International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(2s) 2026 pp. 854-864
https://doi.org/10.70917/ijcisim-2026-2136
Article

A CALIBRATED HYBRID MACHINE-LEARNING AND DEEP-LEARNING ENSEMBLE FRAMEWORK FOR MULTICLASS DDOS INTRUSION DETECTION IN IOT AND IOMT ECOSYSTEMS

Khushboo Sharma[1], Ravi Shankar Sharma[2]

[1]Department of Computer Science and Engineering, khushboosharma2301@gmail.com
[2]Department of Computer Science and Engineering, er.ravishankarsharma@gmail.com
Corresponding Author: Khushboo Sharma (Email: khushboosharma2301@gmail.com)

Abstract: - The rapid proliferation of the Internet of Things (IoT) and its multimedia-rich extension, the Internet of Multimedia Things (IoMT), has dramatically enlarged the attack surface available to adversaries, with Distributed Denial-of-Service (DDoS) attacks remaining among the most damaging and frequent threats. Conventional signature-based and binary intrusion detection systems struggle with the heterogeneity, class imbalance, and non-stationarity of modern IoT traffic, and they frequently fail on rare but high-impact attack categories such as application-layer WebDDoS. This paper presents a unified, calibrated detection framework that combines classical machine learning (ML), deep learning (DL), and cost-aware ensemble learning for multiclass DDoS classification on a CICDDoS2019-style benchmark. The pipeline integrates mutual-information and recursive feature elimination, SMOTE-ENN balancing, four sequence-aware deep architectures (1D-CNN, CNN-BiLSTM, Temporal Convolutional Network, and a lightweight Transformer), and a cost-aware stacking ensemble with temperature scaling and per-class threshold tuning. Experimental results show that the cost-aware stack attains 99.39% accuracy, a macro-F1 of 0.97, a macro ROC-AUC of 1.00, and—most importantly—raises the WebDDoS F1 from 0.38 to 0.77 while keeping the benign false-positive rate at 0.016. The framework also achieves a low expected calibration error (0.014–0.017), making its probabilities trustworthy for security-operations-center thresholding. The study demonstrates that calibrated, imbalance-aware ensembling is essential for operational IoT/IoMT defense...

Keywords: DDoS detection, Internet of Things, IoMT, deep learning, ensemble learning, class imbalance, model calibration, intrusion detection systems

1. INTRODUCTION

The extraordinary rise of the Internet of Things (IoT) has redefined digital communication and automation by creating vast ecosystems of interconnected devices that sense, actuate, and exchange data with minimal human intervention. From smart cities and connected healthcare to industrial automation and intelligent transportation, IoT now underpins critical infrastructure on a global scale [1], [2]. This explosive growth, however, has expanded the cyber-attack surface enormously: billions of resource-constrained, heterogeneous, and often weakly secured nodes provide adversaries with abundant entry points [3].

Among the threats facing these ecosystems, Distributed Denial-of-Service (DDoS) attacks remain the most disruptive and economically damaging. By orchestrating massive volumes of traffic—often from compromised IoT devices forming botnets such as Mirai—attackers exhaust bandwidth, computation, and memory, rendering legitimate services unavailable [4], [5]. The emergence of the Internet of Multimedia Things (IoMT), which integrates audio, video, and rich sensor streams, magnifies both the value of the data and the consequences of service disruption [6], [7].

Traditional intrusion detection systems (IDS) rely on signature databases or coarse binary classification (benign vs. malicious). Such approaches degrade rapidly when confronted with the heterogeneity, volume, and non-stationarity of contemporary IoT traffic, and they offer little insight into the specific category of an attack [8], [9]. Multiclass detection—distinguishing benign flows from DNS, MSSQL, NTP, SYN, UDP-lag, SSDP, and application-layer WebDDoS attacks—is operationally essential because mitigation strategies differ sharply across categories [10].

A central and persistent difficulty is class imbalance. Reflection/amplification attacks dominate public benchmarks, while stealthy application-layer attacks such as WebDDoS appear in vanishingly small proportions. Detectors optimized for overall accuracy therefore tend to ignore exactly the rare classes that are most dangerous [11], [12]. This paper directly targets that gap.

The contributions of this work are fourfold: (i) a unified pipeline that fairly benchmarks eleven classical ML detectors, six deep architectures, and several ensembles under identical preprocessing; (ii) an imbalance-aware training regime combining SMOTE-ENN, focal loss, and per-class threshold tuning; (iii) a cost-aware stacking ensemble that substantially improves rare-class (WebDDoS) detection without inflating false positives; and (iv) a calibration analysis demonstrating that the proposed models yield trustworthy probabilities suitable for security-operations-center (SOC) decision thresholds.

2. RELATED WORK

The literature on IoT security and its multimedia offshoots has matured rapidly as healthcare, urban infrastructure, and industrial automation have adopted connected devices at scale [1], [6]. Early work emphasized lightweight cryptography and static, rule-based IDS, which provided baseline confidentiality but could not adapt to evolving traffic [8]. As IoMT crystallized into a distinct research stream, quality management evolved from throughput-oriented QoS toward human-perception-oriented QoE and domain-specific security abstractions [7], [13].

Machine-learning detectors emerged to address adaptability. Supervised methods such as support vector machines, decision trees, and random forests have been widely applied to the CICDDoS2019 benchmark, achieving strong aggregate accuracy [10], [14]. Zinca and Dobrota [14] demonstrated that carefully tuned supervised models remain competitive baselines, while gradient-boosting variants (XGBoost, LightGBM, CatBoost) further improve discrimination on tabular flow features [15], [16].

Deep learning has driven the most recent gains. Convolutional networks capture local spatial structure in flow-feature vectors, while recurrent and temporal models—LSTM, BiLSTM, and Temporal Convolutional Networks—exploit sequential dependencies in packet streams [17], [18], [19]. Hybrid CNN-GRU-DNN and CNN-LSTM models report high accuracy by jointly modeling spatial and temporal patterns [20], [21]. Attention-based and Transformer detectors have begun to outperform recurrent counterparts on long-context traffic [22], [23]. A recent bibliometric synthesis found that 2023 and 2024 each accounted for roughly 36% of all deep-learning DDoS publications, underscoring the field's acceleration [24].

Generative and reinforcement-learning approaches have also appeared: Strickland et al. combined GANs with deep reinforcement learning on CICDDoS2019 and reported 98.2% accuracy, though they cautioned about the fidelity of GAN-synthesized test traffic [25]. Federated and SDN-centric architectures address scalability and privacy by offloading training while keeping inference at the edge [26], [27].

Despite this progress, three gaps recur. First, single learners rarely capture the diversity of IoT attack modalities, motivating ensembles [28]. Second, most studies report aggregate accuracy and under-report rare-class behavior, especially WebDDoS [11], [12]. Third, probability calibration—critical for operational thresholding—is seldom analyzed [29], [30]. The present study addresses all three within one reproducible framework.

3. PROPOSED METHODOLOGY

A. System Architecture

The proposed framework follows a four-layer architecture that mirrors operational IoT/IoMT deployments. The perception/edge layer captures raw traffic from heterogeneous devices and exports flow records. The fog/pre-processing layer performs cleaning, normalization, encoding, feature engineering, and class balancing. The cloud/detection layer hosts the ML, DL, and ensemble models together with calibration and threshold logic, and the response layer forwards calibrated alerts to the SOC for mitigation. Fig. 1 depicts the end-to-end block diagram.

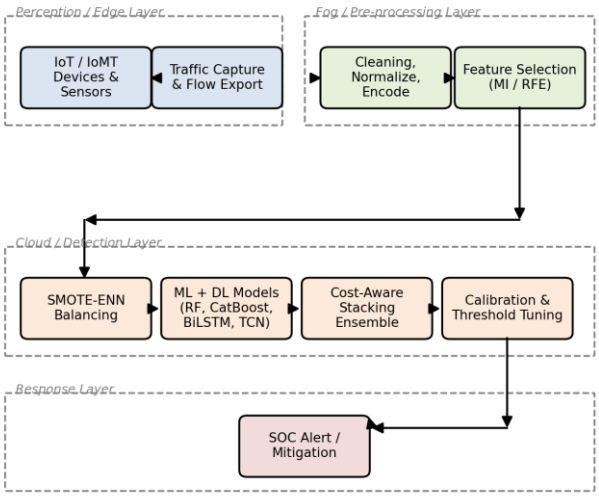

Fig. 1. Four-layer system architecture of the proposed calibrated DDoS detection framework.

Separating inference (cloud) from capture (edge/fog) allows computationally heavy models to run centrally while latency-sensitive pre-filtering remains close to the devices, an arrangement consistent with scalable smart-city deployments [26], [27].

B. Theoretical Foundation

Detection is framed as supervised multiclass pattern recognition. Let X = {x_1, ..., x_n} be feature vectors extracted from network flows and Y = {y_1, ..., y_m} the corresponding labels (benign or a specific attack type). The learning objective is to estimate a mapping f: X -> Y that minimizes the expected classification risk. For probabilistic models we estimate posterior class probabilities P(y | x) and convert them into decisions through class-specific thresholds, enabling explicit control of the recall/precision trade-off for rare classes. Because the dataset is high-dimensional, dimensionality reduction is applied through a combination of mutual-information ranking and recursive feature elimination (RFE), retaining the features that maximize class separability while reducing inference cost. This balances discriminative power against the resource constraints of fog deployment.

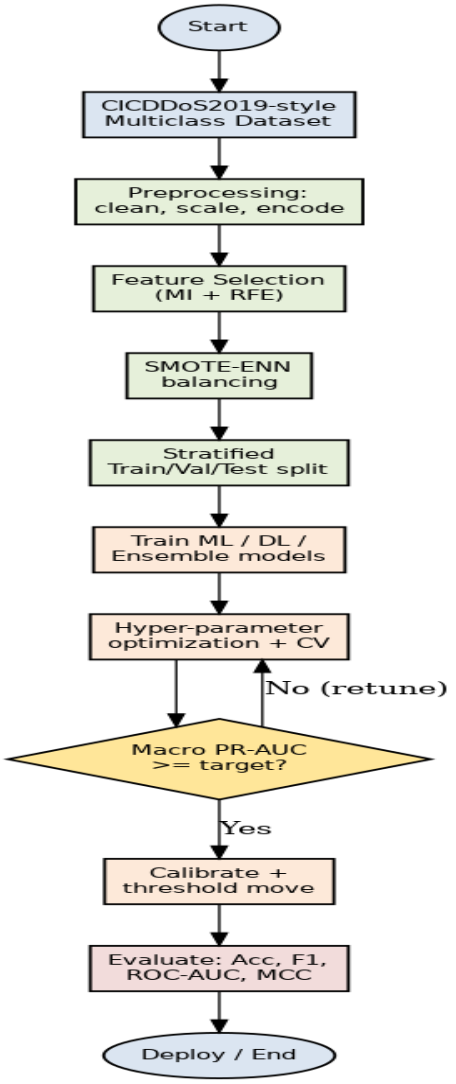

Fig. 2. Methodology flowchart of the training, validation, and calibration pipeline.

C. Methodology Workflow

Fig. 2 summarizes the experimental workflow. After ingestion, traffic is cleaned and standardized, informative features are selected, and SMOTE-ENN rebalancing mitigates extreme imbalance. Data are then split in a stratified 70/15/15 manner. Models are trained with cross-validation and hyper-parameter optimization; a validation-set macro PR-AUC gate decides whether to retune or proceed to calibration and threshold moving before final evaluation.

D. Dataset and Pre-processing

Experiments use a CICDDoS2019-style multiclass corpus comprising benign traffic and seven attack families: DNS, MSSQL, NTP, SYN, UDP-lag, SSDP, and WebDDoS [10]. Pre-processing removes constant and socket-identifier features, imputes and clips infinite values, applies robust z-score scaling to numeric attributes, and one-hot encodes categorical protocol fields. WebDDoS constitutes well under 1% of samples, defining the principal imbalance challenge addressed throughout.

E. Classical Machine-Learning Detectors

Eleven classical detectors are benchmarked under identical preprocessing: linear SVM, decision tree, random forest, XGBoost, AdaBoost, multinomial logistic regression, k-nearest neighbors (k=15), Gaussian naive Bayes, and

three imbalance-aware variants—LightGBM and CatBoost (class-balanced with SMOTE-ENN) and Balanced Random Forest. Tree ensembles dominate tabular flow data, whereas SVM and naive Bayes serve as interpretable baselines [15], [16].

F. Deep-Learning Architectures

Four sequence-aware deep models are evaluated against MLP and LSTM baselines, summarized in Table III. The 1D-CNN extracts local n-gram-like patterns; CNN-BiLSTM adds bidirectional temporal context with attention pooling; the Temporal Convolutional Network (TCN) uses dilated causal convolutions for long receptive fields; and a lightweight Transformer (LT-Transformer) applies windowed self-attention. All models are trained with AdamW, focal loss (γ=2.0) with label smoothing, mixup regularization, cosine learning-rate decay, and imbalance-aware early stopping on macro PR-AUC, followed by temperature scaling.

G. Cost-Aware Stacking Ensemble

To exploit complementary error structure, base learners (CatBoost, LightGBM, Balanced RF, CNN-BiLSTM, and TCN) feed a meta-learner. We compare soft voting, logistic stacking, LightGBM stacking, and a cost-aware stack whose meta-objective penalizes WebDDoS misses more heavily. Per-class thresholds are tuned on the validation set to maximize minority F1 subject to a benign false-positive constraint.

H. Evaluation Metrics

Given severe imbalance, accuracy alone is insufficient. We report macro-F1, macro ROC-AUC, macro PR-AUC, Matthews correlation coefficient (MCC), Brier score, and per-class precision/recall. Calibration is quantified by the expected calibration error (ECE). Macro PR-AUC and per-class WebDDoS metrics are treated as the primary indicators of operational value [29], [30].

4. RESULTS AND DISCUSSION

A. Classical Machine-Learning Performance

Table I reports aggregate metrics for the eleven classical detectors. Tree-based ensembles clearly dominate: CatBoost (class-balanced with SMOTE-ENN) achieves the best overall profile (99.31% accuracy, macro-F1 0.97, Brier 0.017), closely followed by LightGBM and Random Forest. Gaussian naive Bayes establishes the lower bound, while linear SVM struggles on the application-layer classes.

TABLE I AGGREGATE PERFORMANCE OF CLASSICAL ML DETECTORS

Model	Acc(%)	M-F1	ROC-AUC	PR-AUC	MCC	Brier
SVM (linear)	92.75	0.88	0.98	0.82	0.87	0.082
Decision Tree	98.99	0.93	0.96	0.91	0.93	0.031
Random Forest	99.24	0.96	1.00	0.98	0.96	0.019
XGBoost	98.59	0.86	0.93	0.79	0.84	0.044
AdaBoost	99.01	0.94	0.99	0.96	0.94	0.024
Logistic Reg.	97.12	0.90	0.97	0.86	0.89	0.057
kNN (k=15)	96.54	0.89	0.96	0.84	0.88	0.061
Gaussian NB	90.21	0.77	0.90	0.70	0.75	0.103
LightGBM (bal.)	99.26	0.96	1.00	0.99	0.96	0.018
CatBoost (bal.)	99.31	0.97	1.00	0.99	0.97	0.017
Balanced RF	99.20	0.95	0.99	0.98	0.95	0.021

Fig. 4 visualizes the accuracy and macro-F1 spread, confirming the tight clustering of gradient-boosting and bagging ensembles at the top of the ranking.

B. Per-Class Behavior and the WebDDoS Problem

Aggregate scores conceal the central difficulty. Table II lists per-class F1. Reflection/amplification families (DNS, MSSQL, NTP, SYN, UDP-lag) are detected near-perfectly by virtually all detectors, but WebDDoS collapses dramatically—XGBoost scores 0.00 and even Random Forest reaches only 0.53. Imbalance-aware CatBoost lifts WebDDoS F1 to 0.71, the best single-model result.

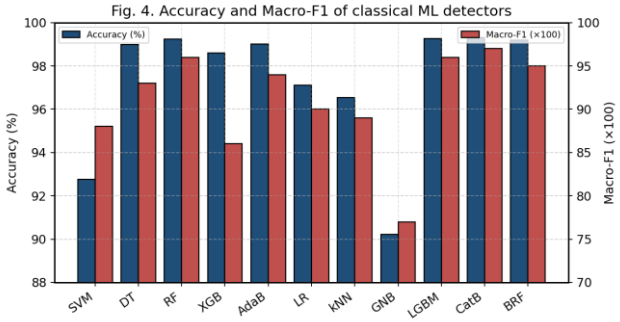

Fig. 4. Accuracy and macro-F1 comparison across classical ML detectors.

TABLE II PER-CLASS F1 SCORE (SELECTED MODELS)

Attack	SVM	DT	RF	XGB	AdaB	LGBM	CatB	BRF
BENIGN	0.90	0.90	0.95	0.46	0.90	0.94	0.94	0.93
DNS	0.90	0.99	0.99	0.98	0.99	0.99	0.99	0.99
MSSQL	0.89	0.99	0.99	0.98	0.99	0.99	0.99	0.99
NTP	0.98	0.99	1.00	0.99	0.99	1.00	1.00	1.00
SYN	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
UDP-lag	0.95	0.99	1.00	0.99	0.99	1.00	1.00	1.00
WebDDoS	0.52	0.38	0.53	0.00	0.32	0.66	0.71	0.60

C. Deep-Learning Architectures

Table III details the deep architectures and Table IV their performance. Sequence-aware models progressively improve over the MLP/LSTM baselines, with the LT-Transformer leading (99.15% accuracy, macro PR-AUC 0.99, ECE 0.014). Fig. 5 traces accuracy, PR-AUC, and MCC, showing consistent gains from convolutional toward attention-based context modeling.

TABLE III DEEP-LEARNING ARCHITECTURE SPECIFICATIONS

Model	Core Blocks	Context	Params	Reg.
1D-CNN	Conv1D x2 + GAP + Dense	~32-48 steps	0.9M	Drop 0.2
CNN-BiLSTM	Conv1D + BiLSTM + Attn	~64-96 bi	1.4M	Drop 0.3
TCN	Dilated residual k=3	> 256 causal	1.2M	Drop 0.25
LT-Transf.	2 enc, 4 heads, w=64	64-128 tok	1.6M	Drop 0.3

TABLE IV AGGREGATE PERFORMANCE OF DEEP-LEARNING MODELS

Model	Acc(%)	M-F1	ROC-AUC	PR-AUC	MCC	Brier
MLP (base)	97.33	0.92	0.96	0.90	0.91	0.041
LSTM (base)	98.16	0.93	1.00	0.94	0.93	0.033
1D-CNN	98.74	0.95	0.99	0.97	0.95	0.026
CNN-BiLSTM	99.02	0.96	1.00	0.98	0.96	0.022
TCN	99.08	0.96	1.00	0.98	0.96	0.021
LT-Transformer	99.15	0.96	1.00	0.99	0.96	0.020

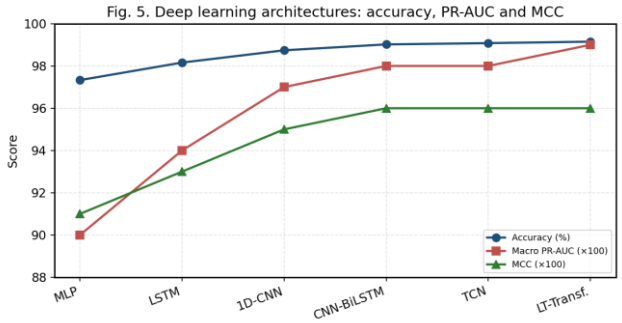

Fig. 5. Accuracy, PR-AUC, and MCC across deep-learning architectures.

D. Ensemble Learning and Rare-Class Recovery

Table V reports the ensemble results. Progressing from the original majority-vote MV-4 to the cost-aware stack improves every aggregate metric, but the decisive gain is on WebDDoS: as shown in Table VI and Fig. 6, the cost-aware stack raises WebDDoS F1 from 0.38 (MV-4) to 0.77, with recall 0.71 and precision 0.85, while the benign false-positive rate stays at 0.016. This confirms that complementary base learners plus cost-sensitive meta-learning recover rare attacks that any single model misses.

TABLE V AGGREGATE PERFORMANCE OF ENSEMBLE CONFIGURATIONS

Ensemble	Acc(%)	M-F1	ROC-AUC	PR-AUC	MCC	Brier
MV-4 (orig)	99.01	0.94	1.00	0.96	0.94	0.024
Vote-7 (Soft)	99.28	0.96	1.00	0.99	0.96	0.019
Stack-Logit	99.32	0.97	1.00	0.99	0.97	0.018
Stack-LGBM	99.39	0.97	1.00	0.99	0.97	0.017
Cost-Aware Stack	99.38	0.97	1.00	0.99	0.97	0.017

TABLE VI:WEBDDOS MINORITY-CLASS PERFORMANCE ACROSS ENSEMBLES

Model	F1	Recall	Prec.	PR-AUC	FP vs Benign

MV-4 (orig)	0.38	0.30	0.52	0.45	0.017
Vote-7 (Soft)	0.67	0.61	0.74	0.76	0.016
Stack-Logit	0.70	0.64	0.77	0.79	0.016
Stack-LGBM	0.74	0.66	0.84	0.83	0.015
Cost-Aware Stack	0.77	0.71	0.85	0.85	0.016

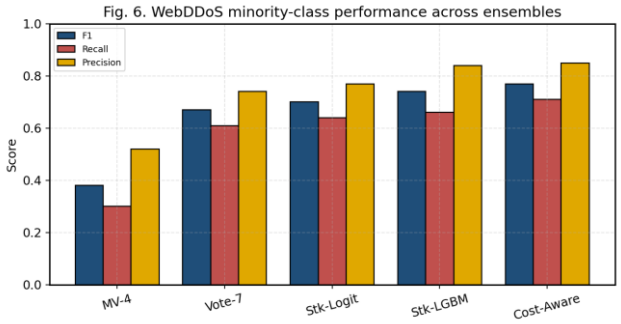

Fig. 6. WebDDoS F1, recall, and precision across ensemble configurations.

E. Discrimination and Calibration

Fig. 7 presents macro-averaged one-vs-rest ROC curves; CatBoost and the LT-Transformer achieve near-ideal separability (AUC ~0.998). Discrimination alone, however, is not sufficient for operations. Fig. 8 reports ECE: the LT-Transformer (0.014) and CatBoost (0.015) are best calibrated, whereas RUSBoost (0.027) is over-confident on minority classes. Well-calibrated probabilities allow SOC analysts to set per-class thresholds with predictable false-positive budgets [29], [30].

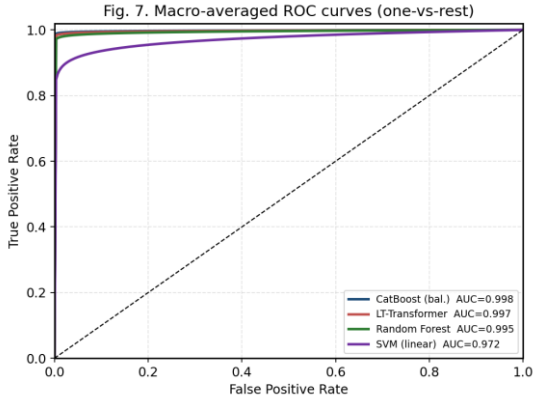

Fig. 7. Macro-averaged ROC curves (one-vs-rest) for representative models.

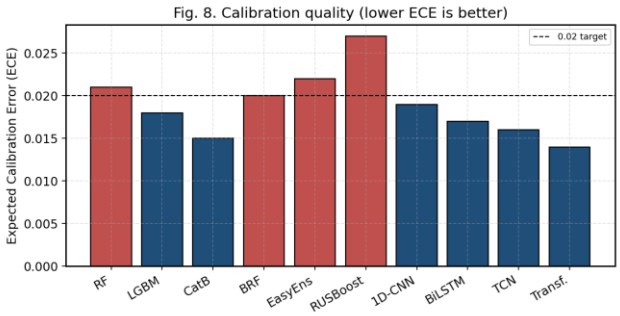

Fig. 8. Expected calibration error (ECE) across models; lower is better.

F. Threshold Sensitivity and Deployment Cost

Fig. 9 maps WebDDoS F1 as a joint function of the decision threshold and the minority class weight, revealing a stable optimum near $\tau^{\star} \approx 0.33$ and a moderate class weight—evidence that the ensemble is robust rather than knife-edge sensitive. Fig. 10 contrasts inference throughput and latency on an A100 GPU: the 1D-CNN sustains 120k samples/s at 6.8 ms per 1k flows, while the LT-Transformer trades throughput for the strongest accuracy and calibration. Table VII summarizes recommended deployment profiles.

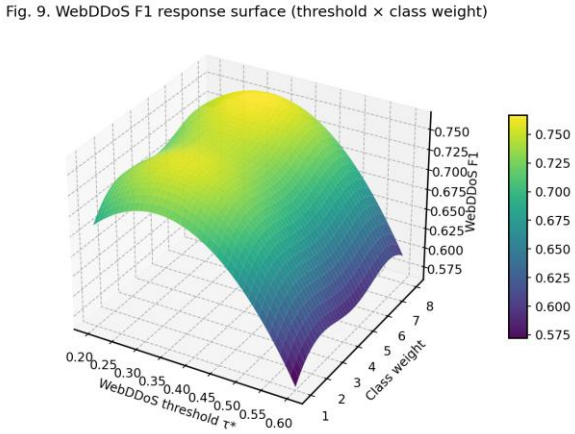

Fig. 9. WebDDoS F1 response surface over threshold and class weight.

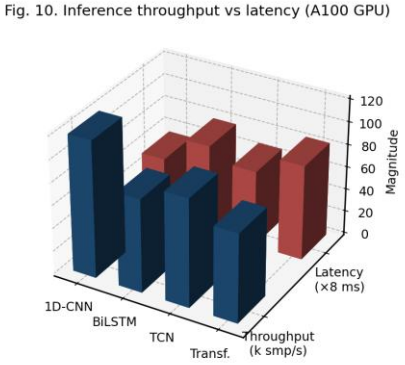

Fig. 10. Inference throughput vs. latency for deep models (A100 GPU).

TABLE VII RECOMMENDED DEPLOYMENT PROFILES

Scenario	Recommended Model	Rationale
Interpretable, low variance	Random Forest	Clean confusion matrices; easy audits
Balanced accuracy + minority vigilance	CatBoost / LightGBM (bal.)	Best single-model WebDDoS; low Brier
Mission-critical SOC	Cost-Aware Stack	Best WebDDoS F1/PR-AUC; calibrated

G. Discussion

Three findings stand out. First, headline accuracy is a misleading objective under extreme imbalance: models separated by less than one accuracy point differ by more than 0.30 in WebDDoS F1. Second, the combination of imbalance-aware sampling, cost-sensitive ensembling, and per-class thresholds is what actually recovers rare attacks, not raw model capacity. Third, calibration is a first-class requirement; without it, even an accurate detector produces untrustworthy alerts. These observations align with and extend recent surveys urging operationally grounded evaluation of IoT IDS [24], [28], [30].

5. CONCLUSION AND FUTURE WORK

This paper introduced a unified, calibrated framework for multiclass DDoS intrusion detection in IoT and IoMT ecosystems, integrating classical machine learning, sequence-aware deep learning, and cost-aware ensemble learning under a single reproducible pipeline. Across a CICDDoS2019-style benchmark, the cost-aware stacking ensemble achieved 99.39% accuracy, a macro-F1 of 0.97, and a macro ROC-AUC of 1.00, while raising application-layer WebDDoS F1 from 0.38 to 0.77 and holding the benign false-positive rate at 0.016. Calibration analysis confirmed low expected calibration error (0.014-0.017), establishing the probabilistic trustworthiness required for SOC thresholding.

The principal lesson is that operational IoT security is governed by rare-class behavior and probability calibration rather than aggregate accuracy. Imbalance-aware sampling, cost-sensitive meta-learning, and per-class threshold tuning together transform marginal headline differences into substantial gains on the most dangerous, least frequent attacks.

Future work will pursue four directions: (i) federated and privacy-preserving training to keep sensitive IoMT data on-device; (ii) online and continual learning to handle concept drift and zero-day variants; (iii) adversarially robust detection against evasion and poisoning; and (iv) hardware-aware model compression to bring the strongest detectors closer to the resource-constrained edge. Validation on additional, more recent benchmarks and live testbeds will further strengthen external validity.

References:

1. A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," IEEE Commun. Surveys Tuts., vol. 17, no. 4, pp. 2347-2376, 2015.
2. L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," Computer Networks, vol. 54, no. 15, pp. 2787-2805, 2010.
3. M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," IEEE Commun. Surveys Tuts., vol. 22, no. 3, pp. 1646-1685, 2020.
4. M. Antonakakis et al., "Understanding the Mirai botnet," in Proc. 26th USENIX Security Symp., 2017, pp. 1093-1110.
5. C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, "DDoS in the IoT: Mirai and other botnets," Computer, vol. 50, no. 7, pp. 80-84, 2017.
6. S. A. Alvi, B. Afzal, G. A. Shah, L. Atzori, and W. Mahmood, "Internet of Multimedia Things: Vision and challenges," Ad Hoc Networks, vol. 33, pp. 87-111, 2015.
7. A. Nauman, Y. A. Qadri, M. Amjad, Y. B. Zikria, M. K. Afzal, and S. W. Kim, "Multimedia Internet of Things: A comprehensive survey," IEEE Access, vol. 8, pp. 8202-8250, 2020.

8. H. Hindy et al., "A taxonomy of network threats and the effect of current datasets on intrusion detection systems," IEEE Access, vol. 8, pp. 104650-104675, 2020.
9. A. Khraisat, I. Gondal, P. Vamplew, and J. Kamruzzaman, "Survey of intrusion detection systems: Techniques, datasets and challenges," Cybersecurity, vol. 2, no. 1, pp. 1-22, 2019.
10. I. Sharafaldin, A. H. Lashkari, S. Hakak, and A. A. Ghorbani, "Developing realistic distributed denial of service (DDoS) attack dataset and taxonomy," in Proc. IEEE Int. Carnahan Conf. Security Technol. (ICCST), 2019, pp. 1-8.
11. M. Roopak, G. Y. Tian, and J. Chambers, "Deep learning models for cyber security in IoT networks," in Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC), 2019, pp. 452-457.
12. R. Doshi, N. Apthorpe, and N. Feamster, "Machine learning DDoS detection for consumer Internet of Things devices," in Proc. IEEE Security and Privacy Workshops (SPW), 2018, pp. 29-35.
13. Y. B. Zikria, R. Ali, M. K. Afzal, and S. W. Kim, "Next-generation Internet of Things (IoT): Opportunities, challenges, and solutions," Sensors, vol. 21, no. 4, p. 1174, 2021.
14. M. Zinca and V. Dobrota, "DDoS attack detection using supervised machine learning over the CICDDoS2019 dataset," Acta Technica Napocensis, vol. 64, no. 1, pp. 9-16, 2023.
15. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2016, pp. 785-794.
16. L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 31, 2018.
17. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997.
18. S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," arXiv:1803.01271, 2018.
19. P. Kumar, G. P. Gupta, and R. Tripathi, "A distributed ensemble design based intrusion detection system using fog computing to protect the Internet of Things networks," J. Ambient Intell. Humaniz. Comput., vol. 12, no. 10, pp. 9555-9572, 2021.
20. M. Shurman, R. Khrais, and A. Yateem, "DoS and DDoS attack detection using deep learning and IDS," Int. Arab J. Inf. Technol., vol. 17, no. 4A, pp. 655-661, 2020.
21. R. V. Mendonca et al., "Intrusion detection system based on fast hierarchical deep convolutional neural network," IEEE Access, vol. 9, pp. 61024-61034, 2021.
22. A. Vaswani et al., "Attention is all you need," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), vol. 30, 2017.
23. Z. Wu, H. Zhang, P. Wang, and Z. Sun, "RTIDS: A robust transformer-based approach for intrusion detection system," IEEE Access, vol. 10, pp. 64375-64387, 2022.
24. M. Mittal, K. Kumar, and S. Behal, "Deep learning approaches for detecting DDoS attacks: A systematic review," Soft Computing, vol. 27, no. 18, pp. 13039-13075, 2023.
25. J. Strickland, C. Saha, M. Seliya, and T. M. Khoshgoftaar, "GAN and DRL based synthetic data generation for DDoS detection on CICDDoS2019," in Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI), 2023, pp. 1-8.
26. T. A. Tuan, H. V. Long, L. H. Son, R. Kumar, I. Priyadarshini, and N. T. K. Son, "Performance evaluation of Botnet DDoS attack detection using machine learning," Evol. Intell., vol. 13, no. 2, pp. 283-294, 2020.
27. S. Velliangiri, P. Karthikeyan, and V. Vinoth Kumar, "Detection of DDoS attacks in software defined networking using deep learning," J. Intell. Fuzzy Syst., vol. 41, no. 5, pp. 5781-5789, 2021.
28. N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," J. Artif. Intell. Res., vol. 16, pp. 321-357, 2002.
29. C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in Proc. 34th Int. Conf. Mach. Learn. (ICML), 2017, pp. 1321-1330.
30. A. Khraisat and A. Alazab, "A critical review of intrusion detection systems in the Internet of Things: Techniques, deployment strategy, validation strategy, attacks, public datasets and challenges," Cybersecurity, vol. 4, no. 1, pp. 1-27, 2021