The Research Analytics
Volume 3, Special Issue 1, Jan 2026
ISSN (Online): 3107-6165

The Research Analytics
(A Peer Reviewed and Open Access Journal)

4

Optimized Multiclass Classification of IoT Cyberattacks Using a Hybrid Random Forest–Gradient Boosting Stacking Ensemble

Mustapha Ismail Kwari[1], N Rajkumar[2], and C Selvarathi[3]

1) Research Scholar, CSE, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India. vtd1224@veltech.edu.in
2) Professor, CSE, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India. nrajkumar@veltech.edu.in
3) Assistant Professor, CSE, Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamilnadu, India .selvarathigce@gmail.com

Pages No: 20-30

Abstract: *Cybersecurity threats have increased due to the Internet of Things' (IoT) growing prominence, precise and computationally efficient intrusion detection solutions are crucial. However, the efficacy of standalone machine learning models is limited by the diverse nature of IoT traffic and the extreme class imbalance in modern datasets. This study proposes an optimised and lightweight stacking ensemble model for multiclass IoT cyberattack classification using the CICIoT2023 dataset, which comprises 712,311 network-flow records covering 37 attack and benign traffic categories. The proposed hybrid architecture combines Gradient Boosting and Random Forest as base learners, with Logistic Regression as a meta-classifier to enhance generalisation and stability. The model is tested against popular baseline classifiers, including Random Forest, XGBoost, and Logistic Regression, which achieve accuracies of 78.0%, 76.8%, and 74.8%, respectively. According to experimental data, the proposed stacking ensemble outperforms all baselines in terms of accuracy and weighted performance measures, achieving an accuracy of 81.3%. According to experimental data, the proposed stacking ensemble outperforms all baselines in terms of accuracy and weighted performance measures, achieving an accuracy of 81.3%. The results show that in resource-constrained IoT and edge environments, lightweight stacking-based ensemble learning offers a feasible and efficient approach for large-scale multiclass IoT intrusion detection.*
Keywords: Internet of Things security; intrusion detection systems; ensemble learning; stacking classifier; CICIoT2023 dataset; Machine Learning.

I. INTRODUCTION

The Internet of Things (IoT), which connects billions of diverse devices in the fields of healthcare, transportation, smart homes, industrial automation, and critical infrastructure, has transformed contemporary digital ecosystems due to its rapid global adoption [1]. IoT networks are vulnerable to a variety of cyberthreats, such as botnets, malware insertion, reconnaissance

attacks, and massive Distributed Denial-of-Service (DDoS) attacks, even though this hyper-connectivity allows for previously unheard-of automation and data-driven decision-making [2]. Intrusion detection is a crucial security need for protecting IoT infrastructure since IoT devices are still extremely susceptible to intrusions due to their limited computational capabilities, shoddy authentication procedures, and varied communication protocols [3].

Traditional signature-based intrusion detection systems (IDS) have proven insufficient in IoT environments due to their inability to detect novel and evolving threat variants. As a result, machine learning (ML), which can analyze traffic patterns and spot unusual behaviors in real-time, has become a feasible alternative. According to recent research, ML-based IDS can significantly improve detection accuracy and flexibility under various network settings [2]. Nevertheless, a number of difficulties still exist in spite of these developments. First, single machine learning classifiers find it challenging to generalise across a variety of attack categories due to the significant unpredictability and nonlinearity of IoT traffic [4]. Second, there is a significant class imbalance in many popular IoT datasets, like N-BaIoT, NSL-KDD, and Bot-IoT, which makes it difficult to identify minority attack classes [5]. Third, despite their accuracy, deep learning-based IDS models frequently need a lot of processing power, making them unsuitable for real-time deployment in IoT contexts with limited resources [6].

Due to its capacity to leverage the advantages of several base learners to achieve better performance than individual models, ensemble learning—especially stacking—has drawn attention as a solution to these constraints. Ensemble-based intrusion detection systems proved significant gains in attack categorisation, stability, and generalisation, particularly in intricate IoT environments. Rawashdeh et al., suggested a stacked ensemble for classifying IoT traffic and found that it was significantly more accurate than conventional machine learning classifiers [2]. Similarly, Maodah et al. showed the resilience of ensemble fusion approaches for cybersecurity applications by merging stacking and voting ensembles, which greatly improves intrusion detection in cloud-based systems [7]. Additionally, ensemble techniques have been successful in identifying botnets like Mirai and BASHLITE, particularly when hybrid models integrate boosting and tree-based algorithms [4].

The emergence of contemporary datasets, like CICIoT2023, which contains 37 attack categories and large-scale realistic IoT traffic, necessitates optimised and computationally efficient models capable of handling severe imbalance and high dimensionality, despite the fact that previous research has focused primarily on binary or limited multiclass classification. The deployment of many current solutions on edge and IoT platforms is compromised by their heavy reliance on deep learning. In order to obtain reliable multiclass classification while preserving computing economy, this work proposes a lightweight and optimised stacking ensemble model that integrates Random Forest, Gradient Boosting, and Logistic Regression. Our objective is to close the gap between IoT intrusion detection performance accuracy and resource efficiency.

II. RELATED WORK

Machine learning and ensemble algorithms can model nonlinear traffic patterns and identify a variety of cyberattacks, they have been thoroughly investigated for IoT and network intrusion detection. Traditional classifiers like SVM, KNN, Decision Trees, and Naïve Bayes were employed in early ML-based IDS techniques to identify malicious activity. These techniques showed respectable performance, but when used with high-dimensional and unbalanced IoT datasets, their capacity for generalisation remained constrained [4]. Classical machine learning techniques including SVM, Naïve Bayes, KNN, Decision Trees, and ANN were used in early IoT intrusion detection studies. These models performed notably well on simpler datasets, but when they were applied to real-world IoT systems with complicated nonlinear behaviours and diverse traffic, their performance declined. The shift towards multi-model ensemble techniques

has been prompted by research on ensemble-based IDS for IoT, which confirms that standalone ML models frequently struggle to generalise well across varied threat types [8].

A. **Classical and Hybrid Ensemble Models**

Classical ensemble techniques, such as bagging and boosting, have been demonstrated in numerous tests to greatly improve IDS performance. BoostedEnML developed an ensemble boosting-based method that used optimised boosting algorithms designed for IoT traffic patterns to improve cyberattack detection [3]. In IoT network intrusion scenarios, gradient boosting-based models have also been shown to perform better than standard classifiers, mainly because of their capacity to capture intricate feature interactions and minimise bias [9]

Multiple algorithms are used in hybrid ensemble models, which have demonstrated immense promise. For instance, the XGB-RF hybrid model achieved remarkably high accuracy for IoT botnet detection by integrating Random Forest for feature selection and XGBoost for classification [4]. Additionally, a number of ensemble-based intrusion detection systems have demonstrated better stability and robustness than standalone models, particularly when handling dynamic attack patterns and noisy traffic samples.

B. **Stacking-Based Ensemble Approaches**

Stacked ensembles have become a potent intrusion detection method. Maodah et al., demonstrated that stacked ensembles, which successfully combine a variety of classifiers to improve model stability and lower false alarm rates, outperform individual learners and basic voting methods in cloud environments [7]. Rawashdeh et al. demonstrated the adaptability of stacking frameworks in IoT security by applying stacking to IoT network traffic and reporting enhanced classification of complicated attack patterns [2].

Similar to this, layered deep neural models offered reliable detection of medical IoT assaults in stacked ensemble deep learning (SE-DL) architectures, despite their high computational cost [6]. In a different study, Vishwakarma and Kesswani introduced StaEn-IDS, an explainable stacking ensemble IDS that combines DNN, RF, and SVM to support transparent decision-making in IoT networks. The model's complexity limits its applicability in resource-constrained settings, despite its strong performance [10].

C. **Multiclass IoT Intrusion Detection Studies**

The majority of IDS research has historically concentrated on coarse multiclass grouping or binary classification (benign vs. assault). However, a far wider range of dangers affects contemporary IoT devices. ML classifiers regularly struggle to detect minority attacks due to significant class imbalance, a common problem in IoT datasets like Bot-IoT, UNSW-NB15, N-BaIoT, and others, according to a recent thorough study on multiclass IoT attack detection [5]. IDS dependability is eventually impacted by this imbalance, which lowers recall for uncommon but crucial threat types like SQL injection, XSS, malware, and data exfiltration attacks.

D. **Research Gap Identified**

Despite significant advancements, several gaps still exist:

1. Real-time deployment in IoT systems is limited by the computational burden of many IDS models, particularly deep learning ensembles.
2. Most studies rely on earlier datasets like NSL-KDD, UNSW-NB15, Bot-IoT, or N-BaIoT; multiclass IoT intrusion detection using large-scale datasets like CICIoT2023 is understudied.
3. Rare attack classes like malware, XSS, SQL injection, and uploading attacks are poorly detected due to inadequate imbalance management.
4. Few studies assess lightweight stacking ensembles that strike a balance between efficiency and performance.

The optimised hybrid stacking model suggested in this work is justified by these gaps.

III. Dataset Description

The CICIoT2023 dataset, a comprehensive and practical benchmark created to assess intrusion detection systems in IoT contexts, is used in this work. CICIoT2023, developed by the Canadian Institute for Cybersecurity, captures heterogeneous IoT network traffic produced by numerous IoT devices under a range of attack and benign scenarios. It is one of the most complete IoT intrusion detection datasets available, containing both packet-level and flow-level information taken from real-time traffic.

The dataset includes 40 features that describe statistical, behavioural, and protocol-level characteristics and 712,311 occurrences of network flow. These features enable precise modelling of IoT communication behaviour by capturing crucial attributes including connection duration, byte counts, packet rates, flow directionality, and flag patterns. The dataset has 37 different classes that contain both benign traffic and a variety of attack categories, including:

- DDoS attacks: UDP Flood, TCP Flood, SYN Flood, PSH-ACK Flood, RST-FIN Flood, HTTP Flood, Slowloris
- Fragmentation-based attacks: ICMP Fragmentation, UDP Fragmentation, ACK Fragmentation
- Mirai botnet variants: Mirai-GREETH Flood, Mirai-GREIP Flood, Mirai-UDPPlain
- Reconnaissance attacks: Host Discovery, Port Scan, Ping Sweep, OS Scan
- Application-level attacks: SQL Injection, XSS, Uploading Attacks, Dictionary Brute Force
- Other critical threats: Backdoor Malware, Browser Hijacking, MITM-ARPSpoofing

The extremely unequal distribution of classes in CICIoT2023 is one of its distinguishing features. While minority classes like Uploading Attack, Backdoor Malware, and SQL Injection have fewer than 20 instances, some attack categories, like DDOS-ICMP_FLOOD and DDOS-RSTFINFLOOD, have tens of thousands of samples. Machine learning algorithms are severely hampered by this imbalance since classifiers often overfit dominant classes and are unable to correctly detect low-frequency but high-risk attacks. This imbalance is confirmed by recent studies on multiclass IoT intrusion detection as a significant barrier to successful threat identification across several categories [5].

The remaining dataset was standardised using z-score normalisation after incorrect entries, such as infinite values and missing data, were eliminated to guarantee consistency and prevent data leakage. To maintain class distribution throughout model training and assessment, a stratified 80/20 train–test split was used.

All things considered, CICIoT2023 offers a solid framework for assessing lightweight and optimised ensemble learning models in actual IoT intrusion scenarios, particularly in multiclass classification tasks, where conventional IDS systems encounter major difficulties.

IV. Methodology

The proposed approach utilises a lightweight and efficient stacking ensemble model to classify 37 types of IoT cyberattacks from the CICIoT2023 dataset. Data preparation, feature scaling, base learner training, and meta-learning with a stacking ensemble are the four main parts of the methodological workflow. The method's focus on computing efficiency makes it appropriate for implementation in IoT scenarios with limited resources.

A. **Data Preprocessing**

The CICIoT2023 dataset contains 40 numerical features and 712,311 network-flow instances. Several preprocessing techniques were used before the model was trained in order to guarantee the quality of the data and avoid inconsistent behaviour during inference.

1. Handling Invalid Values**:**
   To preserve the integrity of the dataset, instances with infinite or missing values were eliminated.

2. Feature–Target Separation**:**
   The input space is made up of the remaining features, while the target variable Label represents one of 37 classes.

3. Stratified Train–Test Split**:**
   An 80/20 stratified split was employed to maintain the characteristics of class imbalance in both categories. To prevent over-representation of minority classes in either split, stratification is particularly important in severely imbalanced datasets.

B **Feature Scaling**

The StandardScaler was used to apply z-score normalisation due to the base learners' sensitivity to feature magnitude, especially Logistic Regression. Scaling speeds up convergence during model training and guarantees numerical stability.

C. **Stacking Ensemble Architecture**

The proposed intrusion detection model uses a two-level stacking ensemble with Logistic Regression (LR) acting as the meta-classifier and Random Forest (RF) and Gradient Boosting (GB) as base learners. By utilising the complementary qualities of several models while maintaining computing efficiency, stacking improves prediction performance.

1. Base Learners

- **Random Forest (RF)**
  - *n_estimators = 50*, *max_depth = 10*
  - Uses several randomised decision trees to capture nonlinear interactions.
  - Bootstrap aggregation keeps overfitting at bay and is robust to noisy data.
- **Gradient Boosting (GB)**
  - *n_estimators = 50*, *learning_rate = 0.1*, *max_depth = 3*
  - sequentially reduces classification errors by focusing on difficult-to-learn instances.
  - useful for datasets with intricate decision boundaries.

While maintaining good learning potential, these lightweight setups (each with 50 trees) drastically reduce computational effort.

2. **Meta-Learner (Level-2 Classifier)**

- **Logistic Regression (LR)** with *max_iter = 500*

- Combines the predictions of base learners.
- Provides a linear, low-complexity decision boundary.
- Ensures the ensemble remains lightweight and fast during inference.

Stacking is implemented without passthrough to minimize feature expansion and memory usage:
To improve generalisation and decrease overfitting, a three-fold cross-validation technique was applied within the stacking layer. The preprocessing phase, parallel base learners, and meta-classification layer are highlighted in Figure 1, which depicts the architecture of the proposed RF–GB–LR stacking ensemble. The preprocessing phase, parallel base learners, and meta-classification layer are highlighted in Figure 1, which depicts the architecture of the proposed RF–GB–LR stacking ensemble.

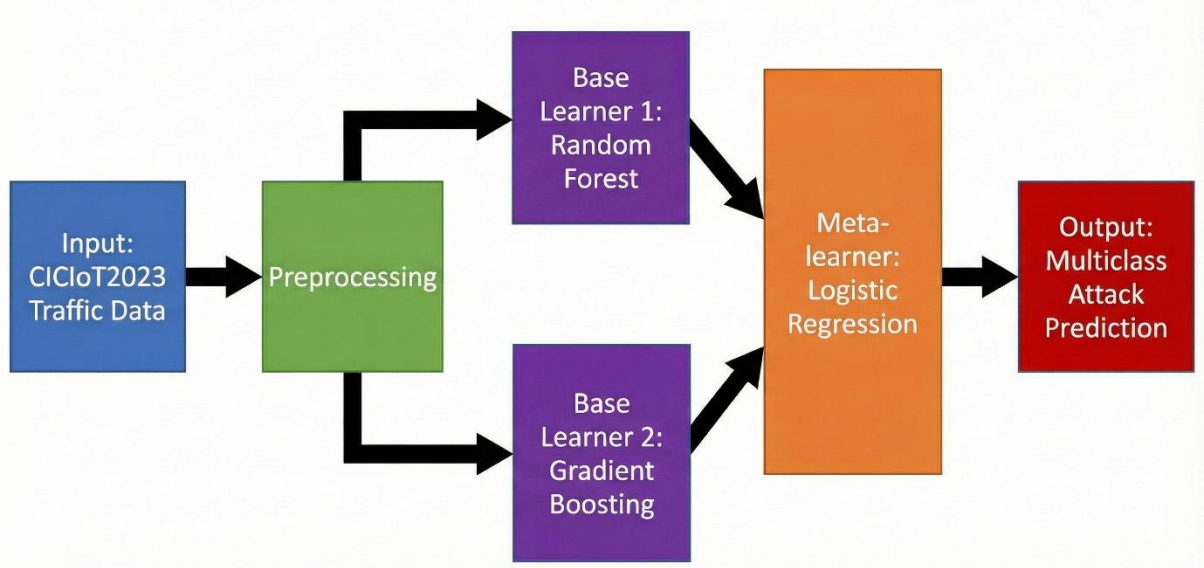

Figure 1. Architecture of the proposed RF–GB–LR stacking ensemble for multiclass IoT intrusion detection.

D. **Model Training and Testing**

The scaled training set was utilised to train the stacking ensemble. The model was assessed using the held-out test set (142,460 samples) after it had been optimised. Performance in prediction was evaluated using:

- Accuracy
- Precision
- Recall
- F1-score
- Per-class evaluation metrics

In multiclass IDS tasks, these metrics are crucial, particularly when class imbalance affects the accuracy of minority attack detection.

F. Baseline Classifiers

Three baseline classifiers—Logistic Regression, Random Forest, and XGBoost—were used to compare the performance of the proposed stacking ensemble in order to verify its efficacy. To evaluate the shortcomings of linear decision boundaries in complicated IoT traffic, logistic regression was added as a linear baseline model. While XGBoost is a powerful boosting-based classifier that is frequently utilised in IDS, Random Forest is a nonlinear ensemble model based on bagging.

To ensure a fair and consistent comparison, all baseline models were trained and assessed using the same preprocessing pipeline, feature scaling technique, and data partitioning as proposed. Because of the significant class imbalance in the CICIoT2023 dataset, performance was evaluated using weighted evaluation measures.

In line with findings from earlier ensemble IDS research, the proposed Hybrid RF–GB–LR stack offers an efficient balance of accuracy, speed, and resource consumption as compared to deep learning stacking architectures that demand significant GPU resources.

V. **Results and Discussion**

The proposed hybrid stacking ensemble model is thoroughly evaluated on the CICIoT2023 dataset in this part, along with an in-depth assessment of how effectively it performed in contrast to baseline classifiers. In order to account for the extremely unbalanced and multiclass nature of IoT traffic, the model is evaluated using accuracy, precision, recall, and F1-score, with a focus on weighted metrics and per-class behaviour. This assessment demonstrates how well the proposed approach handles the vast and intricate threat landscape found in IoT systems.

A. **Overall Performance**

With an overall accuracy of 81.3%, the proposed stacking ensemble showed a remarkable capacity to learn intricate and diverse traffic patterns from the extensive CICIoT2023 dataset as Table 1and Figure 2 illustrate. Given the extreme class imbalance of the dataset—many attack types have less than 20 samples—this performance level is very noteworthy. By combining complementary decision boundaries via a meta-learning framework, the stacking ensemble demonstrates improved generalisation in comparison to individual base learners, such as Random Forest and Gradient Boosting. This result is in line with earlier research on stacking-based intrusion detection, which found that in complex IoT security contexts, integrating heterogeneous classifiers lowers variance and increases predictive stability [2].

Table 1: Performance Comparison of Baseline and Proposed Models

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-score (Weighted)
Logistic Regression (Linear Baseline)	0.748	0.74	0.75	0.71
Random Forest (Baseline)	0.768	0.77	0.77	0.76
XGBoost (Baseline)	0.780	0.78	0.78	0.77
Proposed RF–GB–LR Stacking Ensemble	**0.813**	**0.82**	**0.81**	**0.80**

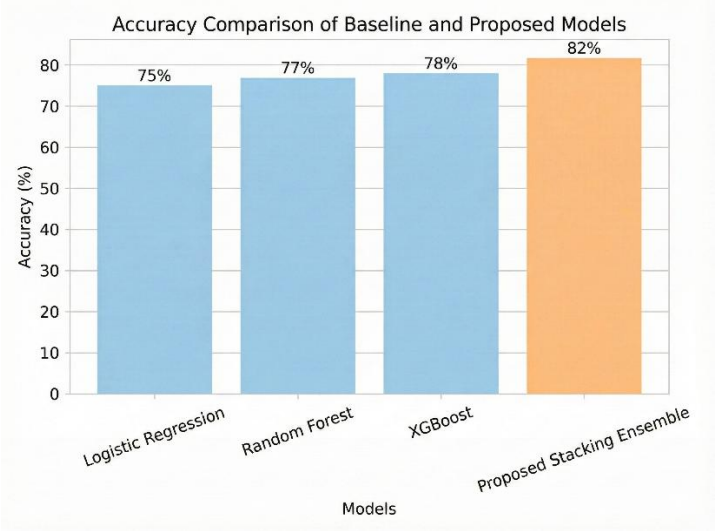

Figure 2: Accuracy comparison of baseline classifiers and the proposed stacking ensemble on the CICIoT2023 dataset.

The advantage of ensemble fusing in challenging IoT intrusion detection tasks is confirmed by the findings, which unequivocally demonstrate that the suggested stacking ensemble performs better than the baseline classifiers across all reported metrics

B. Analysis of Baseline Classifier Behavior

The Logistic Regression baseline showed the poorest performance compared to all other models tested. Although it performed well in identifying common classes like benign traffic and volumetric DDoS attacks, it struggled to detect nonlinear and rare types of attacks. This issue is a natural limitation of linear classifiers when dealing with the diverse and complex nature of IoT traffic, which involves intricate interactions between different features.

Random Forest showed enhanced performance by using ensemble decision trees to capture nonlinear interactions. It demonstrated lower recall for reconnaissance and application-layer assaults with few training samples, but it achieved great detection rates for high-volume DDoS and Mirai threat categories. Similarly, by utilising boosting-based optimisation, XGBoost considerably enhanced overall performance; yet, it continued to struggle with severely rare classes like SQL Injection, XSS, Uploading Attacks, and Backdoor Malware.

These findings reveal that although standalone nonlinear and boosting-based models perform well, they are still susceptible to class imbalance and are unable to adequately handle the variety of IoT attack behaviours.

C. Effectiveness of the Proposed Stacking Ensemble

The superiority of the proposed RF–GB–LR stacking ensemble to incorporate complementary learning behaviours is responsible for its improved performance. Gradient Boosting improves discrimination by concentrating on incorrectly categorised samples, Random Forest adds robustness to noisy and high-dimensional features, and Logistic Regression functions as a lightweight meta-learner that best integrates the predictions of basic models.

Compared to individual classifiers, the stacking ensemble can achieve better multiclass discrimination and more stable decision limits thanks to this integration. Notably, the ensemble reduces misclassification across dominating attack categories without appreciably raising computational complexity, as evidenced by improvements in total weighted precision and F1-score.

D. Per-Class Performance Analysis

The proposed stacking ensemble offers near-perfect detection performance for a number of high-support IoT threat categories, according to a thorough analysis of the classification

findings. Specifically, volumetric DDoS and botnet-related assaults, like DDOS-ICMP_FLOOD, DDOS-RSTFINFLOOD, and DDOS-PSHACK_FLOOD, show recall and F1-scores that are close to 1.00, and MIRAI-GREETH_FLOOD and DDOS-UDP_FLOOD also routinely achieve high detection rates. These findings show that the model well captures the steady traffic structures, high-frequency attack signatures, and recurring flow patterns typical of extensive DDoS and botnet activity. Similar findings, where the majority of classes profit from a large number of training samples and clearly defined statistical properties, have been documented in earlier IoT intrusion detection analyses by Alotaibi et., [3].

The stacking ensemble obtains respectable but significantly reduced recall and F1-scores for attack categories with substantial sample support, such as DNS Spoofing, MITM-ARPSpoofing, and DOS-HTTP_FLOOD. This behaviour implies that despite the model does a good job of generalising across a variety of attack types, it occasionally has difficulties with traffic patterns that statistically overlap with benign behaviour or other attack classes.

Minority attack types, on the other hand, show nearly zero recall and F1-scores. These include Backdoor Malware, SQL Injection, Uploading Attacks, Browser Hijacking, and Recon-PingSweep. Given the extreme class imbalance in the CICIoT2023 dataset—some classes have fewer than 20 samples—this result is anticipated. Robust decision boundaries cannot be learnt by the classifier with such minimal representation. These results low detection rates for low-frequency attack classes in unbalanced environments.

E. Comparison with Existing Ensemble-Based IDS Approaches

The proposed stacking model exhibits competitive and, in some cases, better performance when compared to current ensemble-based intrusion detection techniques. The enhanced detection accuracy reported by earlier stacking-based IDS researches was limited to binary or small-scale multiclass datasets, making them unsuitable for real IoT contexts. Although they performed well on botnet-focused datasets, boosting-based frameworks like BoostedEnML were not tested on extensive multiclass benchmarks like CICIoT2023. Similar to this, stacking-based deep learning models demonstrated great accuracy but were not appropriate for deployment on IoT or edge devices with limited resources due to their reliance on computationally demanding architectures.

On the other hand, a large-scale 37-class IoT dataset is used to assess the proposed stacking ensemble, which demonstrates its appropriateness for real-world IoT security deployments by achieving strong multiclass performance and maintaining low computational complexity. The majority of current ensemble-based intrusion detection techniques use computationally demanding deep learning architectures or rely on binary or small-scale multiclass datasets to attain high accuracy, as seen in Table 2. On the other hand, the proposed stacking ensemble retains competitive performance while being lightweight and computationally economical when tested on the extensive CICIoT2023 dataset, which includes 37 attack categories. This demonstrates how the proposed approach is appropriate for real-world IoT and edge situations, where resource limitations restrict the use of deep learning-based models.

Table 2: Comparison of the Proposed Method with Existing IoT Intrusion Detection Studies

Study	Dataset	No. of Classes	Methodology	Performance	Key Limitations
[2]	IoT Traffic Dataset	Binary / Small Multiclass	Stacking Ensemble (ML)	>95% Accuracy	Evaluated on smaller-scale datasets
[3] (BoostedEnML)	Bot-IoT	Binary / Few Multiclass	Boosted Ensemble ML	>99% Accuracy	Focused mainly on botnet attacks

Study	Dataset	No. of Classes	Methodology	Performance	Key Limitations
[11] (StaEn-IDS)	IoT Dataset	Multiclass	Deep Learning Stacking	>98% Accuracy	High computational cost
[10]	IoMT Dataset	Multiclass	DL-based Ensemble	>97% Accuracy	Not suitable for edge devices
Proposed Method	**CICIoT2023**	**37**	**Lightweight Stacking (RF–GB–LR)**	**81.3% Accuracy**	Minority class imbalance

F. Discussion of Model Strengths

The proposed RF–GB–LR stacking ensemble has a number of significant advantages. First, it uses Random Forest's resilience to noisy and high-dimensional data and Gradient Boosting's potential to refine misclassified instances through sequential learning to achieve high detection accuracy for the main DDoS and Mirai assault categories. Second, CPU-based execution with minimal memory cost and quick inference times is made possible by the architecture's continued lightweight and computational efficiency. Third, with Random Forest capturing nonlinear relationships, Gradient Boosting modelling intricate residual patterns, and Logistic Regression efficiently aggregating base learner outputs through a low-complexity meta-learning layer, the stacking framework guarantees the seamless integration of heterogeneous learners. In light of these features, the proposed method is especially appropriate for real-world IoT and edge contexts where real-time detection is necessary, and computational resources are constrained.

VI. Conclusion and Future Work

This research employed the CICIoT2023 dataset to offer a lightweight and optimised stacking ensemble for multiclass IoT intrusion detection. The proposed technique effectively captured complex and nonlinear attack patterns while preserving minimal computational overhead appropriate for IoTs and edge environments by combining Random Forest and Gradient Boosting as base learners with a Logistic Regression meta-classifier. According to experimental data, the suggested stacking ensemble outperformed all evaluated baseline classifiers, such as Logistic Regression, Random Forest, and XGBoost, with an overall accuracy of 81.3%. In extremely unbalanced multiclass intrusion detection scenarios, the ensemble consistently produced greater weighted precision, recall, and F1-score, demonstrating the superiority of ensemble fusion over solo learning models.

Regardless of the substantial class imbalance present in the dataset, the suggested model shows limited detection capability for extremely low-frequency attack categories despite its strong overall performance. Therefore, in order to enhance minority-class detection, future research will concentrate on using imbalance-mitigation tools, including SMOTE, cost-sensitive learning, and data augmentation techniques. Furthermore, the utilisation of hybrid ensemble architectures and lightweight deep learning will be explored to improve feature representation while maintaining computational efficiency. Finally, in order to evaluate scalability, latency, and practical viability under realistic operating settings, real-time deployment and assessment on edge-based IoT platforms will be examined.

References

[1] M. Koca and I. Avci, “A Novel Hybrid Model Detection of Security Vulnerabilities in Industrial Control Systems and IoT Using GCN+LSTM,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3466391.

[2] "A stacked ensemble approach to identify internet of things network attacks through traffic analysis _ Rawashdeh _ Bulletin of Electrical Engineering and Informatics".
[3] Y. Alotaibi and M. Ilyas, "Ensemble-Learning Framework for Intrusion Detection to Enhance Internet of Things' Devices Security," *Sensors*, vol. 23, no. 12, Jun. 2023, doi: 10.3390/s23125568.
[4] J. Al Faysal *et al.*, "XGB-RF: A Hybrid Machine Learning Approach for IoT Intrusion Detection," *Telecom*, vol. 3, no. 1, pp. 52–69, Mar. 2022, doi: 10.3390/telecom3010003.
[5] "View of Comprehensive Study on Detecting Multi-Class Classification of IoT Attack Using Machine Learning Methods".
[6] E. Alalwany, B. Alsharif, Y. Alotaibi, A. Alfahaid, I. Mahgoub, and M. Ilyas, "Stacking Ensemble Deep Learning for Real-Time Intrusion Detection in IoMT Environments," *Sensors*, vol. 25, no. 3, Feb. 2025, doi: 10.3390/s25030624.
[7] K. A. Maodah, S. Alhomdy, and F. Thabit, "Detecting intrusions in cloud-based ensembles: evaluating voting and stacking methods with machine learning classifiers," *Front Comput Sci*, vol. 7, 2025, doi: 10.3389/fcomp.2025.1623375.
[8] A. Abbas, M. A. Khan, S. Latif, M. Ajaz, A. A. Shah, and J. Ahmad, "A New Ensemble-Based Intrusion Detection System for Internet of Things," *Arab J Sci Eng*, vol. 47, no. 2, pp. 1805–1819, Feb. 2022, doi: 10.1007/s13369-021-06086-5.
[9] M. A. O. Ahmed, Y. AbdelSatar, R. Alotaibi, and O. Reyad, "Enhancing Internet of Things security using performance gradient boosting for network intrusion detection systems," *Alexandria Engineering Journal*, vol. 116, pp. 472–482, Mar. 2025, doi: 10.1016/j.aej.2024.12.106.
[10] M. Vishwakarma and N. Kesswani, "StaEn-IDS: An Explainable Stacking Ensemble Deep Neural Network-Based Intrusion Detection System for IoT," *IEEE Access*, vol. 13, pp. 109713–109728, 2025, doi: 10.1109/ACCESS.2025.3582391.
[11] R. Lazzarini, H. Tianfield, and V. Charissis, "A stacking ensemble of deep learning models for IoT intrusion detection," *Knowl Based Syst*, vol. 279, Nov. 2023, doi: 10.1016/j.knosys.2023.110941.