling to preserve the original class distribution and reduce potential bias or imbalance artifacts. Although full k-fold cross-validation was not feasible due to computational cost, this stratification and the large, diverse test set provide strong evidence that the reported accuracy is robust and not an artifact of a lucky split. Future work will extend the evaluation to cross-dataset validation (e.g., BoT-IoT, CICIoT2023) and k-fold cross-validation to further assess generalization. Feature reduction experiments further proved that the BiLSTM maintains high accuracy even with a compact feature set,

showing its adaptability to resource-constrained environments. Training dynamics indicated smooth convergence with low validation loss and no overfitting, while the final evaluation required only ≈4 ms per sample, demonstrating practical feasibility for live IoT networks. In comparison to traditional ML models [9,10] and federated or distributed frameworks [12–14], the BiLSTM strikes an effective balance between accuracy, simplicity, and scalability, making it a compelling choice for deployment in various IoT settings. Unlike CNN-based approaches that rely on converting network traffic into images, our BiLSTM directly exploits sequential dependencies in raw network flows, which are more representative of IoT communication patterns. Compared to ensemble frameworks such as XGBoost or CatBoost, which often incur significant computational overhead, our approach achieves similar or higher accuracy with lower inference cost, making it more deployable on edge devices. Furthermore, the ability to sustain performance even when the feature set is reduced highlights the practical robustness of the model in constrained environments. While this study was conducted using only a subset of the ToN-IoT dataset, future research should focus on cross-dataset validation, adaptive learning to cope with new and evolving attack patterns, and privacy preserving deployment strategies. Taken together, these findings suggest that the developed BiLSTM acts as a reliable, efficient, and scalable solution for multi-class intrusion detection in IoT networks.

6. Conclusions

The current work proposed a BiLSTM-based Intrusion Detection System (IDS) for IoT networks, experimented on the ToN-IoT dataset. The model achieved 99.94% test accuracy across ten traffic categories, demonstrating outstanding capability in distinguishing normal traffic from diverse cyberattacks, including minority classes such as MITM and ransomware. The proposed approach effectively addresses class imbalance through SMOTE-based training, and its high accuracy was maintained even with a reduced feature set, showing that it works well even on devices with limited resources. The model trained smoothly without overfitting and was able to evaluate new data almost instantly, with a very low delay of around 4 ms per sample. Compared to prior works—including deep learning, traditional ML, ensemble, and federated approaches The BiLSTM model stands out as an excellent choice for securing real-world IoT networks. It strikes a perfect balance between being highly accurate, straightforward to implement, and easy to scale up. Future research should focus on testing on different datasets, adaptive learning for evolving attacks, and deploying while protecting user privacy to further enhance robustness and applicability in dynamic IoT environments. Additionally, SHAP-based explainability will be integrated to provide per-sample feature attribution, enhancing trust in real-world deployments. The model will also be extended into a federated learning framework, enabling privacy-preserving updates across distributed IoT devices without centralizing sensitive traffic data. Future work will focus on further assessing the generalization ability of the proposed IDS. This includes performing cross-dataset validation such as training on ToN-IoT and testing on BoT-IoT or CICIoT2023 and implementing k-fold cross-validation to ensure consistent performance across different data partitions. In conclusion, the proposed BiLSTM-based IDS provides a robust, proficient, and scalable paradigm for multi-class intrusion detection in IoT networks, providing a reliable solution for both research and practical use. This work surpasses previously reported results on the ToN-IoT dataset, providing the highest multiclass accuracy to date and setting a strong baseline for future IDS research in IoT environments.

References

[1] S. Gupta et al., "A Survey on Intrusion Detection System in IoT Networks," Journal of Network and Computer Applications, 2024.
[2] H. Khan et al., "Lightweight CNN-BiLSTM Based Intrusion Detection Systems for Resource-Constrained IoT Devices," arXiv preprint, 2024.
[3] M. Z. Gueriani et al., "Deep Reinforcement Learning for Intrusion Detection in IoT: A Survey," arXiv preprint, 2024.
[4] K. S. Alzahrani et al., "IoT-Based Intrusion Detection System Using New Hybrid Deep Learning Algorithm," Electronics, 2024.
[5] A. R. Syed, M. M. Kamruzzaman, and M. A. Hossain, "Explainable Deep Learning-Based Intrusion Detection System for Internet of Things Networks," IEEE Access, vol. 11, pp. 84760–84775, 2023.
[6] M. Alsulami and A. Dehghantanha, "Leveraging Deep Learning for Intrusion Detection in IoT Through Visualized Network Data," IEEE Access, vol. 11, pp. 58132–58142, 2023.
[7] M. Rehman, H. Malik, and M. A. Khan, "Performance Evaluation of Multiclass Classification Models for ToN-IoT Network Device Datasets," Procedia Computer Science, vol. 218, pp. 164–171, 2023.
[8] Y. Tang et al., "A Distributed Intrusion Detection System Using Machine Learning for IoT Based on Industrial Framework," IEEE Internet of Things Journal, vol. 10, no. 4, pp. 3129–3142, 2023.
[9] M. S. Ahmed and M. A. Elsisi, "A Hybrid Feature Selection Method for ToN-IoT Dataset Using Evolutionary Algorithms and Machine Learning Models," Journal of Network and Computer Applications, vol. 202, p. 103393, 2022.
[10] A. Hussein and H. A. Ali, "Performance Analysis of Supervised Machine Learning Models for Intrusion Detection Using ToN-IoT Dataset," ICT Express, vol. 9, no. 1, pp. 89–96, 2023.
[11] L. Wang, X. Zhou, and H. Yu, "An Ensemble-Learning Framework for Intrusion Detection to Enhance Internet of Things' Devices Security," Computers & Security, vol. 124, p. 102983, 2023.
[12] M. R. Munir, M. H. U. Rehman, and S. A. Madani, "A Distributed Intrusion Detection System Using Machine Learning for IoT Based on Cloud, Fog, and Host Levels," Sensors, vol. 23, no. 4, p. 2345, 2023.
[13] R. H. Amini, S. R. Aghili, and M. GhasemiGol, "Federated Deep Learning for Intrusion Detection in IoT Networks," Neural Computing and Applications, vol. 35, pp. 23925–23938, 2023.
[14] Y. Zhang and Z. Lu, "FEDPG: A Novel Federated Principal Component Analysis Method on Grassmann Manifold," Pattern Recognition Letters, vol. 170, pp. 87–94, 2023.