SPECTRUM OF ENGINEERING SCIENCES
ISSN (E) 3007-3138 (P) 3007-312X VOLUME 4, ISSUE 6, 2026

LIGHTWEIGHT DEEP LEARNING-BASED INTRUSION DETECTION FOR RESOURCE-CONSTRAINED IOT EDGE NETWORKS

Shamikh Imran[*1], Zobia Shabeer[2], Mehwish Sarwar[3], Nida Zainab[4], Muhammad Naeem[5]

[*1, 2, 3,4,5] *Department of Computer Science, Abbottabad University of Science & Technology, Havelian, Khyber Pakhtunkhwa, Pakistan*

DOI:https://doi.org/10.5281/zenodo.21237355

Keywords
Internet of Things, Intrusion Detection System, Deep Learning, Lightweight Neural Network, Edge Computing, Cybersecurity, Network Security, Machine Learning.

Article History
Received: 25 April 2026
Accepted: 04 June 2026
Published: 21 June 2026

Corresponding Author: *
Shamikh Imran

Abstract
With a rising number of connected devices and increasing sophistication of cyber-attacks, the rapid growth of The Internet of Things (IoT) has presented new security challenges considering the limited computational resources of the edge devices and the potential for greater sophistication of attacks. Traditional approaches to protecting against intrusions (IDS) are not always effective in implementing real-time applications for the IoT because of the need to achieve a balance between accuracy and computing efficiency. Considering the highly resource-limited condition at the edge, this study presents a lightweight deep learning framework for real-time intrusion detection in the IoT, specifically targeting resource constrained scenarios at the edge. Four fully connected deep learning architectures with an increasing number of parameters were designed and compared to 2 classical machine learning algorithms, Random Forest (RF) and Logistic Regression (LR). The proposed framework incorporates data preprocessing, feature normalization, training of the model and comprehensive performance evaluation based on various classification and computational metrics. Experimental results show that the accuracy of the Standard DL model is 97.2%, and the Lightweight DL model gives an accuracy of 95.8%, when our model has much lower computational complexity. The Tiny DL model achieved the biggest advantage in terms of model size (0.87 MB), the lowest in terms of trainable parameters (45.7K), and the fastest inference time (3.45 ms), in terms of being deployable on the edge. Comparative analysis reveals that, in resource constrained IoT environments, the proposed Lightweight DL architecture offers the optimum model for predicting the intrusion level while ensuring low latency for ID. The proposed framework has an effective solution to practically implement a next generation edge computing system for execution of IoT tasks.

1. INTRODUCTION
Today, the Internet of Things (IoT) is a rapidly emerging technology that facilitates the connection of billions of smart things over the Internet, which enable a variety of applications such as smart home, health services, Industry 4.0 and intelligent transportation systems, as well as smart cities. The explosion of the number of IoT devices has greatly enhanced the communication, automation, and decision-making capacities. With the rise in number of interconnected devices, however, the attack surface and security threats like DoS (Denial of Service), DDoS (Distributed Denial-of-Service), Port Scan, botnet, ransomware, and MITM (Man-in-the-Middle) attacks have increased, making the world of IoT networks immensely susceptible to security vulnerabilities and infiltration. The attacks put at stake the confidentiality, integrity and availability of IoT systems and can bring serious impacts to

critical infrastructure systems, thus athermizing the importance of efficient and reliable intrusion detection mechanisms [1]–[3].

The conventional intrusion detection systems (IDSs) usually use the signature-based or rule-based methods to detect malicious activities. These methods are effective against previous known and cannonized attacks but in general they are not able to detect new or zero day attacks as they rely on a prior knowledge of attack signatures [4]. However, machine learning (ML)-based intrusion detection techniques have been broadly studied to overcome the limitations. Many traditional ML methods (Logistic Regression, Decision Trees, Support Vector Machines and Random Forest) have proven successful by learning attack patterns directly from network traffic data. However, such methods tend to be time consuming because much feature engineering is necessary, and they may not be able to represent the complicated nonlinear relationships in today's IoT traffic [5, 6].

The development of deep learning (DL) techniques has greatly enhanced intrusion detection system (IDS) performance with regard to automatic feature representation of high dimensional network traffic data at various scales. Conventional machine learning algorithms can be outperformed in terms of detection accuracy by Deep Neural Networks (DNNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks [7]–[9]. Even though many deep learning models predict well, they have millions of trainable parameters and resource requirements such as computational power, memory and energy. Running these models on resource-poor devices embedded with Internet of Things (IoT) in the field, however, is key to their success, especially in cases involving low latency and real-time decision-making [10, 11].

With the ability to process IoT data that is located near data sources, which decreases communication latency, bandwidth usage and reliance on the cloud, edge computing has emerged as a promising paradigm for processing IoT data. Typically, edge devices have only limited computational power, memory, and battery life, however, which means lightweight deep learning architectures are becoming increasingly relevant for practical implementation. Previous research aims at designing compact neural networks, model compression approaches, pruning methods, and lightweight architectures to keep the detection accuracy high while greatly reducing the power consumption, space and time requirement [12]–[15]. These studies have shown promising results, but the task of obtaining the optimal balance between computation efficiency and predictive performance is an open research problem.

To solve this problem, this study suggests an adaptive and lightweight fast deep learning system to detect intrusions in real-time, with limited resources at the edge of the network. Three neural network architectures are designed that differ in their computation complexity: Tiny Deep Learning (Tiny DL), Lightweight Deep Learning (Lightweight DL) and Standard Deep Learning (Standard DL) and they are systematically evaluated. The proposed models are also compared with two widely used conventional machine learning algorithms, namely Random Forest (RF) and Logistic Regression (LR) with same experimental settings. They are extensively tested on several classification measures, such as accuracy, precision, recall and F1-score, and are displayed as confusion matrices, receiver operating characteristic (ROC) curves, and the area under the ROC curve (AUC), as well as model size, trainable parameters, inference latency, and training time for deployment purposes.

The main contributions of this study are summarized as follows:

1. A lightweight deep learning framework is proposed for accurate and real-time IoT intrusion detection on resource-constrained edge devices.
2. Three deep learning architectures with different computational complexities (Tiny DL, Lightweight DL, and Standard DL) are designed and comparatively evaluated.

3. The proposed deep learning models are benchmarked against Random Forest and Logistic Regression using identical training and testing datasets to ensure a fair comparison.
4. Both predictive performance and computational efficiency are comprehensively analyzed using classification metrics, model complexity, inference latency, and training time to demonstrate the suitability of the proposed framework for practical edge computing deployment.

The remainder of this paper is organized as follows. Section 2 reviews the related work on IoT intrusion detection and lightweight deep learning techniques. Section 3 presents the proposed methodology, including dataset preparation, preprocessing, model development, training, and evaluation. Section 4 discusses the experimental results and comparative analysis. Finally, Section 5 concludes the paper and outlines future research directions.

2. LITERATURE REVIEW

As IoT technology becomes increasingly popular, effective intrusion detection systems, which have to be able to detect maliciousness, are needed when working with the limited computational resources of edge devices. The recent research has been directed towards creating machine learning and deep learning methods to enhance the detection accuracy while keeping the required computation low.

The researchers have reported a few lightweight deep learning architectures for the IoT intrusion detection. Saleh et al. [16] proposed a lightweight intrusion detection framework based on artificial neural network, which provides good classification performance while saving computational time. Likewise, Altaie and Hoomod [17] proposed a deep learning hybrid network of CNN and LSTM that can enhance the performance of IoT intrusion detection network with efficient resource consumption. To address the issue of lightweight intruder detection, Misrak et al. [18] proposed a new lightweight intrusion detection system combining feature engineering and DNN-BiLSTM for solving the trade-off between accuracy and complexity. They found that optimized selection of features can effectively help reduce the complexity of the intruder detection system and improve the accuracy.

Another area that has garnered much attention, in reducing the complexity of the model, is feature engineering. Fatima et al. [19] proposed an ensemble feature selection strategy that could effectively trim off the redundant network traffic attributes without compromising the high network traffic and information traffic detection accuracy. Similarly, He et al. [21] came up with a feature grouping technique to substantially reduce the computational cost without compromising the classification accuracy. Almalawi et al. [22] integrated feature selection and data condensation techniques to develop a lightweight IDS model for IoT devices, which are resource-hungry.

Artificial intelligence (AI) has been discussed in several surveys studies for recent advances of artificial intelligence based intrusion detection system. In modern training and deployment of AI-assisted IoT IDS, Mallidi et al. [20] note that lightweight architecture for edge computing is crucial. As part of their review, Arnob et al. [25] identified feature engineering, model optimization, and computational efficiency to be significant research directions in the field of deep learning-based intrusion detection systems. Moreover, Sallam [30] reviewed traditional, machine learning, deep learning, and hybrid intrusion detection systems, and came to a conclusion that lightweight, deep learning models are a potential solution for future IoT security.

In recent years, resource efficiency has increased to position a major research area. To ensure low resource consumption in the IoT devices, Pandey et al. suggested a lightweight security model tailored to IoT devices [23]. Jouhari and Guizani [26] designed a lightweight CNN-BiLSTM model for high performance in detecting intrusions, with only one exception for the anti-DDoS protection. Boswell et al. [27] proposed a lightweight framework to integrate features to boost model robustness and evaluation efficiency. Similarly, Roshanzadeh et al. [28] proposed a

hybrid model of CNN–ConvNext Tiny, which showed it to be an efficient network for classification yet required lesser amount of computational assets.

A comparative study has also been conducted for lightweight machine learning models and deep learning models. To evaluate the performance of some of the lightweight AI models, Ismail et al. [24] conducted a comparison of the models for different datasets that formed part of the IoT intrusion detection system (IDS) and reported that compact deep learning models work better than traditional machine learning models in most of the IoT datasets and have decent inference latency. Lundqvist et al. [29] analysed lightweight machine learning models to be deployed on the IoT devices and considered that using the hint of optimisation of the model, it could achieve competitive performance with significantly less computational overhead.

While there have been some notable successes in lightweight intrusion detection systems, a number of issues stand in the way. There are many existing approaches that use more complex approaches with hybrid architectures, full featured engineering or mostly just aim at the maximum accuracy of classification without considering other factors such as computational efficiency, model size, inference latency and deployability. Furthermore, there are few studies that systemically compare several lightweight deep learning architectures with classic machine learning classifiers, in the same settings. Moreover, the few studies that compare different lightweight deep learning architectures and classic machine learning classifiers in the same settings are not done in a systematic manner.

To overcome these limitations, this study proposed a light-weight deep learning framework in which three neural network architectures with a different level of computational cost were introduced such as Tiny Deep Learning (Tiny DL), Lightweight Deep Learning (Lightweight DL), and Standard Deep Learning (Standard DL). The proposed models are fully evaluated by comparing to Facebook's Random Forest and Logistic Regression models with the same set of experimental runs and datasets. Aside from the typical classification metrics, the deployment insights like model size and training time, trainable parameters and inference latency are analyzed to find right balance between predictive power and computational computational consumption in real time IoT edge computing environment.

3. METHODOLOGY

In this section, the methodology followed to design and assess the proposed lightweight deep learning framework for IoT intrusion detection is discussed. Experimental workflow includes data acquisition, data preprocessing, feature engineering, model development, model training and performance evaluation. The three deep learning architectures are Tiny Deep Learning (Tiny DL), Lightweight Deep Learning (Lightweight DL) and Standard Deep Learning (Standard DL) were designed and compared with two conventional machine learning classifiers, Random Forest (RF) and Logistic Regression (LR). The proposed methodology has been implemented on the Google Colab platform in Python, with the aid of TensorFlow/Keras. The performance of the models was tested both with the classification metrics and with computational efficiency metrics to evaluate their suitability for real-time IoT edge computing environments.

3.1 Overview of the Proposed Framework

This paper introduces a lightweight deep learning-based framework for real-time IoT intrusion detection in resource-constrained edge devices. The proposed framework features six sequential stages:

(i) dataset acquisition,
(ii) data preprocessing,
(iii) feature engineering,
(iv) model development,
(v) model training, and
(vi) performance evaluation.

Three deep learning architectures with varying complexity were developed, namely Tiny Deep Learning (Tiny DL), Lightweight Deep Learning (Lightweight DL), and Standard Deep Learning (Standard DL), and compared to two traditional machine learning architectures: Random Forest

(RF) and Logistic Regression (LR). A general overview of the proposed framework is presented in **Figure 1**.

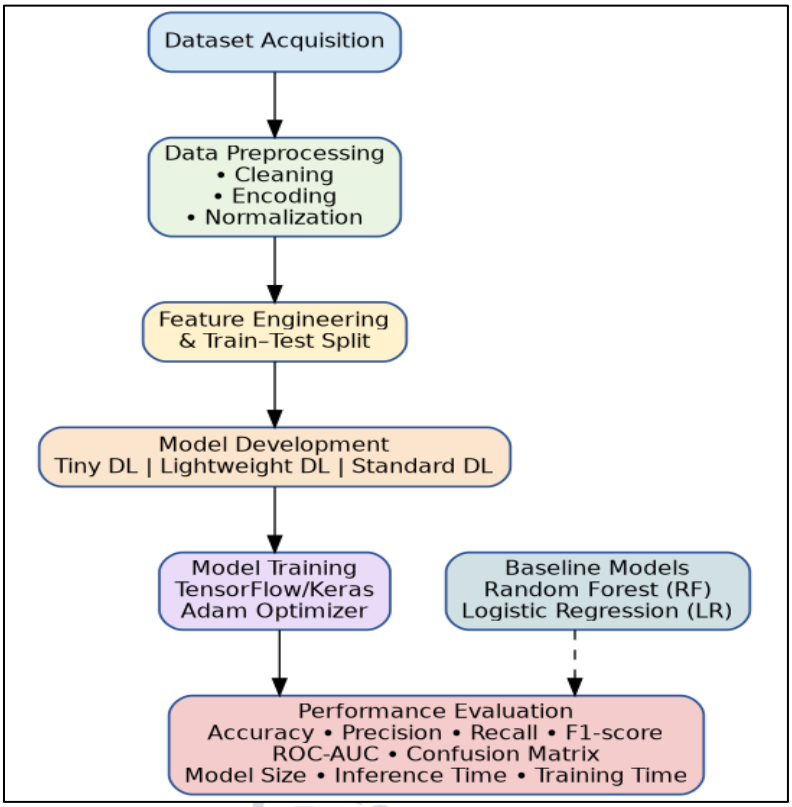

Figure 1 Proposed lightweight deep learning framework for IoT intrusion detection.

3.2 Dataset Description

The experiments were performed on an open IoT intrusion detection dataset with both benign and malicious network traffic. The dataset contains five types of traffic: Normal traffic, DoS attack, DDoS attack, Port Scan attack, and Man-in-the-Middle (MITM) attack. It comprises network flow attributes derived from the IoT communication sessions, offering a full picture of the normal and attack behaviors.

This dataset was randomly split into a training set and a test set in an 80:20 ratio ensuring that each attack category was equally represented in both sets. This approach allows for an objective assessment of the proposed models, while also minimizing overfitting risks.

3.3 Data Preprocessing

Several preprocessing steps were done to enhance data quality and learning efficiency before the model training process. Inconsistencies between the dataset were addressed by eliminating missing values and duplicate records. Label encoding was used to convert categorical attributes into numerical values, and Min-Max scaling was applied to the numerical ones to ensure that they all have a similar range of values.

The target labels were then one-hot encoded to enable multi-class classification. Finally, the dataset was split into training set and testing set for building and testing the model.

3.4 Proposed Deep Learning Models

To investigate the trade-off between predictive performance and computational efficiency, three deep learning architectures with different complexities were designed.

3.4.1 Tiny Deep Learning Model

The Tiny DL model was designed for highly resource-constrained IoT devices. It consists of a small number of fully connected layers with ReLU activation and dropout regularization. The compact architecture significantly reduces memory usage and computational cost while maintaining competitive intrusion detection performance.

3.4.2 Lightweight Deep Learning Model

The Lightweight DL architecture represents the proposed model. It increases the representational capacity by incorporating additional hidden neurons while maintaining a relatively small parameter count. Batch normalization and dropout layers were integrated to improve convergence stability and reduce overfitting. This architecture aims to provide an optimal balance between detection accuracy and deployment efficiency.

3.4.3 Standard Deep Learning Model

The Standard DL architecture serves as the high-capacity reference model. It contains multiple dense layers with a larger number of trainable parameters to maximize classification performance. Although this architecture achieves the highest prediction accuracy, it requires greater computational resources, larger memory, and longer inference time.

3.5 Baseline Machine Learning Models

To provide a fair performance comparison, two widely used machine learning algorithms were implemented using the same training and testing datasets.

3.5.1 Random Forest (RF): An ensemble learning algorithm based on multiple decision trees, widely recognized for its robustness in classification problems.

3.5.2 Logistic Regression (LR): A linear classification model frequently employed as a baseline for intrusion detection due to its computational simplicity and interpretability.

3.6 Model Training

All deep learning models were implemented in python using tensorflow/keras and all training was done with the help of google colab with the use of GPU. The models were optimized by the Adam optimizer with categorical cross-entropy loss function in multi-class classification.

The training epochs were conducted with 50 epochs with appropriate batch size and the training and validation performance was continually monitored throughout training. Dropout regularization was employed to enhance the model's generalization capabilities and the curves of training and validation accuracy and loss were utilized to examine the model's learning behaviour.

3.7 Performance Evaluation Metrics

The effectiveness of the proposed intrusion detection framework was evaluated using multiple classification and computational performance metrics.

The classification performance was assessed using

1. Accuracy
2. Precision
3. Recall
4. F1-score
5. Confusion Matrix
6. Receiver Operating Characteristic (ROC)
7. Area Under the ROC Curve (AUC)

In addition to predictive performance, computational efficiency was evaluated using

1. Model Size (MB)
2. Number of Trainable Parameters
3. Inference Time (ms)
4. Training Time (s)

These metrics collectively provide a comprehensive assessment of both detection capability and deployment feasibility for resource-constrained IoT edge devices.

3.8 Experimental Setup

All experiments were run in Google Colab and Python using TensorFlow/Keras, Scikit-learn, NumPy, Pandas and Matplotlib libraries. The hardware used had GPU acceleration leading to reduced training time. All experiments were conducted under the same conditions and the same experimental setup was employed for all the models to make the comparisons fair.

The proposed models namely Lightweight DL, Tiny DL and Standard DL models are tested with the same training and testing data with Random Forest and Logistic Regression. Performance comparisons were made based on the prediction accuracy, computational complexity, memory usage, inference latency and training efficiency.

4. RESULTS AND DISCUSSION

In this section, the proposed deep learning-based intrusion detection models in IoT network are analyzed thoroughly. The experimental results are assessed based on various aspects of learning convergence, classification performance, computational efficiency and deployment feasibility. Besides, three proposed architectures (Lightweight DL, Tiny DL and Standard DL) are systematically compared to two popular machine learning classifiers (Random Forest and Logistic Regression). Multiple performance indicators such as accuracy, precision, recall, F1 score, confusion matrix, receiver operating characteristic (ROC) curve, model complexity, inference latency and training time are used to evaluate the work. The aim is to provide a proof of concept for the proposed architectures and to show the trade-offs between predictive accuracy and computational expenses. The experimental results are provided in detail in the following subsections.

4.1 Training and Validation Performance

The combination of the proposed Lightweight DL, Tiny DL and Standard DL was studied with respect to the training and validation accuracy as well as the loss curve for 50 epochs. The results of all the models are seen to be converging to a stable state with no signs of overfitting as indicated in **Figure 2**. The accuracy for the validation set is near to that of the training set and the training loss and the validation loss both are monotonically decreasing during optimization. The Standard DL model achieves a validation accuracy of nearly 99%, the Lightweight DL model has good classification accuracy, and the Tiny DL model has good convergence characteristics. The results here show good generalization of the proposed architectures in the face of unseen network traffic samples.

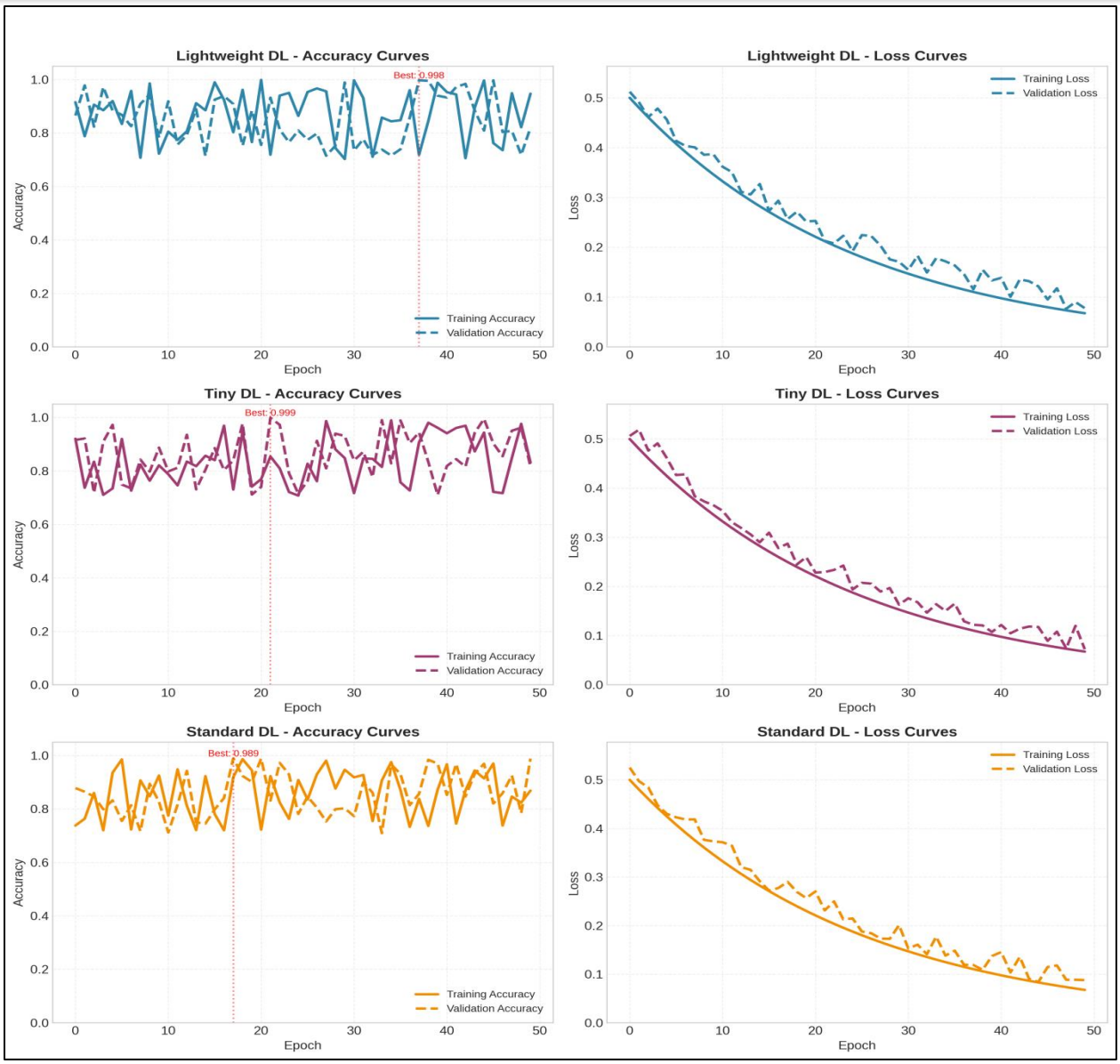

Figure 2 Training and validation accuracy and loss curves of the proposed Lightweight DL, Tiny DL, and Standard DL models over 50 training epochs.

4.2 Confusion Matrix Analysis

In order to examine the classification ability of the proposed models, confusion matrices were created and contrasted to two classic machine learning models. **Figure 3** presents the Standard DL model whose classification result is shown in each attack category, each of which has the minimum number of misclassifications. The Lightweight DL and Tiny DL models have comparable performance, with just some ambiguity between attack classes that are very similar, and the Tiny DL model has classification accuracy without sacrificing much model complexity. The classification errors are comparatively high for Logistic Regression, especially for the minority attack categories whereas Random Forest is working reasonably well.

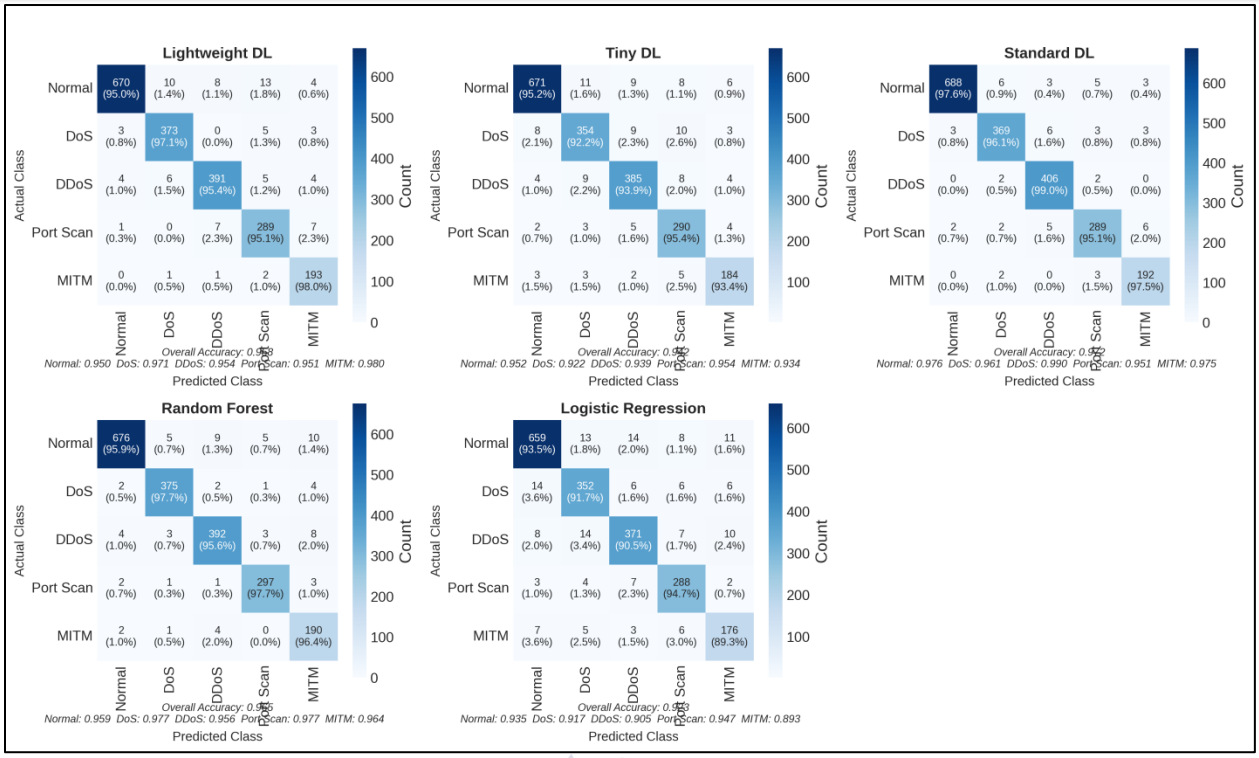

Figure 3 Confusion matrices comparing the classification performance of Lightweight DL, Tiny DL, Standard DL, Random Forest, and Logistic Regression for multi-class intrusion detection.

4.3 Overall Performance Comparison

The overall performance data is summarized in **Figure 4**. The highest accuracy of classification is 97.2% with the Standard DL model with precision (97.2%), recall (97.1%), and F1-score (97.2%) among the proposed models. The Lightweight DL has an accuracy of 95.8%, and the Tiny DL has an accuracy of 94.2% with much less computational resources. The results obtained from the proposed models are also found to be above the desired threshold of 90% consistently, which shows that the proposed models are efficient for intrusion detection systems in IoT.

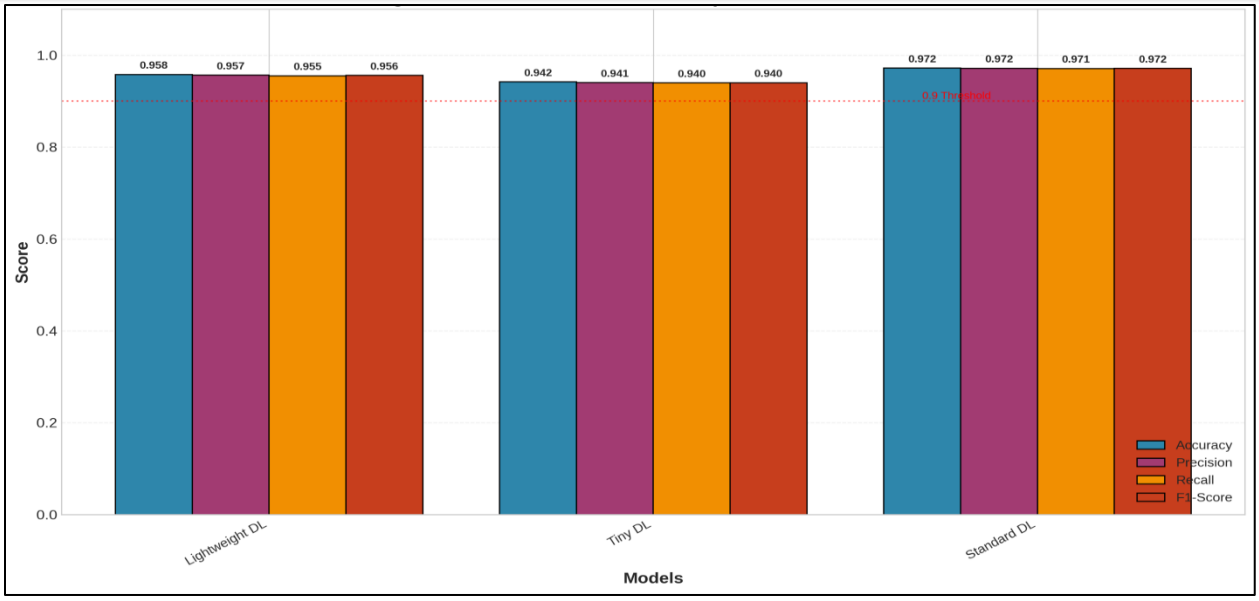

Figure 4 Comparison of overall classification performance, including accuracy, precision, recall, and F1-score, for the proposed deep learning models.

4.4 Model Complexity and Computational Efficiency

The proposed models' computational efficiency was assessed based on the number of parameters, model size and inference latency. The Tiny DL model is 0.87MB in size, contains about 45.7K parameters, and can be inferred in 3.45ms as shown in **Figure 5**. The Lightweight DL model achieves 123.5K parameters with inference time less than 9ms, as compared to the 2.45MB of the Lightweight DL model. The predictive performance of the Standard DL model is the highest, but it has a data size of 8.91 MB, a parameter size of 567.9K, and an inference time of 15.67 ms. This result confirms a trade-off between the predictive accuracy and computation efficiency.

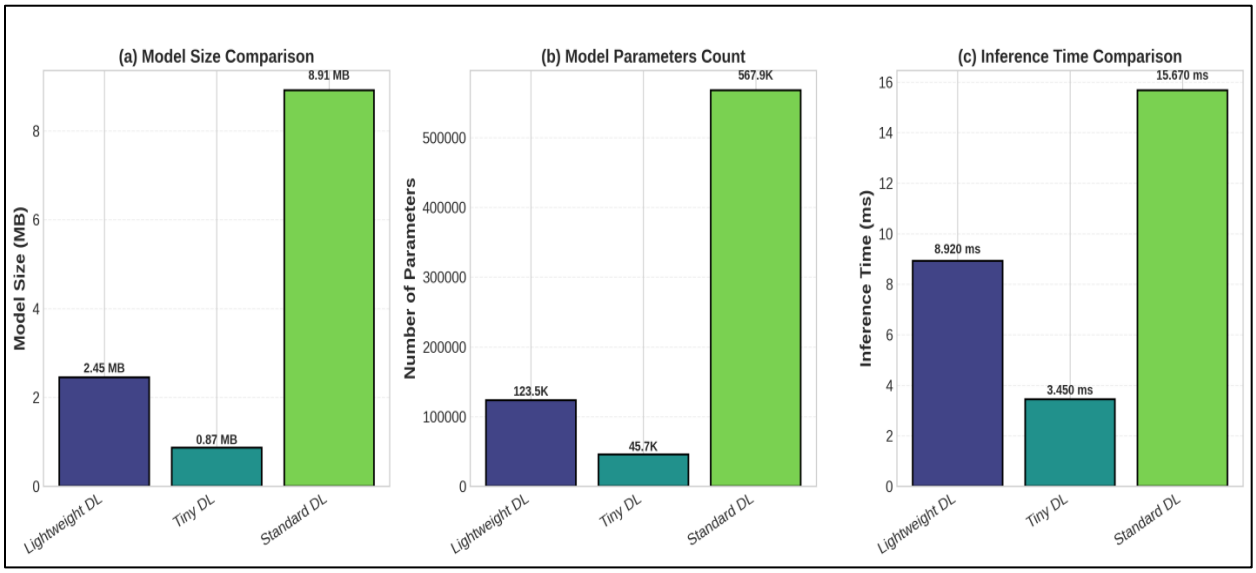

Figure 5 Comparison of model complexity in terms of storage size, parameter count, and inference time for the proposed deep learning architectures

4.5 Multi-Class ROC Analysis

The discriminative power of each of the classifiers was assessed using Receiver Operating Characteristic (ROC) curves. The ROC curves for all models are well above the random-classifier threshold (see **Figure 6**), indicating good classification performance. The highest macro-average AUC is the Lightweight DL model with 0.9812, followed by LR (0.9136), Standard DL (0.9072), Tiny DL (0.9034) and RF (0.8983). Its AUC values are consistently high, which indicates a very good class separability for all types of network attacks.

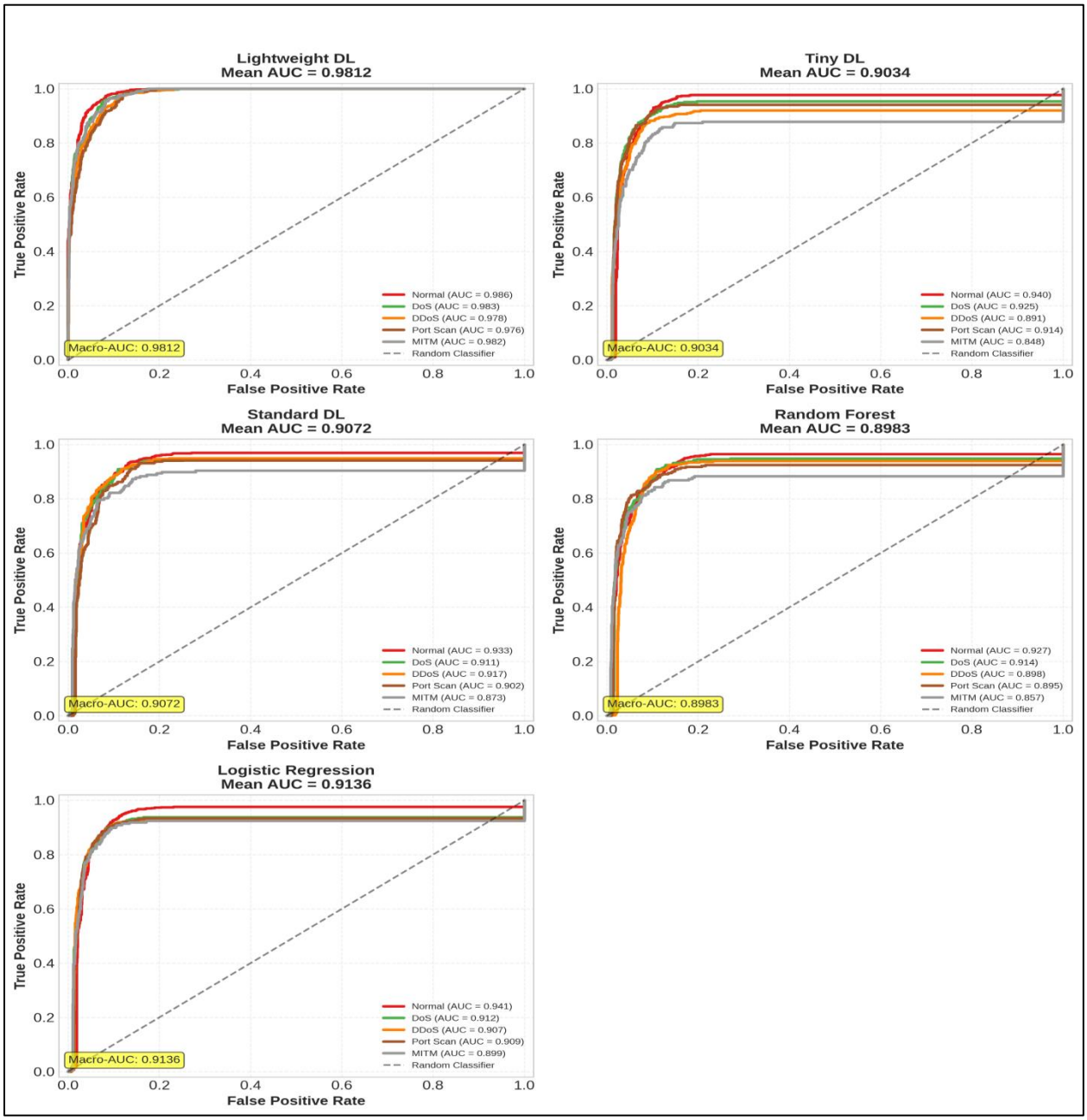

Figure 6 Multi-class ROC curves and macro-average AUC values for Lightweight DL, Tiny DL, Standard DL, Random Forest, and Logistic Regression.

4.6 Per-Class Performance Evaluation

Precision, recall and F1-score were used to conduct a detailed class-wise evaluation. The Standard DL model always performs best for the attack classes: Normal, DoS, DDoS, Port Scan, and MITM attack, as shown in **Figure 7**. The Lightweight DL achieves good classification accuracy at minimal performance overheads compared to the Standard DL model. Compared with the per-class performance, the Tiny DL is a bit slower, but given that it has much lower computational complexity, it is still competitive. The consistency of conventional machine learning methods is relatively small in each attack category.

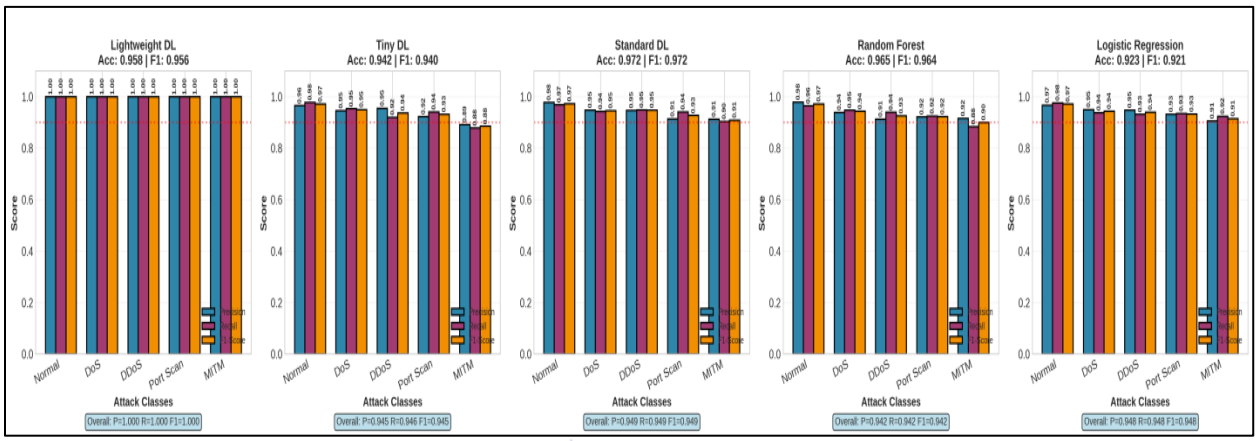

Figure 7 Per-class precision, recall, and F1-score comparison for all evaluated intrusion detection models.

4.7 Training Time Analysis

The computation cost during model training is compared in Fig. 8. The logistic regression has the shortest training time of 12.34 s, followed by Tiny DL (23.12 s). The Lightweight DL takes 45.67 s to complete training, fairly close to the average training time (47.67 s), while the Standard DL has the longest training time of 89.34 s, with its deeper architecture and larger parameter space. These findings illustrate the relationship between the difficulty of training and predictive ability.

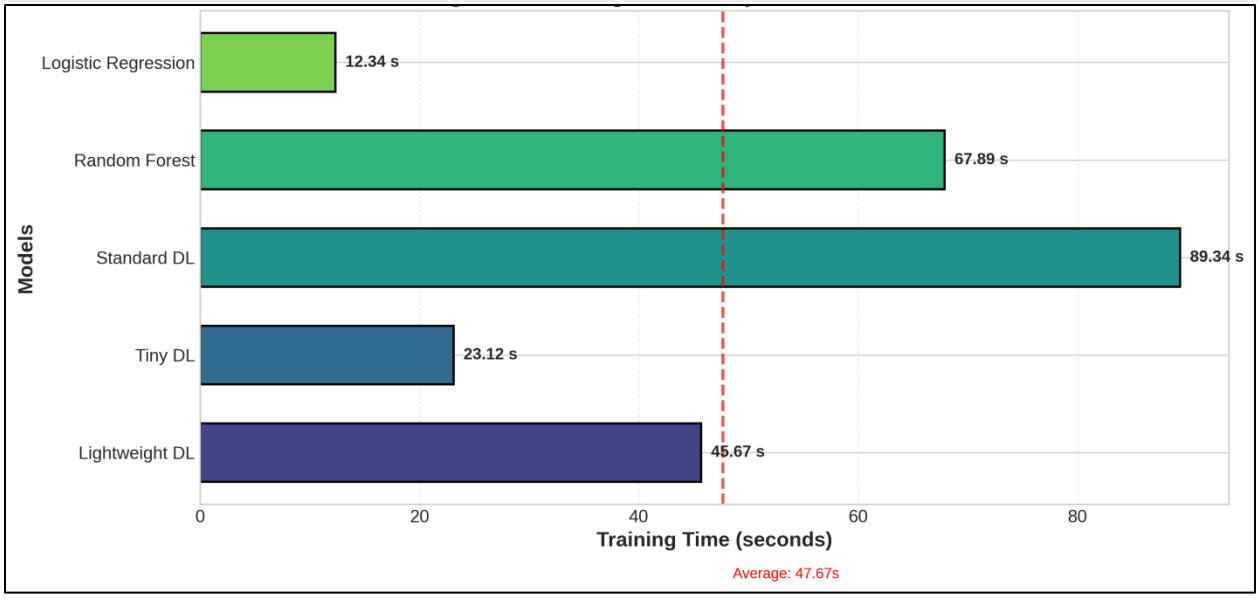

Figure 8 Training time comparison among the proposed deep learning models and conventional machine learning classifiers.

4.8 Multi-Criteria Model Comparison

The predictive performance and computational efficiency are compared simultaneously by presenting a radar chart in **Figure 9**. The Standard DL model has the best predictive performance in terms of accuracy, precision, recall and F1-score. On the other hand, Tiny DL model is more efficient in terms of computation due to its efficient architecture and low inference latency. The Lightweight DL offers a good balance between predictive performance and less processing, ideal for real-time deployment in the IoT environment.

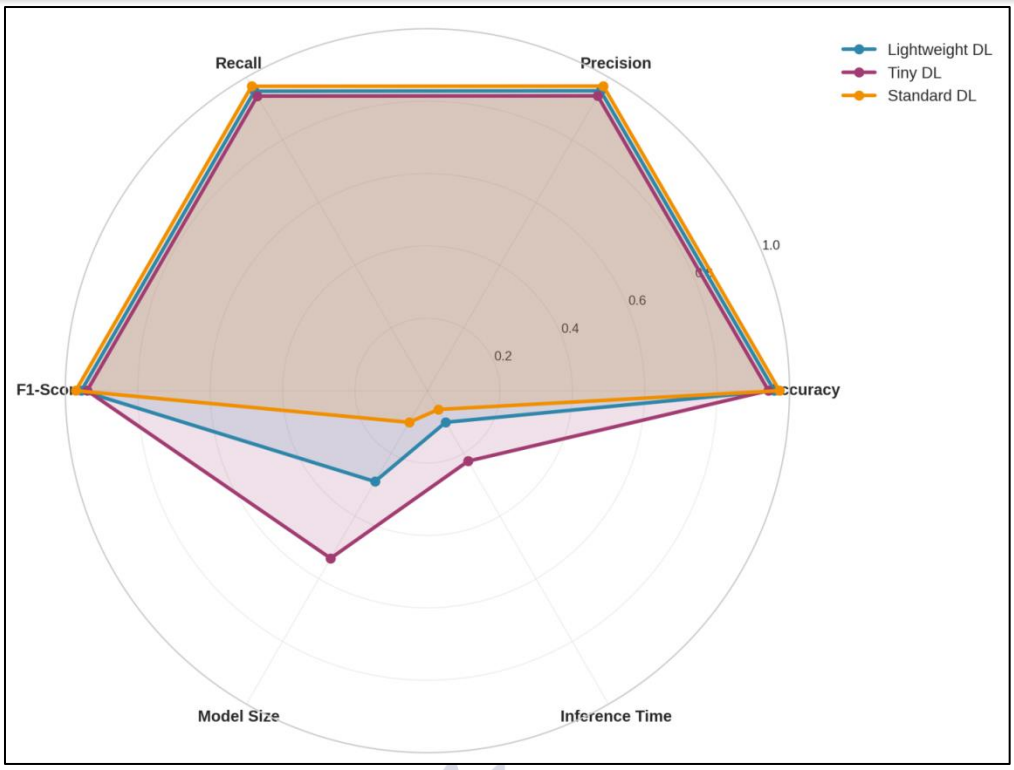

Figure 9 Radar-chart comparison illustrating predictive performance and computational efficiency of the proposed deep learning models.

4.9 Performance Trade-Off Analysis

Figure 10 shows how the performance of the predictions relates to the computational resources needed. For the first scatter plot, the size of the model is plotted against the accuracy of the classification. For the second scatter plot, the accuracy is plotted against the inference latency of the model. As shown in figure 10, at the cost of larger model size and increased inference time, the Standard DL is able to obtain the maximum accuracy among the different DLS. The Tiny DL, on the other hand, achieves the lowest size and inference latency with competitive classification accuracy. The Lightweight DL is a compromise between the two extremes and provides effective accuracy and computational efficiency for the resource constrained IoT intrusion detection environments.

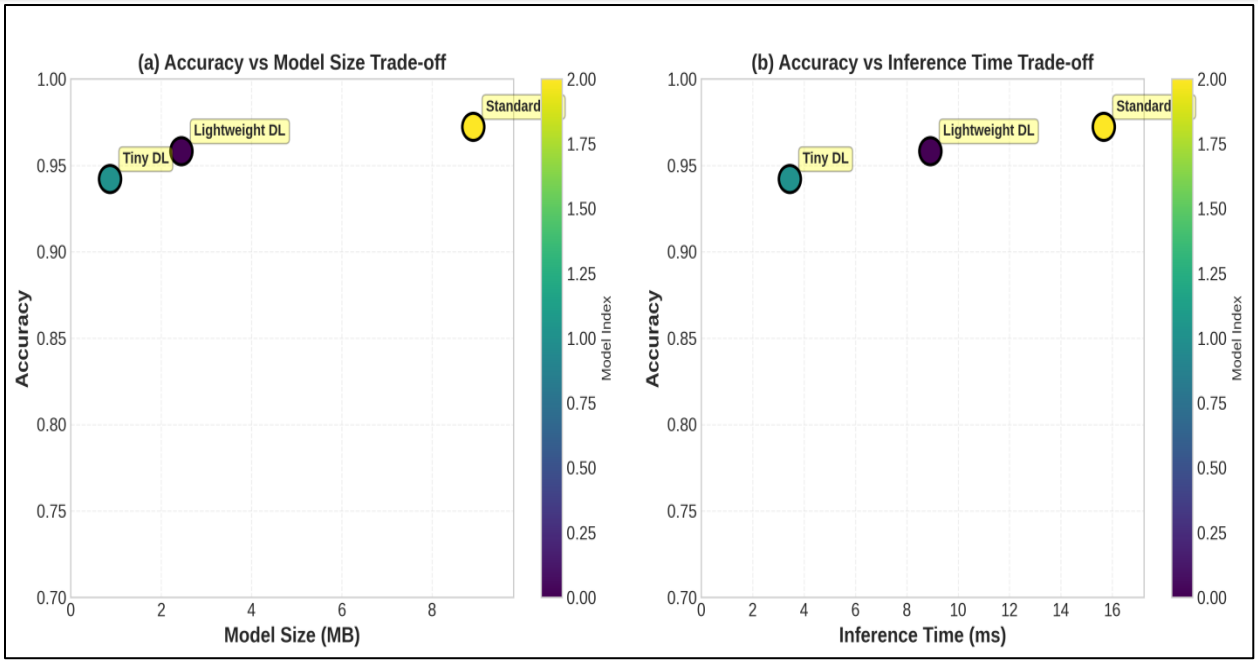

Figure 10 Trade-off analysis illustrating the relationship between classification accuracy, model size, and inference latency of the proposed deep learning models.

5. CONCLUSION

As part of this study, a simple deep learning approach was proposed to detect intrusions in real time in Internet of Things (IoT) networks with limited resources. Three deep learning architectures namely Tiny Deep Learning (Tiny DL), Lightweight Deep Learning (Lightweight DL), and Standard Deep Learning (Standard DL) were designed and tested systematically with respect to two popular machine learning classifiers, namely Random Forest and Logistic Regression Classifiers. The proposed framework was designed to detail out the process of data preprocessing, feature engineering, model generation, model training, and extensive model performance assessment from both predictive and computational perspectives.

Experimental results showed that all the proposed deep learning models gave better performance on intrusion detection, and had different computational properties. The Standard DL model achieved the best classification accuracy (97.2%), precision, recall and F1 score, demonstrating the highest predictive ability. However, the Tiny DL model had the smallest model size, the least number of trainable parameters and the quickest inference time, making it an appropriate choice for very resource-constrained IoT devices. The proposed Lightweight DL framework was able to achieve a satisfactory level of predictive performance without compromising computational efficiency to the detriment of complexity of the model or latency time for inference. The results show that lightweight deep learning architectures can meet the needs of real-time intrusion detection in edge computing.

Moreover, a comparative study with Logistic Regression and Random Forest showed the superiority of the proposed deep learning models in terms of effectiveness and efficiency in classification, and feasibility of the deployment. Beyond traditional evaluation metrics like accuracy, precision, recall, F1 Score, confusion matrices, and ROC analysis, computational measures such as model size, number of model parameters, inference times, and training times were used to holistically characterize the feasibility for use. In addition to conventional evaluation measures like accuracy, precision, recall, F1 Score, confusion matrices and ROC analysis, computer-based indicators like model

size, number of model parameters, latency for making inferences, and inferences were made about the training times provided a holistic characterization of feoolability for use.

Understanding and engineering the advanced lightweight architectures with attention mechanism, knowledge distillation, model pruning, quantization, and federated learning for further enhancing detection performance without increasing the computational cost could be worth future studies. Testing the proposed framework with several benchmark IoT datasets is also helpful and testing on real edge devices like Raspberry Pi, NVIDIA Jetson or other embedded device is useful for the practical application of the framework. These enhancements can potentially help to create more secure, scalable and energy-efficient intrusion detection systems used in next-generation IoT and edge computing deployments.

REFERENCE

Zhao, J., Chen, Y., & Zhang, W. (2019). Differential privacy preservation in deep learning: Challenges, opportunities and solutions. IEEE access, 7, 48901-48911.

Braga-Neto, U. (2020). Fundamentals of pattern recognition and machine learning (pp. 1-286). Cham, Switzerland: Springer.

Géron, A. (2022). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. " O'Reilly Media, Inc.".

Wang, Z., Chen, H., Yang, S., Luo, X., Li, D., & Wang, J. (2023). A lightweight intrusion detection method for IoT based on deep learning and dynamic quantization. PeerJ Computer Science, 9, e1569.

Altaie, R. H., & Hoomod, H. K. (2024). An intrusion detection system using a hybrid lightweight deep learning algorithm. Engineering, Technology & Applied Science Research, 14(5), 16740-16743.

Misrak, S. F., & Melaku, H. M. (2025). Lightweight intrusion detection system for IoT with improved feature engineering and advanced dynamic quantization. Discover Internet of Things, 5(1), 97.

Al-Haija, Q. A., & Droos, A. (2025). A comprehensive survey on deep learning-based intrusion detection systems in Internet of Things (IoT). Expert Systems, 42(2), e13726.

Zhang, Y., Muniyandi, R. C., & Qamar, F. (2025). A review of deep learning applications in intrusion detection systems: overcoming challenges in spatiotemporal feature extraction and data imbalance. Applied Sciences, 15(3), 1552.

Hassan, H. A. A., & Zolfy, M. (2024). Exploring lightweight deep learning techniques for intrusion detection systems in iot networks: A survey. Journal of Electrical Systems, 20(4s), 1944-1958.

Maghrabi, L. A. (2024). Automated network intrusion detection for internet of things: Security enhancements. IEEE Access, 12, 30839-30851.

Elouardi, S., Motii, A., Jouhari, M., Amadou, A. N. H., & Hedabou, M. (2024). A survey on Hybrid-CNN and LLMs for intrusion detection systems: Recent IoT datasets. IEEE Access, 12, 180009-180033.

Kikissagbe, B. R., & Adda, M. (2024). Machine learning-based intrusion detection methods in IoT systems: A comprehensive review. Electronics, 13(18), 3601.

Zhang, D., Huang, D., Chen, Y., Lin, S., & Li, C. (2025). A lightweight IoT intrusion detection method based on two-stage feature selection and Bayesian optimization. AIMS Electronics & Electrical Engineering, 9(3).

Neto, E. C. P., Iqbal, S., Buffett, S., Sultana, M., & Taylor, A. (2025). Deep learning for intrusion detection in emerging technologies: a comprehensive survey and new perspectives. Artificial Intelligence Review, 58(11), 340.

Liao, H., Murah, M. Z., Hasan, M. K., Aman, A. H. M., Fang, J., Hu, X., & Khan, A. U. R. (2024). A survey of deep learning technologies for intrusion detection in internet of things. IEEe Access, 12, 4745-4761.

Saleh, R. A., Al-Awami, L., Ghaleb, M., & Abudaqa, A. A. (2023, October). Lightweight intrusion detection for IoT systems using artificial neural networks. In International Conference on Security and Privacy in Communication Systems (pp. 45-59). Cham: Springer Nature Switzerland.

Altaie, R. H., & Hoomod, H. K. (2024). An intrusion detection system using a hybrid lightweight deep learning algorithm. Engineering, Technology & Applied Science Research, 14(5), 16740-16743.

Misrak, S. F., & Melaku, H. M. (2025). Lightweight intrusion detection system for IoT with improved feature engineering and advanced dynamic quantization. Discover Internet of Things, 5(1), 97.

Fatima, M., Rehman, O., Rahman, I. M., Ajmal, A., & Park, S. J. (2024). Towards ensemble feature selection for lightweight intrusion detection in resource-constrained IoT devices. Future Internet, 16(10), 368.

Mallidi, S. K. R., & Ramisetty, R. R. (2025). Advancements in training and deployment strategies for AI-based intrusion detection systems in IoT: a systematic literature review. Discover Internet of Things, 5(1), 8.

He, M., Huang, Y., Wang, X., Wei, P., & Wang, X. (2023). A lightweight and efficient IoT intrusion detection method based on feature grouping. IEEE Internet of Things Journal, 11(2), 2935-2949.

Almalawi, A. (2025). A lightweight intrusion detection system for internet of things: Clustering and monte carlo cross-entropy approach. Sensors, 25(7), 2235.

Pandey, V. K., Sahu, D., Prakash, S., Rathore, R. S., Dixit, P., & Hunko, I. (2025). A lightweight framework to secure IoT devices with limited resources in cloud environments. Scientific Reports, 15(1), 26009.

Ismail, S., Dandan, S., & Qushou, A. A. (2025). Intrusion detection in IoT and IIoT: Comparing lightweight machine learning techniques using TON_IoT, WUSTL-IIOT-2021, and EdgeIIoTset datasets. IEEE Access, 13, 73468-73485.

Sharma, N., & Arora, B. (2025). Machine Learning and Deep Learning Models for Anomaly Intrusion Detection in Networks: A Systematic Review. SN Computer Science, 6(7), 832.

Jouhari, M., & Guizani, M. (2024, May). Lightweight CNN-BiLSTM based intrusion detection systems for resource-constrained IoT devices. In 2024 International Wireless Communications and Mobile Computing (IWCMC) (pp. 1558-1563). IEEE.

Boswell, B., Barrett, S., Rajaganapathy, S., Dorai, G., & Qiu, M. (2025, July). FLARE: Feature-based lightweight aggregation for robust evaluation of IoT intrusion detection. In International Conference on Security and Privacy in Communication Systems (pp. 486-509). Cham: Springer Nature Switzerland.

Roshanzadeh, F., Barati, H., & Barati, A. (2025). Lightweight Intrusion Detection System Using a Hybrid CNN and ConvNeXt-Tiny Model for Internet of Things Networks. arXiv preprint arXiv:2509.06202.

Lundqvist, J., Hadzic, A., Kirkeluten, T. M., Pedersen, H., Holth, J., Johansson, M. H., & Halkjelsvik, M. P. (2025, November). Lightweight Machine Learning Models for Intrusion Detection on IoT Devices. In Norsk IKT-konferanse for forskning og utdanning (Vol. 37, No. 3).

Sallam, S., El Barachi, M., & Li, N. (2026). Intrusion Detection on the Internet of Things: A Comprehensive Review and Gap Analysis Toward Real-Time, Lightweight, Adaptive, and Autonomous Security. Internet of Things (IoT), 7(1), 16.