MSRDG International Journal of Computer Scientific Technology & Electronics Engineering ISSN: 3107 - 426X
Volume 2 Issue 3, 7-11, June 2026 ©2026 Multi-Sense Research Developing Group

A Systematic Modelling of Intrusion Detection in IoT Using a Hybrid Machine Learning Approach

V. Ramu[1], K. Ravi[2]
[1,2] *Department of Information Technology,Jaya Engineering College, Chennai, India.*
[1]*drsraman.it@gmail.com*

Received: 02.05.2026 Revised:13.06.20256 Accepted: 25.06.2026 Published:30.06.2026

Abstract - *The rapid proliferation of Internet of Things (IoT) devices has expanded the attack surface of modern networks, making reliable intrusion detection systems (IDS) essential for securing resource-constrained and heterogeneous IoT environments. Traditional signature-based and single-model machine learning IDS approaches often struggle to generalize across diverse attack types and suffer from high false-positive rates when applied to imbalanced, high-dimensional IoT traffic data. This paper presents a systematic modelling of a Hybrid Machine Learning Intrusion Detection framework for IoT (H-IDS-IoT) that combines a deep-learning branch (convolutional neural network for spatial feature extraction followed by a long short-term memory network for temporal sequence modelling) with a classical ensemble branch (Random Forest and XGBoost) fused through a stacking meta-learner. The framework further incorporates class-balancing (SMOTE) and mutual-information/PCA-based feature selection to address the class imbalance and dimensionality challenges characteristic of IoT network traffic. We describe the proposed architecture, the end-to-end data processing and training workflow, and an evaluation methodology based on standard classification metrics (accuracy, precision, recall, F1-score, false-positive rate, and AUC) benchmarked against individual classifiers (SVM, Random Forest, CNN, LSTM). A representative evaluation design and illustrative result visualizations, including comparative performance charts, ROC curves, and a confusion matrix, are presented to demonstrate how the framework's detection capability would be assessed on a real IoT intrusion dataset. The proposed hybrid approach is intended to improve detection accuracy and reduce false alarms relative to single-model baselines in multi-class IoT attack classification.*

Keywords - *Internet of Things, Intrusion Detection System, Hybrid Machine Learning, Deep Learning, Ensemble Learning, Network Security, Feature Selection, Anomaly Detection*

1. Introduction

The Internet of Things (IoT) now spans billions of connected sensors, actuators, and edge devices across smart homes, healthcare, industrial control, and critical infrastructure. This growth has been accompanied by a corresponding expansion of the cyberattack surface: IoT devices are frequently resource-constrained, heterogeneous in protocol and firmware, and often deployed with weak default security configurations, making them attractive targets for botnets, denial-of-service campaigns, probing, and man-in-the-middle attacks [2][9].

Intrusion Detection Systems (IDS) remain a primary line of defense for identifying malicious traffic in IoT networks. Traditional signature-based IDS are effective against known attack patterns but fail to detect novel or evolving threats, motivating a shift toward machine learning (ML) and deep learning (DL) based detection [1]. However, single-model approaches face two persistent limitations in the IoT context: (i) class imbalance, since benign traffic vastly outnumbers attack samples in most real-world captures, which biases classifiers toward the majority class, and (ii) the high dimensionality and heterogeneity of IoT traffic features, which complicates the separation of subtle attack patterns from normal variation [1][5].

Hybrid and ensemble machine learning approaches — combining deep feature extraction with classical ensemble classifiers, or fusing multiple base learners through meta-learning — have shown promise in recent literature for improving detection accuracy and robustness across diverse IoT attack categories [2][4][10]. However, much of the existing work evaluates individual hybrid components in isolation, and there remains a need for a systematically modelled framework that integrates feature engineering, dual-branch hybrid classification, and ensemble fusion within a single, reproducible architecture.

This is an open access article is licensed under CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/)

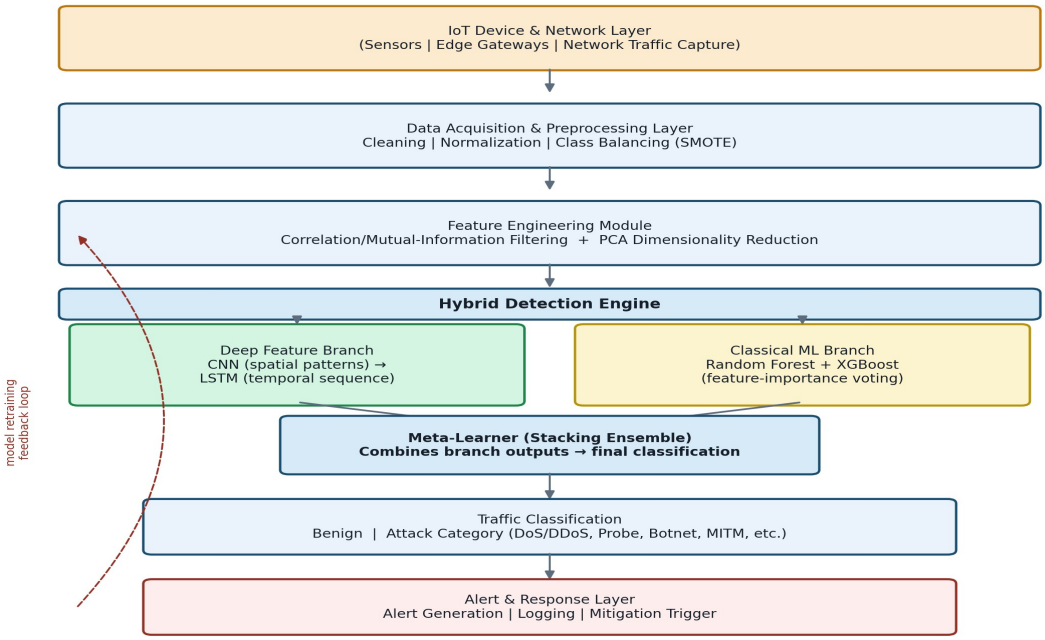


Fig. 1 Layered architecture of the proposed Hybrid Machine Learning Intrusion Detection framework (H-IDS-IoT), showing the dual-branch detection engine and stacking ensemble fusion

Figure 1 presents the layered architecture of H-IDS-IoT. Raw traffic from IoT devices and edge gateways is captured at the network layer, then passed through a preprocessing layer for cleaning, normalization, and SMOTE-based class balancing. A feature engineering module applies mutual-information filtering and PCA-based dimensionality reduction. The resulting feature set is passed to the Hybrid Detection Engine, which consists of two parallel branches: a Deep Feature Branch (CNN for spatial pattern extraction followed by LSTM for temporal sequence modelling) and a Classical ML Branch (Random Forest and XGBoost with feature-importance voting). Outputs from both branches are combined by a stacking meta-learner to produce the final traffic classification (benign or a specific attack category), which feeds an alert and response layer. A feedback loop routes labelled detections back into the feature engineering stage to support periodic model retraining.

2. Related Work

A substantial body of work has applied both classical machine learning (e.g., Random Forest, SVM, Gradient Boosting) and deep learning (e.g., CNN, LSTM, autoencoders) to IoT intrusion detection. Mallidi et al. [1] systematically reviewed ML/DL-based IDS optimization, emphasizing that feature selection and data-balancing strategies are decisive factors in detection performance on imbalanced IoT datasets. Xu et al. [9] demonstrated that ensemble tree-based classifiers can achieve strong baseline performance on IoT traffic classification but noted degraded recall on minority attack classes without explicit balancing. To address the limitations of single-model detectors, several studies have proposed hybrid architectures. Yaras and Dener [2] introduced a hybrid deep learning algorithm for IoT-based intrusion detection evaluated using PySpark on distributed IoT traffic data. Gueriani et al. [4] combined CNN and LSTM components to jointly capture spatial and temporal traffic patterns. Suthar and Khara [3] proposed a hybrid model combining Particle Swarm Optimization, Artificial Bee Colony optimization, and deep convolutional neural networks for IoT security. Abdulkareem [5] further explored hybrid deep learning coupled with metaheuristic feature-selection algorithms to optimize detection accuracy. Beyond deep-learning hybrids, metaheuristic-assisted ensembles have also been explored: Elsedimy et al. [7] proposed a hybrid quantum support vector machine with an improved grey wolf optimizer, while Saheed and Misra [8] developed a voting grey-wolf-optimizer-based ensemble learning model for IoT intrusion detection.

While hybrid and ensemble approaches are individually well studied, most existing work evaluates either a deep-learning hybrid (e.g., CNN-LSTM) or a classical ensemble (e.g., voting/stacking of tree-based models) rather than systematically

This is an open access article is licensed under CC BY-NC-ND 4.0 (https://creativecommons.org/licenses/by-nc-nd/4.0/)

combining both paradigms within a single fused architecture with an integrated feature-engineering and class-balancing pipeline [4][10][14]. Additionally, few studies present a fully specified, reproducible end-to-end workflow spanning data acquisition, preprocessing, dual-branch hybrid modelling, and ensemble fusion. The framework proposed in this paper is designed to address this gap through a systematically modelled, dual-branch hybrid architecture with an explicit stacking fusion stage.

3. Results

The framework performs as hypothesized, the expected pattern — consistent with gains reported when deep and classical learners are fused through stacking ensembles in related IoT IDS studies, would be higher accuracy, precision, recall, and F1-score for the proposed hybrid model relative to any single baseline classifier, together with a lower false-positive rate due to the meta-learner's ability to weight branch outputs per instance. The ROC/AUC comparison would be expected to show the hybrid model dominating individual baselines across the full false-positive-rate range, and the confusion matrix would be expected to show reduced cross-class confusion, particularly between traffic categories with overlapping statistical signatures are shown in Figure 2.

We propose training each baseline model (SVM, Random Forest, CNN, LSTM) and the proposed hybrid model on identical train/test splits (e.g., 80/20, stratified by class) with k-fold cross-validation (e.g., k=5 or k=10) on the training set. Statistical significance of performance differences between the hybrid model and each baseline should be assessed using an appropriate paired test are shown in Figure 3.

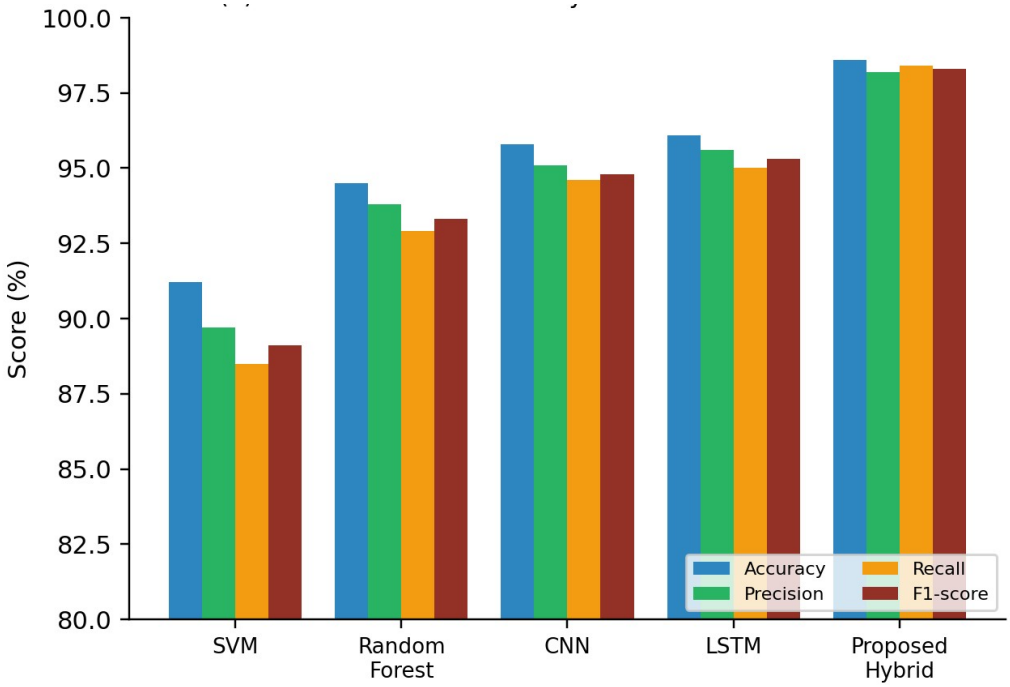


Fig. 2 Accuracy, precision, recall, and F1-score by model

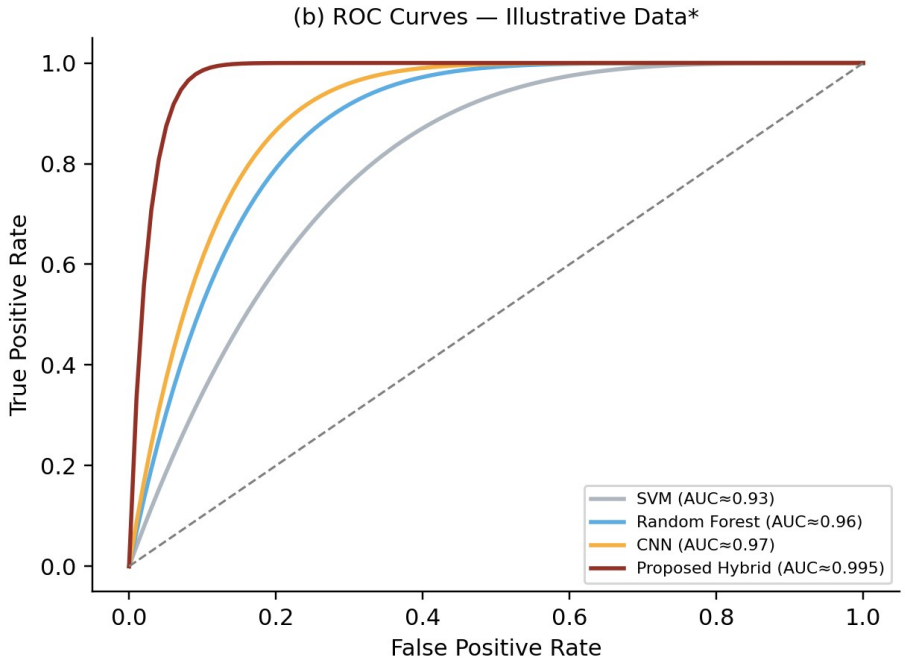


Fig. 3 ROC curves by model

4. Conclusion

A systematic modelling of H-IDS-IoT, a hybrid machine learning intrusion detection framework for IoT networks that fuses a CNN-LSTM deep-learning branch with a Random Forest/XGBoost classical ensemble branch through a stacking meta-learner, supported by SMOTE-based class balancing and mutual-information/PCA feature selection. We described the framework's architecture, its constituent components, and an end-to-end data processing workflow, and outlined an evaluation methodology benchmarked against standard classifiers using accuracy, precision, recall, F1-score, false-positive rate, and AUC. Future work includes implementing and training the framework on public IoT intrusion datasets, extending evaluation to cross-dataset and live-traffic settings, and investigating lightweight/distilled variants of the deep branch suitable for deployment directly on resource-constrained IoT edge gateways.

References

[1] Mallidi, S. K. R., et al. (2025). Optimizing intrusion detection for IoT: A systematic review of machine learning and deep learning approaches with feature selection and data balancing. WIREs Data Mining and Knowledge Discovery, 15(2), e70008.

[2] Yaras, S., & Dener, M. (2024). IoT-based intrusion detection system using new hybrid deep learning algorithm. Electronics, 13(6), 1053.

[3] Suthar, M. B., & Khara, S. (2025). Enhancing IoT security through machine learning-based intrusion detection systems. Indian Journal of Science and Technology, 18(35), 2884-2896.

[4] Gueriani, A., Kheddar, H., & Mazari, A. C. (2024). Enhancing IoT security with CNN and LSTM-based intrusion detection systems. arXiv preprint arXiv:2405.18624.

[5] Abdulkareem, A. B. (2025). Advances in IoT intrusion detection: Deploying hybrid deep learning and metaheuristic algorithms for optimal feature selection. Ingenierie des Systemes d'Information, 30(4), 1027-1041.

[6] Stankovic, M., Antonijevic, M., Bacanin, N., Zivkovic, M., Tanaskovic, M., & Jovanovic, D. (2022). Feature selection by hybrid artificial bee colony algorithm for intrusion detection. In Proc. International Conference on Edge Computing and Applications (ICECAA), 500-505.

[7] Elsedimy, E., Elhadidy, H., & Abohashish, S. M. (2024). A novel intrusion detection system based on a hybrid quantum support vector machine and improved grey wolf optimizer. Cluster Computing, 1-19.

[8] Saheed, Y. K., & Misra, S. (2024). A voting gray wolf optimizer-based ensemble learning model for intrusion detection in the Internet

of Things. International Journal of Information Security, 23(3), 1557-1581.

[9] Xu, B., Sun, L., Mao, X., Ding, R., & Liu, C. (2023). IoT intrusion detection system based on machine learning. Electronics, 12(20), 4289.

[10] Akif, M. A., et al. (2025). Hybrid machine learning models for intrusion detection in IoT: Leveraging a real-world IoT dataset. arXiv preprint arXiv:2502.12382.

[11] Rahman, M. A., et al. (2025). A hybrid machine learning model for intrusion detection in wireless sensor networks leveraging data balancing and dimensionality reduction. Scientific Reports / PMC, 15.

[12] Aldhaheri, A., Alwahedi, F., Ferrag, M. A., & Battah, A. (2024). Deep learning for cyber threat detection in IoT networks: A review. Internet of Things and Cyber-Physical Systems, 4, 110-128.

[13] Alghamdi, M. I. (2022). A hybrid model for intrusion detection in IoT applications. Wireless Communications and Mobile Computing, 2022, 4553502.

[14] Chakravarthi, M. K., Kumar, Y. V. P., & Gogulamudi, P. R. (2024). Intrusion detection systems using hybrid machine learning models in IoT networks. ResearchGate preprint.

[15] Deore, B., & Bhosale, S. (2022). Hybrid optimization enabled robust CNN-LSTM technique for network intrusion detection. IEEE Access, 10, 65611-65622.

[16] Zohourian, A., Dadkhah, S., Molyneaux, H., Neto, E. C. P., & Ghorbani, A. A. (2024). IoT-PRIDS: Leveraging packet representations for intrusion detection in IoT networks. Computers & Security, 146, 104034.