IoT Detection and Classification Techniques: A Review of Algorithms and Applications

Manar Bashar Mortatha
College of Science, Diyala University, Diyala, Iraq
College of Education for Pure Sciences, Wasit University
Al-Kut, Wasit, Iraq
scicomphd232407@uodiyala.edu.iq

Dhahir Abdulhade Abdulah
College of Science, Diyala University
Diyala, Iraq
dhahair@uodiyala.edu.iq

Abstract*—The Internet of Things (IoT) currently interconnects smart cities, healthcare systems, and industrial automation. However, the presence of heterogeneous, resource-constrained devices and diverse protocol stacks increases the difficulty of securing and reliably operating IoT systems. This survey reviews intrusion detection alongside traffic and device classification, covering classical statistical approaches, machine learning (ML), deep learning (DL), and hybrid or metaheuristic pipelines. We discuss supervised, unsupervised, and semi-supervised ML methods (e.g., Random Forest and SVM), DL architectures (e.g., CNNs and LSTMs), and hybrid schemes that combine metaheuristic feature selection with efficient classifiers to improve scalability, interpretability, and robustness. Beyond reporting accuracy, we synthesize evidence on deployment-centric trade-offs, highlighting that many studies omit key operational metrics such as the false-positive rate, inference latency, memory footprint, and energy consumption. For practical deployment, detectors should be lightweight and secure, achieve a robust macro-F1 score with low false positives, and meet strict latency and footprint constraints while remaining maintainable and, where possible, explainable. Based on the reviewed literature, we identify research gaps and outline future directions for building versatile and resilient IoT intrusion-detection and classification systems.

Keywords— Internet of Things (IoT); intrusion detection; traffic classification; machine learning; deep learning; metaheuristic feature selection; edge computing; latency

I. INTRODUCTION

The Internet of Things (IoT) has emerged as a key enabler across healthcare, smart cities, and industrial automation, driven by increased device connectivity and technologies such as 5G and LPWANs [1]. IoT devices collect, share, and process data to facilitate intelligent decision-making, supporting remote patient monitoring [2], dynamic traffic management [3], and IIoT predictive analytics [4]. However, this proliferation raises significant challenges in security, scalability, and data trust, making anomaly detection and device classification essential for robust IoT operations [5], [7].

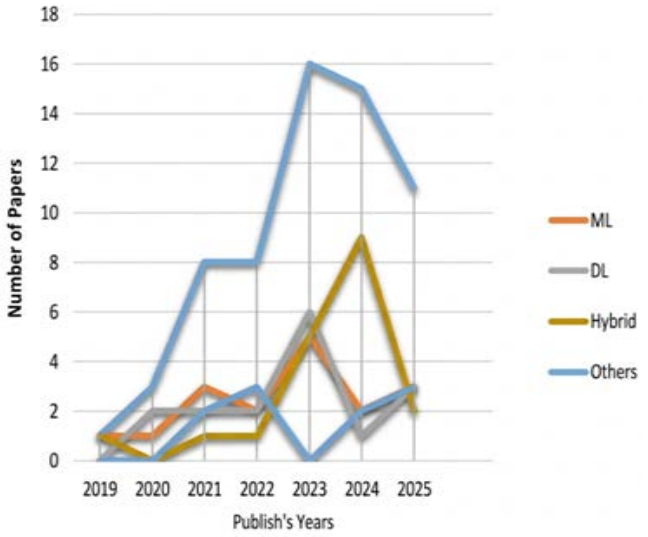

Fig. 1. Distribution of reviewed IoT detection & classification publications by method (2019–2025)

A diverse range of detection and classification methods has been proposed, from statistical and classical ML to DL and hybrid metaheuristic pipelines. Lightweight models are often required for resource-constrained IoT devices [9].

Motivation. Despite numerous accuracy-focused results, practitioners lack a deployment-oriented synthesis for heterogeneous, resource-constrained IoT. This review emphasizes modern datasets, practical edge/fog/cloud model placement, and a balanced accuracy-vs-footprint perspective [5], [6], [19], [22], [33].

Main contributions. This review makes the following contributions: (1) A concise, deployment-oriented mapping of ML, DL, and hybrid/metaheuristic IDS techniques to edge, fog, and cloud tiers; (2) A critical analysis of dataset recency and cross-dataset robustness, moving beyond mere headline accuracy; (3) A compact trade-off comparison (accuracy vs. compute/latency vs. interpretability) intended for operational use; (4) A practical workflow figure (Fig. 2) and summary tables (Tables 1–3) that serve as a design checklist. Unlike prior surveys that focus on accuracy rankings, this review is explicitly deployment-aware: it integrates operational constraints (latency, memory footprint, FPR) and maps each technique to its practical edge/fog/cloud placement tier—providing actionable guidance for real IoT system design.

A. Review Scope and Method

We systematically reviewed peer-reviewed IoT/IIoT/IoMT IDS and traffic/device-classification studies (2019–2025) indexed in Scopus, IEEE Xplore, and Web of Science (searched to January 2026), retaining 62 papers after applying inclusion/exclusion criteria. Unlike prior surveys, data extraction explicitly captured deployment indicators

(latency, model size, compute cost, FPR) alongside standard metrics—enabling our deployment-tier mapping across edge, fog, and cloud tiers presented in Tables 1–5.

II. Background and Fundamentals

This section briefly reviews core concepts only to establish consistent terminology; we deliberately avoid lengthy textbook-style explanations and instead emphasize implications for evaluation and deployment in later sections.

A. IoT Detection and Classification Definitions

Detection covers identification of anomalies, attacks, or hardware failures; classification categorizes devices, traffic, or behaviors. Together they underpin IoT security: anomaly-based detection identifies unknown attack patterns [5], while device classification supports resource allocation and protocol management [7], [10]. Privacy and resource constraints are driving federated learning and edge/fog lightweight models [19], [22], [33].

- Resource Constraints

IoT nodes operate under strict computational budgets (limited CPU, RAM, battery), constraining model complexity; lightweight models (e.g., pruned DTs, compressed NNs) are preferred for edge-tier deployment [9].

- Heterogeneity

The IoT environment encompasses diverse hardware, protocols, and configurations, making it difficult to train models that generalize across platforms [11], [21].

- Scalability

Large-scale IoT deployments require hierarchical IDS architectures where lightweight edge models filter traffic before cloud-tier correlation to maintain throughput and low FPR [6].

- Security and Privacy

The distributed nature of IoT makes these networks frequent targets; effective IDS must achieve real-time detection while preserving user privacy [7], [12].

III. Detection and Classification Techniques

A key observation is the increasing shift from accuracy-driven to deployment-aware evaluation, reflecting recognition that high benchmark performance does not guarantee practical effectiveness under real IoT constraints (latency, scalability, resource consumption).

Detection and classification techniques are grouped into classical statistical, ML, DL, and hybrid methods [16]; choice depends on use-case and data characteristics.

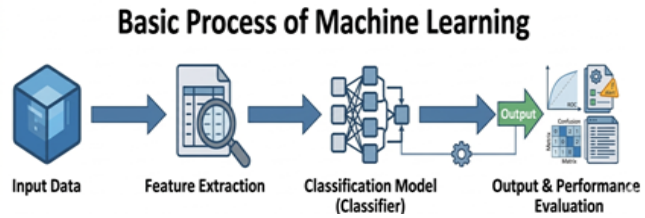

Fig. 2. ML-based IoT detection workflow [21]

A. Machine-Learning–Based IoT Detection and Classification

ML has proven highly effective for IoT IDS and device classification across both legacy (NSL-KDD, CICIDS2017) and modern datasets (IoT-23, ToN_IoT, CIC-IoT2022, MQTT-IoT-IDS2020, Edge-IIoTset) [19], [22], [32], [33]. For resource-constrained nodes, tree-based models with feature selection achieve near-DL accuracy at lower latency and energy cost [7], [12]. Table 1 reviews ML-based methods with deployment indicators.

B. Comparative Synthesis and Practical Trade-offs (ML)

Beyond headline accuracy, ML methods tend to succeed when (i) discriminative flow/behavioral features are available, (ii) class imbalance is handled, and (iii) evaluation avoids device/time leakage. Tree ensembles (RF/GB/XGBoost/LightGBM) deliver strong macro-F1 with low inference cost, making them suitable for edge gateways, while SVM may incur higher cost on large feature spaces.

When near-perfect scores are reported, they should be interpreted cautiously: IoT datasets often contain repeated flows or temporally correlated sessions, and random splits can inflate accuracy compared with time-based or device-disjoint splits. FPR is often more operationally critical than accuracy in high-volume IoT networks. We therefore prioritize macro-F1, FPR, and latency/footprint indicators; where studies omit these, we rely on model-family expectations (Table 1).

To substantiate the deployment-oriented perspective, we summarize typical trade-offs across model families in Table 2 and provide a conceptual accuracy–cost visualization in Fig. 3.

TABLE I. Typical Deployment Trade-offs by Model Family (Qualitative)

Model family	Typical tier	Latency	Footprint	Interpretability	Notes
Classical ML (trees, SVM, LR)	Edge/ Fog	Low–Medium	Low	Medium–High	Strong baselines; good for gateways; benefit from feature engineering.
Deep learning (CNN/LSTM/ Transformers)	Fog/Cloud (sometimes Edge)	Medium–High	Medium–High	Low–Medium	High accuracy; needs careful compression/acceleration for edge.
Hybrid/metaheuristic + ML/DL	Edge/ Fog + Cloud retraining	Low–Medium	Low–Medium	Medium	Feature selection reduces cost; cloud can handle periodic retraining/tuning.

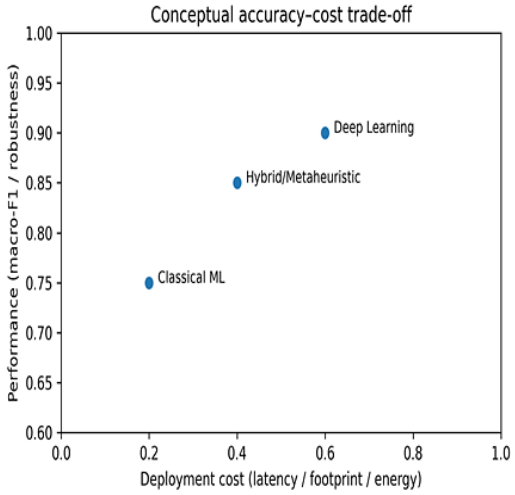

Fig. 3. Conceptual trade-off between performance and deployment cost across model families

1) Supervised Learning

Among supervised ML methods, gradient-boosted ensembles and feature-selection wrappers consistently dominate IoT IDS benchmarks, combining strong macro-F1 with low inference overhead [18].

SVM separates normal from anomalous traffic by margin maximization [19]. With CorrAUC feature selection on Bot-IoT, tree-based learners reached ~99–100% accuracy while SVM maintained >96%, confirming compact feature sets suffice for high-fidelity detection.

Logistic Regression (LR) is a lightweight baseline and stable component in voting ensembles, though its linear boundary limits performance on complex anomalies [25], [32], [34].

2) Unsupervised Learning

Unsupervised methods (K-Means, DBSCAN) trade accuracy for labeling independence, making them valuable for zero-day detection and novel device profiling [28], [29], [30]. Semi-supervised approaches combine a small labeled set with unlabeled data; One-Class SVM, for instance, fits a boundary on normal instances and flags deviations, achieving 95–98% accuracy on IoT traffic [17], [18], [20].

TABLE II. SUMMARY OF ML-BASED LITERATURE FOR IOT'S DETECTION AND CLASSIFICATION.

Ref.	Best Technique	Dataset	Results (Accuracy %)	Metrics (in %)	Year
[20]	RF + Grid-Search (best); others: NB 98.9%, KNN 97.9%, LDA 97.1%, SVM 95.7%	DDoS-SDN	Acc: 99.9%	P=99.99, R=99.99, F1=99.99	2025
[31]	RF (best); SVM and KNN also tested on two datasets	WUSTL EHMS 2020; ICU-IoMT	Acc: 80–86% (best: RF on ICU-IoMT)	P: 85%, R: 84%, F1: 83%	2024
[24]	DT+GA (best); RF, SVM, KNN also compared	IoTID20	Acc: 99.9% (DT+GA)	P: 98.8%, R: 99.2%, F1: 99%	2024
[22]	XGBoost (best); SVM and DCNN also evaluated	IoT-23; NSL-KDD; ToN_IoT	Acc: 99.98% / 80.3% / 99.90%	P=99.99, R=99.98, F1=99.99	2024
[32]	RF (best); LR, NB, AdaBoost, SVM also tested	CIDAD (CoAP-IoT)	Acc: 99.99% (RF)	P: 99.53, R: 100, F1: 99.76	2023
[30]	Unsupervised (K-Means, DBSCAN); Supervised (RF, SVM, DT)	Custom IoT (8 devices)	Supervised Acc: 87.5–99.1%	Macro P: 82.8, Macro R: 79.9	2023
[33]	LR / RF / ANN (best); NB lowest	Ds2OS traffic traces	Acc: 99.4% (LR/RF)	P: 99.34–99.89, F1: 94–96	2023
[34]	LR with LogitBoost	Network traffic (41 IoT devices)	Acc: 99.7–99.9%	F1: 99.7–99.9%	2021
[19]	CorrAUC + DT (best); RF, SVM, NB also tested	Bot-IoT	Acc: 99.9% (DT)	P: 99, R: 99, Spec: 98.95%	2020
[18]	LightGBM (best); Prophet, VAR also compared	Smart-home appliance + weather	Acc: 71.8% (LightGBM)	MAE ≈ 28.2	2019

C. Deep-Learning–Based IoT Detection and Classification

DL models report very high detection accuracy on modern IoT datasets using CNN/LSTM ensembles [41], [42], [44], though their computational demands motivate hierarchical edge–cloud pipelines with lightweight front-end models.

- Comparative Synthesis and Deployment Implications (DL)

DL models often outperform shallow learners when raw inputs are high-dimensional (e.g., packet sequences, multivariate sensor streams) or when temporal structure matters. However, their success is highly context-dependent: CNN-based models can excel on flow-feature tensors but may degrade under traffic drift or device changes unless retrained; sequence models (LSTM/TCN/Transformers) better capture temporal dependencies but may be sensitive to windowing choices and class imbalance.

From a deployment perspective, DL introduces additional constraints: without compression (quantization/pruning/distillation), memory footprint and inference latency can exceed edge budgets. Therefore, many practical IoT deployments adopt tiered processing—a lightweight model for first-line triage at edge/fog, with heavier DL correlation in the cloud.

To avoid over-claiming, we emphasize studies that report FPR, latency, or cross-dataset testing, and we qualify legacy-dataset results (e.g., KDDCup-99/NSL-KDD) as limited evidence of modern IoT generalization.

DL learns rich hierarchical representations from large heterogeneous telemetry where shallow ML overfit, demonstrating reduced false positives in anomaly-based IDS [36], CNN-based IoT detection [37], [45], and medical IoT monitoring [38]. On ToN_IoT, the DIS-IoT stacking ensemble [39] (MLP+DNN+CNN+LSTM) reports 99.6% accuracy, 0.2% FPR, and F1=0.994. On MQTT-IoT-IDS2020, compact models report ≈99.6% precision/recall/F1 with ≤0.03% FPR [42]. For legacy corpora, C2-LSTM on

KDDCup-99 reports >99.5% accuracy, converging ~20% faster than C-LSTM [40].

Transfer-learning CNN backbones on IoT-DS-2 yield ≥99.60% TPR across 15 attack categories [44], and GA-guided DL pipelines in IIoT jointly optimize feature selection and classification [49].

TABLE III. SUMMARY OF DL-BASED LITERATURE FOR IOT'S DETECTION AND CLASSIFICATION.

Ref.	Best Technique	Dataset	Results (Accuracy %)	Metrics (in %)	Year
[46]	FFNN (best); LSTM, RandNN also compared	CIC-IoT2022	FFNN Acc: 99.93%	Acc: 99.93% (FFNN), 99.85% (LSTM), 96.42% (RandNN)	2023
[41]	Ensemble DL (LSTM+CNN+ ANN) best; Binary LSTM also tested	IoT-23 (20 GB)	Ensemble Acc: 99.93%	P: 99.1, R: 99.6, F1: 99.6	2023
[42]	Deep Learning with Bi-flow + Uni-flow Features	MQTT-IoT-IDS2020	Bi-flow Acc: 99.56%, Uni-flow: 99.67%	P: 99.60, R: 99.60, F1: 99.60	2023
[40]	CNN + C2-LSTM Fusion Model	KDDCup-99 (10%)	High acc + improved spatio-temporal features	P: 99.8, R: 99.7, AUC: 99.8	2023
[39]	DIS-IoT Stacking Ensemble (MLP, DNN, CNN, LSTM)	ToN_IoT; CICIDS2017; SwaT	Binary Acc: 99.6; Multi-class: 99.7	P: 99.7, R: 99.7, F1: 99.7	2023
[47]	Voting Classifier (LR, RF, NB, ANN, 1D CNN)	ToN_IoT Telemetry (multi-device)	Acc: 99.7%	P: 99.8, R: 99.8, F1: 98.9	2022
[43]	CuDNN-LSTM (best); TCN, BI-LSTM also tested	SWaT (Singapore Univ.)	CuDNN-LSTM: lowest RMSE	RMSE: 0.035–0.049, F1: 0.99	2022
[44]	CNN1D (best acc); CNN2D, CNN3D, Transfer Learning also compared	BoT-IoT; MQTT-IoT-IDS2020; IoT-23; IoT-DS-1/2	Max Acc: 99.85%	P: 99.80, R: 99.82, F1: 99.83	2021

IV. HYBRID APPROACHES FOR IOT DETECTION AND CLASSIFICATION

High-dimensional telemetry and real-time constraints challenge classical IDS pipelines [50]. Hybrid designs that combine ML/DL with metaheuristic optimization address this gap and show consistently strong results on modern IoT benchmarks [51], [52]. Table 4 summarizes the reviewed hybrid literature.

A. Metaheuristic-Based Feature-Selection Optimization

Hybrid IDS pairing metaheuristic feature selection (WOA, PSO, GWO, GA) with efficient classifiers often approach DL accuracy at lower computational cost [52], [62]. For example, WOA+XGBoost achieved 99.66% accuracy and 0.23% FPR on a URL/traffic dataset [52]; binary Gravitational Search+GWO reduced UNSW-NB15 from 42 to 4 features enabling RF to reach 99.41% accuracy (F1=99.33%, FPR=0.03%); and GA-aided transfer learning on Edge-IIoTset tuned hyperparameters for multi-class IIoT attacks [54]. Such hybrids suit edge/fog first-line triage while the cloud handles global retraining.

B. Hybrid Methods in Feature Selection

Hybrid feature selection combining statistical filters with metaheuristic search is highly effective: one approach [53] ranked ToN-IoT features via Chi-Square and mutual information, then used NSGA-II to converge on 13 features, yielding 99.48% accuracy with RF. Recursive Feature Elimination (RFE) can similarly extract optimal subsets to enhance classifier performance [56].

C. Classification with Machine Learning and Deep Learning Models

After feature selection, multiple classifiers complement each other: bagged trees/RF reach ~99.8% on NSL-KDD, SVM achieves ~99.2%, and DT provides fast interpretable rules [57]. CNN–BiLSTM captures spatio-temporal patterns for multi-class threats [58], while PSO-optimized XGBoost/KNN exceed 89% in vehicular CAN-bus networks [59].

Together, RF/SVM baselines, DT interpretability, CNN–BiLSTM for complex threats, and PSO-optimized models form a scalable pipeline for diverse IoT environments including vehicular networks.

- Critical Synthesis of Hybrid/Metaheuristic Pipelines

Feature subsets or parameter configurations and scoring them using a wrapper objective (often validation F1-score or AUC). This mechanism can significantly reduce dimensionality, accelerate inference speed, and enhance robustness by eliminating noisy or redundant features.

However, if the wrapper evaluation leaks device or time information, the optimizer may over-specialize to dataset artifacts, inflating accuracy. The search phase itself is computationally expensive and should be treated as an offline training step at fog/cloud tier.

TABLE IV. SUMMARY OF HYBRID APPROACH-BASED LITERATURE FOR IOT'S DETECTION AND CLASSIFICATION

Ref.	Best Technique	Dataset	Results (Accuracy %)	Metrics (in %)	Year
[54]	DTL + CNN (Xception, VGG16/19, InceptionResNetV2) + GA Hyperopt + Bootstrap Ensemble	Edge-IIoTset	Acc: 100%	P: 100%, R: 100%, F1: 100%	2024
[59]	XGBoost + GSAPSO (best); KNN + GSAPSO also tested	CAN dataset (Chevrolet Impala)	Acc: 79.11% (XGBoost+ GSAPSO)	Cohen's Kappa = 0.076	2024
[52]	XGBoost + Modified Whale Optimization Algorithm (MWOA)	ISCX-URL-2016	Acc: 99.66	P=99.80, R=99.57, F1=99.69, FPR=0.23	2023
[53]	BGSA + BGWO Feature Selection + Random Forest	UNSW-NB15	Acc: 99.41%	DR: 99.09%, F1: 99.33%	2023

Ref.	Best Technique	Dataset	Results (Accuracy %)	Metrics (in %)	Year
[58]	SAEHO + CNN-DBN (best); SU-CMO + Bi-LSTM-GRU also tested	UNSW-NB15	Acc: 91.6% (SAEHO+ CNN-DBN)	Rand Index: 99.2%, MCC: 72.0%	2023
[60]	GTO + BSA Feature Selection + KNN; tested on 3 datasets	NSL-KDD; CICIDS-2017; UNSW-NB15; BoT-IoT	Acc: 95.5964% (NSL-KDD) to 98.7915% (CICIDS)	Sens/Spec per dataset	2023
[55]	RFE + Hybrid Metaheuristic (EOSCA+GOA) + ML classifiers	N-BaIoT	Acc: 99.86%	P: 99.94%, R: 99.94%, F1: 99.86%	2023
[57]	GA + DT (best); SVM, Ensemble Classifier also tested	NSL-KDD	Ensemble best performing	High accuracy across attack types	2020

V. Challenges and Future Directions

A. Practical Deployment Challenges

Moving from benchmarks to real IoT deployments requires careful placement across edge, fog/gateway, and cloud tiers (Table 5).

TABLE V. Deployment tiers and typical responsibilities for IoT detection/classification.

Tier	Primary goal	Typical models	Notes
Edge (device / sensor / micro-gateway)	Real-time triage	Lightweight ML, compressed DL	Strict latency/energy; prioritize low FPR.
Fog/Gateway (local aggregation)	Local response & filtering	Ensembles, hybrid FS+ML	Balances speed and accuracy; can maintain context windows.
Cloud (central analytics)	Global correlation & retraining	DL, large ensembles, continual learning	Handles heavy compute; distributes updated models.

- Scalability and resource constraints. IDS must scale across heterogeneous, resource-constrained devices while maintaining throughput and low FPR, requiring model compression and hierarchical edge/fog/cloud architectures.
- Real-time detection under tight latency budgets. Many IoT/IIoT applications cannot tolerate long detection delays, motivating model compression (pruning/quantization) and tiered edge/fog/cloud placement where lightweight ML handles first-line triage and heavier DL is offloaded upstream [61]. However, latency is still inconsistently reported across the reviewed studies, particularly when hardware platforms and deployment settings differ.

Deployment-related reporting remains inconsistent: most studies report only aggregate training/testing time while omitting footprint, energy, and hardware-normalized latency, which limits fair comparison.

- Model lifecycle management and concept drift. Attack strategies and normal behavior evolve over time, complicating model updates due to intermittent connectivity and regulatory constraints on raw data transfer. Incremental/federated learning and scheduled retraining are therefore essential [59], [60].
- Data scarcity and labeling costs. Scarce high-quality labeled traffic motivates semi-supervised, unsupervised, and transfer learning approaches [52], [62].
- Integration overhead and interoperability. Embedding models into IoT gateways introduces complexity: incompatible ML frameworks, limited accelerator support, and the need for standardized alert APIs. Lightweight models (DT/LR) are easier to operationalize; DL requires ONNX/TensorRT and careful monitoring [62].

B. Future Directions

Hybrid pipelines combining metaheuristic feature selection with efficient ML/DL classifiers remain highly promising for resource-constrained nodes. However, future research must become increasingly deployment-aware. We envision several key directions:

- Edge/Fog-Centric Design: Developing models with an "edge-first" paradigm—incorporating explicit latency, memory, and energy budgets—while reserving cloud resources for global retraining, cross-site correlation, and long-term storage, can significantly enhance responsiveness and reduce backbone network traffic [59], [61].
- Lifelong and Federated Learning: Federated learning with secure aggregation enables continuous model adaptation without centralizing raw data, preserving privacy across distributed IoT environments [60], [62].
- Leakage-Resistant Evaluation: Future studies must employ device-disjoint, time-based splits and validate on at least one external dataset. Near-perfect results from random splits often reflect data leakage rather than genuine robustness.

Overall, hybrid frameworks synergizing optimized feature selection, efficient classification, and explicit deployment constraints are uniquely positioned to address high-dimensional data, computational complexity, and evolving threats across heterogeneous IoT ecosystems.

B. Threats to Validity and Limitations of This Review

Key limitations include heterogeneous evaluation protocols across studies, frequent omission of deployment metrics (latency, energy, model size), and limited generalizability of results from legacy datasets to modern IoT traffic. To mitigate these, we grouped analysis by dataset and task type, highlighted deployment indicators where provided, and prioritized contemporary IoT datasets when drawing conclusions.

VI. Conclusion

This survey reviewed 62 peer-reviewed studies (2019–2025) on IoT intrusion detection and traffic/device classification from a deployment-aware perspective—synthesizing classical ML, deep learning, and hybrid metaheuristic pipelines not only by accuracy but by their practical suitability across edge, fog, and cloud tiers.

For resource-constrained edge nodes, tree-based ML models (Random Forest, XGBoost, LightGBM) paired with metaheuristic feature selection consistently deliver near-DL accuracy with substantially lower latency and memory footprint, making them the practical first-line choice for gateway-level triage. Deep learning architectures—particularly CNN/LSTM ensembles—outperform classical ML on high-dimensional or temporally structured inputs but require compression and tiered offloading to be deployable at the edge. Hybrid pipelines that combine metaheuristic optimizers (WOA, PSO, GWO, GA) with efficient classifiers represent the strongest balance between accuracy, interpretability, and computational cost, achieving state-of-the-art results on multiple IoT benchmarks while remaining deployable under realistic resource budgets.

A critical finding is the widespread omission of deployment-oriented metrics: most studies report only accuracy or F1-score, while inference latency, memory footprint, FPR, and energy consumption are rarely quantified. This review bridges theoretical performance and real-world deployment by aligning detection/classification techniques with operational IoT constraints, offering practitioners a model-selection guide beyond benchmark accuracy. Future research must adopt standardized deployment reporting and device-disjoint, time-based evaluation splits to ensure genuine generalizability.

Open challenges include managing concept drift via incremental and federated learning, reducing labeling costs through self-supervised strategies, and ensuring interoperability across heterogeneous IoT stacks. Closing these gaps demands closer collaboration between ML researchers and IoT practitioners. This review contributes a deployment-aware synthesis that goes beyond benchmark accuracy—directly supporting the design of IDS solutions that are not only accurate but lightweight, interpretable, and deployable in real heterogeneous IoT environments.

REFERENCES

[1] M. Islam et al., “Future industrial applications: Exploring LPWAN-driven IoT protocols,” Sensors, vol. 24, 2024.

[2] P. Matthew et al., “A review of the state of the art for the Internet of Medical Things,” Sci, vol. 7, 2025.

[3] K. M. Al-Obaidi et al., “A review of using IoT for energy efficient buildings and cities: A built environment perspective,” Energies, vol. 15, 2022.

[4] M. Javaid et al., “Upgrading the manufacturing sector via applications of Industrial Internet of Things (IIoT),” Sensors International, vol. 2, 2021.

[5] A. Khraisat and A. Alazab, “A critical review of intrusion detection systems in the Internet of Things,” Cybersecurity, vol. 4, 2021.

[6] A. Chatterjee and B. S. Ahmed, “IoT anomaly detection methods and applications: A survey,” Internet of Things, vol. 19, 2022.

[7] P. Das, “AI-based network management for IoT devices,” Computational Engineering and Technology Innovations, vol. 1, 2024.

[8] S. H. Rafique et al., “Machine learning and deep learning techniques for IoT network anomaly detection,” Sensors, vol. 24, 2024.

[9] R. Ranpara et al., “A computational framework for IoT security integrating deep learning-based semantic algorithms,” Scientific Reports, vol. 15, 2025.

[10] Y. Liu et al., “Machine learning for the detection and identification of IoT devices,” IEEE Internet of Things Journal, vol. 9, 2021.

[11] M. Noaman et al., “Challenges in integration of heterogeneous IoT,” Scientific Programming, vol. 2022, 2022.

[12] G. Kołaczek, “Internet of Things (IoT) technologies in cybersecurity,” Applied Sciences, vol. 15, 2025.

[13] A. Duraj et al., “Detection of anomalies in data streams using the LSTM-CNN model,” Sensors, vol. 25, 2025.

[14] F. Kateb et al., “Improved security for IoT-based healthcare systems using deep learning,” Scientific Reports, vol. 15, 2025.

[15] X. Liu et al., “Federated learning-oriented edge computing framework for IIoT,” Sensors, vol. 24, 2024.

[16] Q. Abu Al-Haija and S. Zein-Sabatto, “Deep-learning-based detection and classification system for cyber-attacks,” Electronics, vol. 9, 2020.

[17] M. Hasan et al., “Attack and anomaly detection in IoT sensors using machine learning,” Internet of Things, vol. 7, 2019.

[18] A. Malki et al., “Machine learning approach for anomaly detection and forecasting,” Alexandria Engineering Journal, vol. 61, 2022.

[19] M. Shafiq et al., “CorrAUC: Bot-IoT traffic detection using ML,” IEEE Internet of Things Journal, vol. 8, 2020.

[20] M. S. Sawah et al., “DDoS classification based on random forest,” Scientific Reports, vol. 15, 2025.

[21] S. H. Rafique et al., “Machine learning and deep learning techniques for IoT network anomaly detection,” Sensors, vol. 24, 2024.

[22] M. Balega et al., “Enhancing IoT security: Optimizing anomaly detection through machine learning,” Electronics, vol. 13, 2024.

[23] H. Ma et al., “An IoT intrusion detection framework based on feature selection and large language models,” Scientific Reports, vol. 15, 2025.

[24] E. Altulaihan et al., “Anomaly detection IDS for detecting DoS attacks in IoT networks,” Sensors, vol. 24, 2024.

[25] K. Rahman et al., “Cognitive lightweight logistic regression-based IDS for IoT-enabled FANET,” Mobile Information Systems, vol. 2023, 2023.

[26] T. Zhukabayeva et al., “An edge-computing-based framework for intrusion detection in IIoT,” Sensors, vol. 25, 2025.

[27] M. L. Ali et al., “Deep learning vs. machine learning for intrusion detection,” Applied Sciences, vol. 15, 2025.

[28] L. Best et al., “A hybrid approach using K-means and Naive Bayes for IoT anomaly detection,” arXiv, 2022.

[29] J. Sharma et al., “Enhancing IoT anomaly detection with DBSCAN,” Springer, 2024.

[30] C. Koball et al., “IoT device identification using unsupervised machine learning,” Information, vol. 14, 2023.

[31] D. Alsalman, “A comparative study of anomaly detection techniques for IoT security,” IEEE Access, vol. 12, 2024.

[32] L. Vigoya et al., “ML algorithms for validation of IoT anomaly detection dataset,” Applied Sciences, vol. 13, 2023.

[33] I. Mukherjee et al., “Simulation and modeling for anomaly detection in IoT,” Int. J. Wireless Inf. Networks, vol. 30, 2023.

[34] I. Cvitić et al., “Ensemble ML for classification of IoT devices,” Int. J. Machine Learning and Cybernetics, vol. 12, 2021.

[35] T. Talaei Khoei and N. Kaabouch, “Comparative analysis of supervised and unsupervised IDS models,” Information, vol. 14, 2023.

[36] K. Al Jallad et al., “Anomaly detection optimization using big data and deep learning,” Journal of Big Data, vol. 7, 2020.

[37] M. Kodyš et al., “Intrusion detection in IoT using CNN,” IEEE, 2021.

[38] M. R. Islam et al., “Deep learning-based IoT system for health monitoring,” Sensors, vol. 23, 2023.

[39] R. Lazzarini et al., “Stacking ensemble of deep learning models for IoT intrusion detection,” Knowledge-Based Systems, vol. 279, 2023.

[40] C. Li et al., “An anomaly detection approach based on LSTM for IoT,” Security and Communication Networks, 2023.

[41] R. Alghamdi and M. Bellaiche, “Ensemble deep learning-based IDS for IoT,” Cybersecurity, vol. 6, 2023.

[42] A. F. Otoom et al., “Deep learning for brute-force attack detection in IoT,” Procedia Computer Science, vol. 220, 2023.

[43] S. Gopali and A. Siami Namin, “Deep learning-based time-series anomaly detection in IoT,” Electronics, vol. 11, 2022.

[44] I. Ullah and Q. H. Mahmoud, “Deep learning-based model for anomaly detection in IoT,” IEEE Access, vol. 9, 2021.

[45] A. Gueriani et al., “Enhancing IoT security with CNN and LSTM-based IDS,” arXiv, 2024.

[46] S. A. Bakhsh et al., “Enhancing IoT security through deep learning-powered IDS,” Internet of Things, vol. 24, 2023.

[47] S. Tanzila et al., “Securing IoT systems in smart cities using deep learning,” Discrete Dynamics in Nature and Society, 2022.

[48] S. Markkandeyan et al., “Hybrid deep learning-based cyber security threat detection model,” Cyber Security and Applications, vol. 3, 2025.

[49] N. Alkhafaji et al., “Genetic algorithm and deep learning for cyber-attack detection in IIoT,” Arabian Journal for Science and Engineering, 2024.

[50] S. Jamshidi et al., “Machine learning techniques in IDS for IoT,” arXiv, 2025.

[51] A. Amouri et al., “Enhancing intrusion detection using ensemble approach,” IEEE, 2024.

[52] R. R. N. Al Ogaili et al., “Malware detection using whale optimization algorithm,” Wireless Networks, vol. 30, 2024.

[53] A. K. Dey et al., “Metaheuristic-based ensemble feature selection for cyber threat detection,” Decision Analytics Journal, vol. 7, 2023.

[54] S. Latif et al., “DTL-IDS: Intrusion detection using deep transfer learning,” Journal of Network and Computer Applications, vol. 221, 2024.

[55] E. P. Krishna and A. Thangavelu, “Attack detection using hybrid metaheuristic algorithms,” Int. J. System Assurance Engineering and Management, 2021.

[56] Y. Yin et al., “Hybrid feature selection for intrusion detection,” Journal of Big Data, vol. 10, 2023.

[57] T. Saba et al., “Intrusion detection system using advanced ML for IoT,” IT Professional, vol. 23, 2021.

[58] S. Sagu et al., “Metaheuristic optimization algorithms for secure IoT,” Sustainability, vol. 15, 2023.

[59] P. Dakic et al., “Intrusion detection using metaheuristic optimization in IoT systems,” Scientific Reports, vol. 14, 2024.

[60] S. S. Kareem et al., “Feature selection model using hybrid metaheuristic algorithms for IoT IDS,” Sensors, vol. 22, 2022.

[61] L. Almuqren et al., “Hybrid metaheuristics with ML-based botnet detection in IoT,” IEEE Access, vol. 11, 2023.

[62] R. Alkanhel et al., “Network intrusion detection based on feature selection and hybrid metaheuristic optimization,” Computers, Materials & Continua, vol. 74, 2023.