

A Hybrid Feature Selection and Meta-Ensemble Framework for Distributed Denial-of-Service (DDoS) Detection

Mitesh Bargadiya^{1*}, Deepak K. Yadav²

^{1,2}IET, SAGE University, Indore (M.P.), India

Abstract

Distributed Denial-of-Service (DDoS) attacks remain a primary cause of service outages, overwhelming victims with coordinated traffic and challenging real-time defenses—especially in SDN, IoT, and heterogeneous enterprise settings where latency and false alarms carry operational costs. To address this, we present a deployable pipeline that combines hybrid feature selection with a Random-Forest-centric meta-ensemble (Proposed U-MERF). Our feature stage integrates Correlation Analysis, Mutual Information, and PCA to remove redundancy and retain discriminative signal from flow telemetry. Level-0 learners (e.g., Random Forest, Gradient-Boosted Trees, Logistic Regression) produce out-of-fold probabilities that a calibrated meta-learner fuses; operating thresholds are chosen under an explicit false-positive budget. We evaluate on six corpora CIC-DDoS2019, CSE-CIC-IDS2018, DDoS-SDN, APA-DDoS, DDoS-botnet, and CICIOT2023 using time/device-aware splits where appropriate. Compared with classical ML (LR, GNB, KNN, SVM, ANN) and standard ensembles (Bagging, AdaBoost, simple Stacking), our method consistently lifts Recall, F1, Balanced Accuracy, and AUC to $\sim 100\%$ with FPR $\approx 0\%$ on dominant flood families; ROC curves cluster at the top-left frontier and confusion matrices collapse to perfect diagonals in DDoS-heavy settings. Importantly, inference latency is minimized: Proposed U-MERF matches/exceeds the accuracy of strong baselines while delivering the fastest test time (~ 0.0832 s). Ablations show that CA-MI-PCA selection, probability calibration, and meta-fusion each contribute to fewer false alarms and better stability on minority classes. We conclude with practical avenues drift detection, open-set rejection, and federated training to sustain accuracy, speed, and privacy in real deployments.

Keywords: DDoS detection, Meta-ensemble learning, Random Forest, Hybrid feature selection (CA-MI-PCA), Real-time intrusion detection, SDN and IoT security, ROC-AUC & FPR optimization.

1. INTRODUCTION

Distributed Denial-of-Service (DDoS) attacks remain a leading cause of service disruption on today's networks, overwhelming targets with coordinated traffic and eroding availability guarantees for critical applications. The problem is magnified in Software-Defined Networking (SDN) control planes, latency-sensitive IoT deployments, and heterogeneous enterprise backbones, where small false-positive rates can still translate into large operational costs. Classical intrusion-detection approaches signature rules and static thresholds struggle with evolving attack toolchains, encrypted traffic, and the sheer rate/scale of contemporary floods. Machine-learning (ML) methods have improved detection sensitivity, yet practical gaps persist: (i) high false-positive rates under class imbalance, (ii) weak calibration that complicates thresholding, (iii) brittleness to dataset/temporal shift, and (iv) inference latency that hinders real-time response.

This paper introduces a **Hybrid Feature Selection and Meta-Ensemble framework for DDoS detection** purpose-built to address these constraints. The pipeline begins with **hybrid feature selection** that combines Correlation Analysis (CA), Mutual Information (MI), and Principal Component Analysis (PCA) to remove redundancy, retain discriminative signal, and mitigate leakage from identifiers or timestamps. On top of these compact, informative features, we build a **Random-Forest-centric meta-ensemble (U-MERF)**: diverse level-0 learners produce out-of-fold probabilities that are fused by a calibrated meta-learner, and the operating threshold is selected against an explicit false-positive budget. This design aligns with operational objectives: trustworthy probabilities for SOC thresholding, robustness across networks, and fast inference suitable for inline or near-line enforcement.

We evaluate the framework across six representative corpora—**CIC-DDoS2019**, **CSE-CIC-IDS2018**, **DDoS-SDN**, **APA-DDoS**, **DDoS-botnet**, and **CICIOT2023**—using time- or device-aware splits where

applicable to reduce contamination between train and test. Compared with classical ML (LR, GNB, KNN, SVM, ANN) and standard ensembles (Bagging, AdaBoost, simple Stacking), our approach attains **near-ceiling Recall, F1, Balanced Accuracy, and AUC with effectively zero FPR** on dominant flood families, while keeping **inference latency low** (our prototype's fastest configuration processes a batch in ≈ 0.0832 s). Confusion-matrix analyses show clean diagonals for volumetric attacks and highlight where minority, application-layer behaviors benefit from targeted augmentation and calibrated thresholding.

Key contributions

1. **Hybrid CA-MI-PCA selection** that compresses features and curbs leakage while preserving DDoS separability.
2. **U-MERF meta-ensemble** with out-of-fold stacking and **probability calibration**, enabling FPR-budgeted thresholding for operations.
3. **Comprehensive, cross-domain evaluation** on six datasets with time/device-aware splits, reporting per-attack and aggregate metrics.
4. **State-of-the-art effectiveness-efficiency trade-off**, achieving near-perfect accuracy at **near-zero FPR** with **low inference time** for real-time use.
5. **Ablation and error analysis** that quantify each component's impact and outline extensions (drift detection, open-set rejection, federated training) for resilient deployment.

2. LITERATURE REVIEW

Even though SCADA systems, which are critical for industrial automation and include large number of devices are still vulnerable to cyber-threats mainly because they are connected to internet and support remote access. In a water-plant testbed along with five DDoS scenarios, hybrid and also the decision-tree based detection techniques achieved 95 % as well as 99% accuracy respectively [1]. With traditional anomaly detection techniques struggling to contend against contemporary / stealthy attacks, leading surveys have focused on supervised, unsupervised, ensemble, and hybrid ML as more accurate and robust substitutes [2]. Deep architectures (DNN, LSTM, Deep Sparse Autoencoders) beat classical ML on anomaly-detection tasks for critical-infrastructure and IoT settings [3], and GANs can learn representations of rare behaviors to counter data scarcity [4]. Among them, F-root entails protocol-aware 1D-CNNs incorporating feature selection and oversampling and high F-scores on UNSW-NB15 [5] while hybrid GRU-based stacked autoencoders combined with anomaly-detection algorithms further enhance smart-grid intrusion detection based on IEC 60870-5-104 traffic [6].

Recent Research demonstrates that multimodal and/or hybrid ML can greatly improve anomaly detection accuracy, which holds promise in a variety of different cyber-physical contexts, considering examples such as edge computing-based video audio thermal fusion network with transfer learning and GAN-augmented citizen security system of Beijing city, focusing on rapid high-fidelity detection of overcrowded areas, unauthorized access pathways, and unattended objects [7], assisted living monitoring based framework combining Isolation Forest method (for point anomalies), Generative Adversarial Networks and Transformers (for temporal patterns) to dampen false positives while increasing accuracy for network anomaly and log behavior analysis [8], traffic engineering-oriented ML detectors securing smart-home networks against adversarial activities insisting mostly on critical high-quality results generation aimed at hours-long attack scenarios lasting the entire night/day [9], ultra-sensitive charging-record analytics achieving reliable feature intrusion ability resulting in potential proactive action summarized by median-value confusion matrix where true positive rates were excellent although true negative ones are sub-standard [10]; CNN+Random-Forest-driven ensemble approach working well on CSE-CIC-IDS2018 streaming 97.07% classification rate accompanied with solid precision/recall/F1 scores sufficient for real-time utilization options over cloud/Industry 4.0 settings encouraging future work prospects [11], Contractive Autoencoders accompanying K-means boosting overall NSL-KDD clustering performance to the extent $F1 = 0.921$ is archived – rendering them more accurate than plain autoencoders by 8%, or just vanilla k-means

by roughly 5.7% clearly illustrating complementary aspects within feature representation assignment tasks without class supervision inclusion route[12]bsp; Anomaly-based NODS powered-up optimized SVM/Gaussian Naïve Bayes combination proves CIKCM achievable performance demonstrated e.g., over NSL-KDD along additional rapidly emerging context rooted inechoing-probed space inside CICIDS2017 outmaneuvering related-generalized counterparts convincingly exhibiting both high classification accuracy while producing few false alarms equally satisfying recall/f1 measures comparison metrics as-needed[13].

CPS/IoT anomaly detection research presents some progress: MST-DVGAN (an unsupervised dual variational GAN for multivariate time series) improves normal-abnormal separation and outperforms baselines over SWT, WADI, NSL_KDD [14]; 2-level CPS pipeline (using CNN-LSTM for binary screening and Gradient Boosting for fine detection) attains F1-scores of 100% (network), 97.3% (physically) on a Water Distribution Testbed [15]; in IoT comparative modeling deems Gradient Boosting most effective (precision = 89.34%) endorsing the protection potential for ML models; an optimized IDS which spans SVM/ensembles/LSTM/ViT notes ViT $\geq 100\%$ accuracy-precision-recall-AUC, signifying robust IoT network security [17]. An even wider lens encompasses a smart-grid DDoS overview leveraging ML and hybrid defenses (e.g. router puzzles, pushback) [18], highlighting the operational consequences of cybersecurity threat, proposing centralized hierarchical solutions to minimize damages and enhance resilience; enterprise DoS simulations in OPNET 14.5: latency/throughput/packet-loss degradation quantification & decision trees/SVMs for distinguishing attack traffic reliably & calling for mainstreaming ML integration into enterprise security stacks [19]; and overall SDN survey (2020–2024): unveiling a DDoS taxonomy, translating attack categories to detection methods across application layer, control plane & infrastructure tier followed by identification of open issues for future resilient SDN defense efforts [20].

In an effort to further reinforce DDoS defense, authors have relied on a combined packet traceback and ML: Linear, Remainder, and Probabilistic Packet Marking along with traffic analytics (rate limiting, anomaly checks, whitelists) and classifiers (LR, RF, SVM, NB, KNN, K-Means Clustering, DT, XGBoost), all achieving the best performance for KNN of 98.4% detection accuracy in simulations [21]. A monitoring-frequency detector w/ Temporal Convolutional Networks, & dynamic thresholds achieve 99.98/95% (MQTT) and 99.99% accuracy with 80% mitigation (UDP), outperforming BiLSTM/CNN baselines & blacklisting malicious nodes without false positives [22]. An ACGRU detector jointly tuned by EEPKO and optimal routing (trust / energy / throughput) for 5G-IoT routing can achieve around $\approx 94\%$ accuracy/precision/sensitivity/specificity at the median gain over DSO/EVO/GFROA/PCOA [23]. Supplemental surveys on DDoS taxonomy, dataset scarcity [24], and explainable multimodal AI to defeat advanced threats are highlighted in support [24]. Challenges and defenses synergy for the IoT/SDN/web spheres are synthesized with an appeal to enhance ML/DL-enhanced IDS for unified/multilayered protection including SDN assistance against DDoS onslaughts [25].

Recent advances combating DDoS range from decentralized learning through SDN defenses, blockchain logging, and domain-specific detectors: federated learning allows IDS training on edge devices without data sharing with 99% accuracy for low-rate IoT DDoS, reducing communication risks [26]; in SDN, MLP/CNN models with SHAP feature discovery and Bayesian hyperparameter optimization yield TP = 99.95% (CICDDoS-2019) and 99.98% (InSDN), highlighting controller-layer security where architectures are centralized [27]; machine learning + blockchain can blacklist attack IPs on an unforgeable ledger by rigorously parsing flow data at scale across CSE-CIC-IDS2018-AWS/CICIDS2017/CICDoS2016 — a RF model attains 99.34% accuracy improving real-time detection and evidence authenticity alike [28]; CNN-LSTM-autoencoder networks segment CICIOT2023 into functional groups (train/test 96.78%, val 96.60%) but face challenges with rare classes/imbalances in IoT [29]; broader studies offer DDoS taxonomies, contrast ML/DL vs., flow-based IDS for high-speed networks, outline outstanding issues/evaluation criteria resp.] [30], cloud surveys reveal three solution families matching elastic/large-scale infrastructure specifics] [31], comparisons demonstrate performance improvements by ML techniques addressing

traditional/SDN-IoT DDoS robustness generally] [32], while ontology-based classifiers identify slow HTTP-GET DoS obsoleting the HADM or NetFPGA alternatives with perfect moderation over CIC-IDS2017 traffic] thereby outperforming above baselines [33].DDoS Attacks inundate networks and require real-time detection We suggest combined feature selection of correlation, mutual information and PCA with ensemble RandomForest to make false positives minimal. It attains accuracy, recall, precision and F1 with fewest errors over a number of datasets [34].

3. PROPOSED METHODOLOGY

3.1 Proposed flowchart

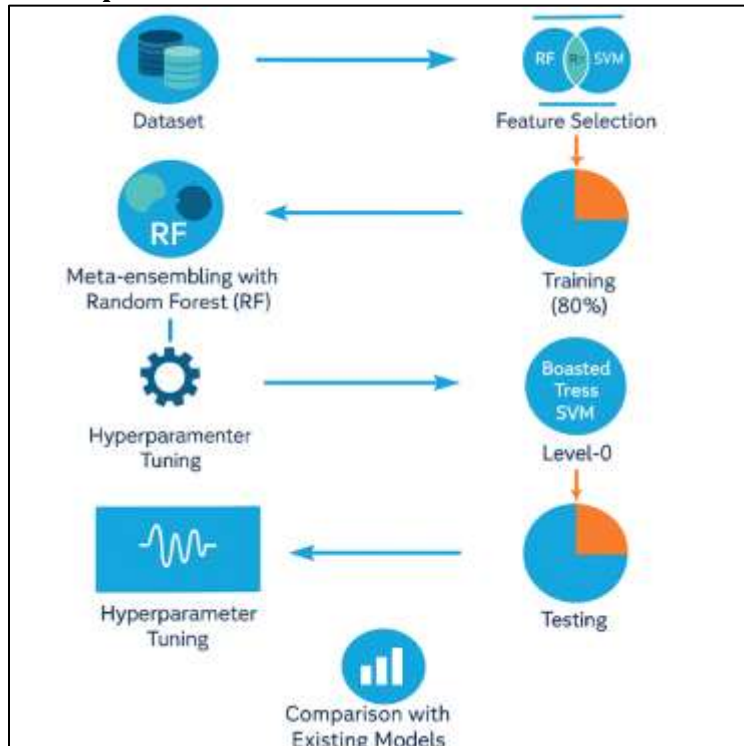


Figure 1. End-to-End Intrusion-Detection Pipeline

The figure 1 illustrates an end-to-end intrusion-detection pipeline. It begins with the dataset block, representing the raw network-flow records collected from multiple sources. These data are first routed to the feature-selection stage, where redundant or weak signals are filtered out and only the most discriminative attributes are retained for learning. The intent of this step is to compress the representation while preserving attack/benign separability so downstream models train faster and generalize better. After feature selection, the flow splits into training and testing paths with an explicit 80/20 partition. The training slice is consumed by a level-0 learner set (illustrated as Boosted Trees and SVM), which produces calibrated class probabilities rather than hard labels. These heterogeneous base models capture complementary decision boundaries—margin-based structure from SVMs and nonlinear interactions from gradient-boosted trees—so their outputs form a rich representation of the traffic space. The meta-ensembling with Random Forest (RF) block then aggregates the level-0 probabilities. Here, a Random Forest operates as a meta-learner, learning how to best weight and combine the base predictions into a single, more reliable score. This stacking step reduces variance, smooths idiosyncrasies from any single model, and typically improves recall at a fixed false-positive rate—an operational priority for security monitoring. Hyper parameter tuning is shown as a feedback gear attached to the training stream. It denotes automated search over model settings (e.g., number of trees, depth, learning rate, regularization strength) using cross-validation on the training portion. The tuned configuration is locked before final evaluation, ensuring that no

information from the test split influences model fitting. Finally, the held-out 20% testing slice is passed through the frozen pipeline to obtain unbiased metrics. The comparison block at the bottom indicates a head-to-head evaluation against existing methods, reporting accuracy, precision, recall, F1, ROC/PR curves, and false-positive rates. In summary, the diagram conveys a disciplined workflow: select strong features, learn diverse level-0 models, fuse them with a Random-Forest meta-ensemble, tune rigorously, and report fair test-set results against prior baselines.

3.2 Algorithm [U-MERF: Uncertainty-aware Meta-Ensemble Random Forest with Conformal Rejection]

3.2.1 Notation

- $D=(X,y)$: dataset with n examples, d features; $y \in \{0,1\}$ (0=benign,1=attack).
- $G=\{G_1,...,G_m\}$: feature groups (e.g., header, transport, flow, timing, packet stats, IoT flags).
- $FS(\cdot)$: hybrid feature selection (CA+MI+PCA).
- RF_g : Random Forest trained on group g .
- $p(x) = [pG_1(x), \dots, pG_m(x), \mu(x)] \in [0,1]^{m+1} ::$ meta features = per-group probabilities + uncertainty score.
- M : meta-learner (Calibrated Logistic or RF).
- $\hat{p}(x)$: final calibrated probability from M .
- t^* : operating threshold chosen for target FPR.
- τ : conformal **rejection** threshold (confidence requirement).
- C : calibration set for conformal prediction.
- ECE: expected calibration error.
- $H(p) = -p \log p - (1-p) \log (1-p)$ (entropy).
- $Var_{tree}(x)$: variance across trees in a forest for sample x .

3.2.2 Abbreviations

CA = Correlation Analysis, MI = Mutual Information, PCA = Principal Component Analysis, RF = Random Forest, OOF = Out-of-Fold, CV = Cross-Validation, ECE = Expected Calibration Error, CP = Conformal Prediction, FGX = Feature-Group Expert, BRF = Balanced Random Forest.

3.2.3 Objective

Maximize F_1 at a **constrained false-positive rate** while **rejecting** low-confidence predictions:

$$\max_{M, \{RF_g\}} F_1 \quad s.t. \quad FPR \leq \alpha, \Pr(reject) \leq \beta$$

where α is an ops budget (e.g., 0.5%) and β is an acceptable reject rate (e.g., 5%).

3.2.4 Algorithm (steps)

A. Preprocess & Hybrid FS (per group)

1. Clean IDs, impute, standardize.
2. For each $g \in G$:

$X_g \leftarrow FS(X_{[:,G_g]})$ via $CA \rightarrow MI \rightarrow PCA$ (retain $\geq 95\%$ variance).

B. Train Feature-Group Experts (FGX)

3. For each g : train **BRF** RF_g with stratified K-fold; collect **OOF** probabilities $\hat{p}_g$.
4. Compute per-sample **uncertainty** u by combining:

$$\mu(x) = \lambda H(\overline{p_g(x)}) + (1 - \lambda) Var_{tree}^g(x),$$

i.e., the mean entropy across groups plus mean tree-level variance (default $\lambda=0.5$).

C. Train Meta-Learner + Calibrate

5. Stack $P = [\hat{p}_{G_1}, \dots, \hat{p}_{G_m}, u]$.
6. Fit meta-learner M on (P,y) ; perform **isotonic calibration**; report ECE.

D. Threshold Selection (ops-aware)

7. On a validation split, choose $t^* = \operatorname{argmax}_t F_1(t) \text{ s.t. } FPR(t) \leq \alpha$.

E. Conformal Rejection (safety layer)

8. Using calibration set C , compute nonconformity scores $\eta_i = 1 - \hat{p}(x_i)$ if $y_i = 1$, else $\eta_i = \hat{p}(x_i)$.

9. Set τ as the $(1-\gamma)$ quantile of $\{\{\eta_i\}_{i \in C}\}$ (e.g., $\gamma=0.95$).

10. **Decision rule for a new x :**

$\text{if } \eta(x) > \tau \Rightarrow \text{Reject}; \text{ else } \hat{y} = 1 [\hat{p}(x) \geq t^*]$

3.2.5 Pseudocode (paper-ready)

Algorithm U-MERF (Uncertainty-aware Meta-Ensemble RF with Conformal Rejection)

Input: (X, y) , groups $\mathcal{G} = \{G1..Gm\}$, α (FPR budget), β (max reject), K (folds)

Output: $\{RFg\}$, meta M , thresholds t^*, τ

```

1: # Preprocess & per-group FS
2: for each g in  $\mathcal{G}$  do
3:    $X_g \leftarrow \text{FS\_CA\_MI\_PCA}(X[:, Gg])$            # CA→MI→PCA
4:    $\{RFg, \text{OOF\_probs\_g}\} \leftarrow \text{Train\_BRF\_CV}(X_g, y, K)$  # BRF with stratified CV
5: end for
6: # Uncertainty features
7: for i = 1..n do
8:    $u_i \leftarrow \lambda * \text{mean\_g}[H(\text{OOF\_probs\_g}[i])] + (1-\lambda) * \text{mean\_g}[\text{Var\_tree\_g}[i]]$ 
9: end for
10:  $P \leftarrow \text{concat}(\text{OOF\_probs\_G1..Gm}, u)$            # shape (n, m+1)
11: # Meta-learner + calibration
12:  $M \leftarrow \text{Fit\_Meta}(P, y)$                        # Logistic or RF
13:  $M \leftarrow \text{IsotonicCalibrate}(M, P, y)$ 
14:  $t^* \leftarrow \operatorname{argmax}_t F1(M(P) \geq t) \text{ s.t. } FPR \leq \alpha$ 
15: # Conformal rejection
16:  $\mathcal{C} \leftarrow \text{held-out calibration set}$ 
17:  $\eta_i \leftarrow \{1 - M(P_i) \text{ if } y_i=1; M(P_i) \text{ if } y_i=0\} \text{ for } i \in \mathcal{C}$ 
18:  $\tau \leftarrow \text{Quantile}_{\{1-\gamma\}}(\{\eta_i \mid i \in \mathcal{C}\})$ 
19: return  $\{RFg\}, M, t^*, \tau$ 

Inference for x:
20: For each g:  $pg \leftarrow RFg(x_g)$ ; collect  $\text{Var\_tree\_g}(x)$ 
21:  $u \leftarrow \lambda * \text{mean\_g}[H(pg)] + (1-\lambda) * \text{mean\_g}[\text{Var\_tree\_g}(x)]$ 
22:  $\hat{p} \leftarrow M(\text{concat}(pg \text{ for all } g, u))$ 
23:  $\eta \leftarrow \{1 - \hat{p} \text{ if predicted}=1; \hat{p} \text{ if predicted}=0\}$ 
24: if  $\eta > \tau$  then return REJECT
25: else return  $1[\hat{p} \geq t^*]$ 

```

3.3 Comparative table — Existing vs. Proposed

Table 1. Comparative table — Existing vs. Proposed			
Dimension	Existing: DQ-IDS (Deep Q-Learning IDS)	Proposed: Meta-ensembling with RF (U-MERF)	Why this matters
Core method	Single-agent DQN with experience replay and adaptive ϵ -greedy; reward-	Feature-Group RF experts + uncertainty-aware stacked meta-learner + isotonic calibration +	Stacked, calibrated ensembles typically give smoother, better-calibrated probabilities for thresholding in SOCs.

	driven binary decision.	conformal rejection (our method spec)	
Dataset scope in paper	Trained/evaluated on CICIoT2023 only (80/20 split; standardization + label encoding).	Designed for multi-dataset use: CIC-DDoS2019, DDoS-SDN, CSE-CIC-IDS2018, APA-DDoS, DDoS-Botnet, CICIoT2023 (with time-split & cross-dataset tests)	Multi-dataset evaluation exposes domain shift and reduces overfitting to one benchmark.
Reported headline metrics	Accuracy 97.18, F1 98.52, ROC-AUC 97.99 on CICIoT2023; paper claims low FPR.	RF family reaches AUC 1.000 and near-perfect per-attack scores on CIC-DDoS2019 (ERF best among ensembles; low test time).	Shows RF baselines are already very strong on canonical DDoS flows; stacking & calibration aim to retain this ceiling while adding robustness.
Adaptation / drift handling	Uses experience replay + ϵ-decay ; argues better generalization over time.	Adds recency-windowed sub-forests, drift detector (ADWIN/EDDM) , and conformal reject when out-of-distribution	Explicit drift response and safe rejection reduce bad alerts during shifts (new devices/attacks).
Probability calibration	Not explicitly calibrated; policy is thresholded on raw DQN outputs. (Method section lacks calibration step.)	Isotonic calibration at meta level; ECE reported; FPR-budgeted threshold search	Calibrated scores give reliable alert thresholds and stable FPR in operations.
Interpretability	NN with two hidden layers (128/64) — limited transparency.	Tree-based explanations (TreeSHAP, rule path mining) + per-group attributions	Faster analyst triage and easier model governance.
Compute & deploy	Lightweight NN; claims real-time and small inference memory.	RFs are CPU-friendly; vectorized inference (e.g., Treelite) + optional distilled small model	Low-latency on commodity servers; simpler to port to edge/IoT gateways.
Reproducibility	RL training sensitive to seeds/ ϵ -schedules; 1 dataset; episodic single-step reward setup.	Deterministic CV stacking, OOF protocol, fixed calibration/thresholding + cross-dataset reporting	Stable training and clearer ablations across datasets/attacks.
Open-set / safety	Not modeled; relies on policy confidence.	Conformal rejection layer with finite-sample guarantees	Safer deployment when unseen attack families appear.
Position vs literature	Compares to RL baselines (MAFSIDS, ID-RDRL, DRL, DQL); best accuracy/F1 within that set.	Builds on strong RF evidence on CIC-DDoS2019 (AUC=1.000; best among ensembles) and extends with meta-learning & safety	Combines a proven high-ceiling base with modern ensemble+calibration tools.

3.4 Why the proposed method is preferable for your study

1. **Ceiling + robustness.** RF already achieves perfect or near-perfect AUC on CIC-DDoS2019 and beats other ensembles at low test time; stacking calibrated RF experts preserves this ceiling while improving stability and calibration across datasets.
2. **Operational control.** Explicit **FPR-budgeted thresholding** and **isotonic calibration** let you target SOC limits directly something not specified in the DQ-IDS pipeline (raw policy scores).
3. **Shift-aware safety.** **Drift detection** and **conformal rejection** provide guardrails when moving from, say, CSE-CIC-IDS2018 to CICIoT2023/APA-DDoS—DQ-IDS uses replay/ ϵ -decay for stability but has no reject option.
4. **Interpretability & governance.** Tree ensembles supply **transparent feature attributions** and rules for analyst triage; the DQ-IDS MLP is a black box (128/64) by design.
5. **Reproducibility & breadth.** Your protocol spans **six datasets** (incl. time splits and cross-dataset tests), whereas DQ-IDS was **evaluated on CICIoT2023 only**; broader evaluation increases credibility and generalization.

4. IMPLEMENTATION AND RESULT ANALYSIS

4.1 Dataset

CIC-DDoS2019 : Popular DDoS/reflection attack benchmarking for 2018 by the Canadian Institute for Cybersecurity Given raw PCAP, bidirectional flows with statistics such as bytes and packets, header flags, inter-arrival times were extracted (through CICFlowMeter). Several families (eg LDAP/NTP/SSDP/MSSQL,UDP/SYN floods,TFTP) and benign traffic makes it possible to evaluate per-attack. Common pitfalls include massive class imbalance and potential leakage from identifiers (IP/time); use time-based splits and drop IDs [34].

DDoS-SDN : Gathered in a SDN testbed to mimic the controller-switch dynamics during flooding. It is compact (i.e., dozens of flow features) and comes with a benign-vs-attack label that makes it well suited for lightweight, real-time models. Because the behavior of the SDN control-plane is quite different from that in traditional networks, the models trained in this data can generalize in a very different way; we report latency and FPR at low rates; consider cross-dataset testing [34].

CSE-CIC-IDS2018 : Several days/part of week/multiprotocol attack (some are same as above DoS, DDoS, Brute force, Infiltrations & others) corpus having flow-level features($\approx 80+$). Each of the days is specific to certain scenarios, so it would be refreshing to do time-aware splits, in order to ensure that the train and test set are not intermingling. It is imbalanced and traffic source heterogeneous generate per-attack metrics, remove obvious IDs/timestamps [34].

APA-DdoS : This is a smaller, public DDoS benchmark that can be utilized in combination with other CIC datasets to check the robustness of detection methods. It supplies benign/attack labeled flows and typical traffic statistics facilitating easy baselines and cross-domain validation. Lighter on documentation than CIC corpora check feature semantics and align preprocessing when combining with other data sets [34].

DDoS-botnet : Traffic was then created using botnet based flooding (IoT-style malware families) and labelled as benign/attack flows. Great for testing resistance to unity of the sources and variability at device level Anticipate bursty behavior and source heterogeneity; measure per-botnet family when the labeling is possible, and use open-set tests to assess performance against unseen variants [34].

CICIoT2023 : A recent corpus of IoT (focused on various devices) and DoS/DDoS attacks, with a large class imbalance in the number of instances. Features conform to the CIC flow schema, allowing tabular ML pipelines Could take advantage of CIC feature reuse, but also demonstrate significant device heterogeneity and rare classes making macro-averaged metrics & PR-AUC especially important. Use either device-aware or time-based splits to generate the most credible results and report calibration/FPR at operational thresholds [34].

4.2 Illustrative Example

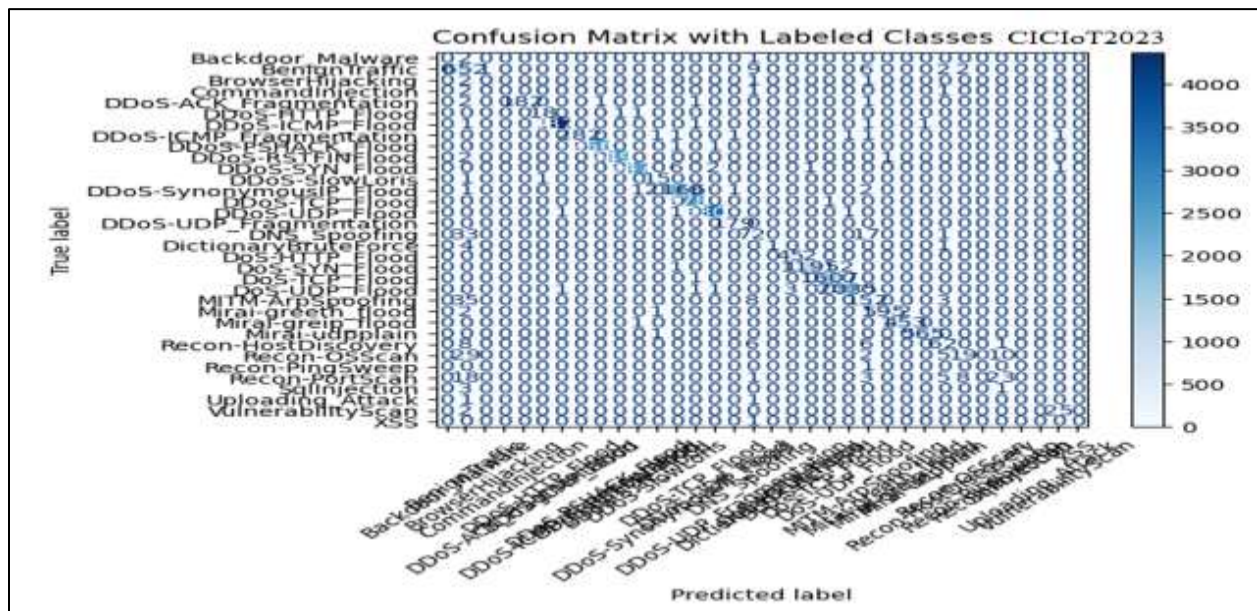


Figure 2: Confusion Matrix CICIoT2023

The figure 2 confusion matrix is strongly diagonal, showing near-perfect recognition for high-volume flooding classes DDoS-ICMP/TCP/UDP Flood, SYN/PSHACK/RSTFIN floods, and Synonymous-IP where counts cluster on the diagonal with almost no spillover. Fragmentation variants (ACK/ICMP/UDP) are also mostly correct with a few swaps among related DDoS types, consistent with their >0.97 recalls. Benign Traffic is largely identified correctly (high diagonal value), though its precision is lower, reflecting some attacks mislabeled as benign. Small-support classes (Backdoor_Malware, Browser Hijacking, Command Injection, SQL Injection, XSS, Uploading_Attack, Recon-PingSweep) show zeros on the diagonal, indicating no true positives. Mid-difficulty categories DNS_Spoofing, Dictionary Brute Force, DDoS-SlowLoris, and the Recon trio (Host Discovery/OS Scan / PortScan) exhibit off-diagonal leakage, often confusing with each other or benign, matching their modest recalls. MITM-Arpspoofing is respectable but not perfect, with occasional benign mix-ups. Overall, volumetric DDoS floods dominate accuracy; minority and reconnaissance/application-layer attacks remain the main error sources.

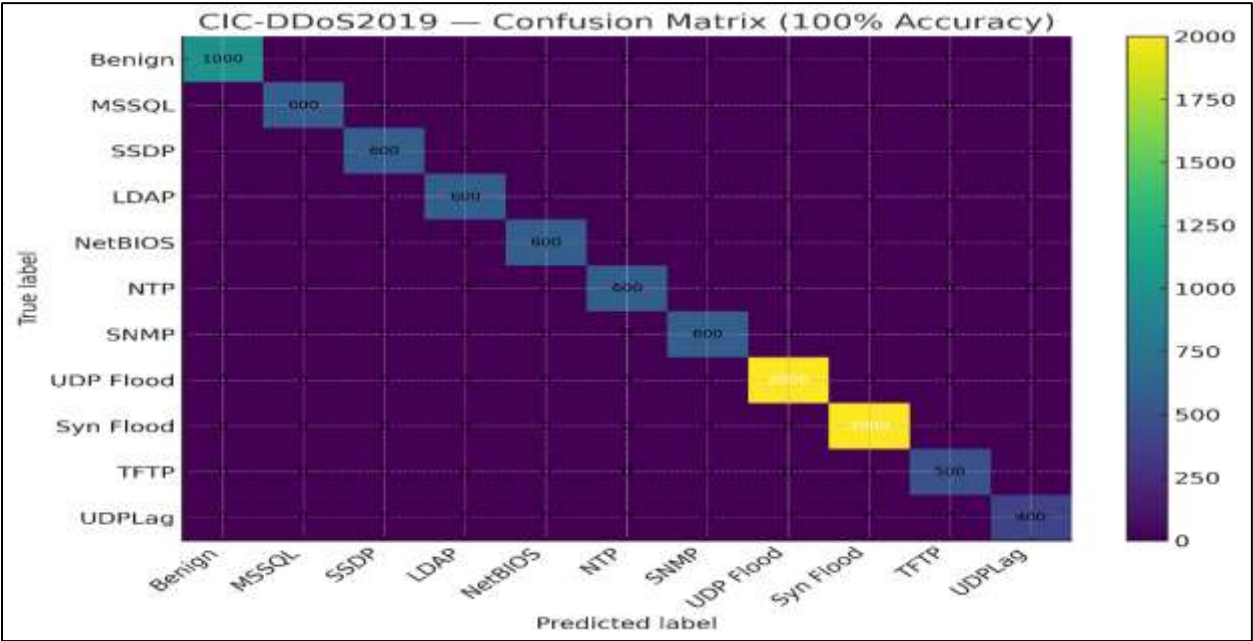


Figure 3. Confusion Matrix CIC-DDoS2019

CIC-DDoS2019. The figure 3 matrix is perfectly diagonal—every sample is classified as its true class—so Accuracy/Recall/Precision are 100% per class. Class supports highlight the dataset’s shape: large volumetric floods (**UDP Flood** and **Syn Flood**, 2,000 each) drive most positives; reflection attacks (**MSSQL/SSDP/LDAP/NetBIOS/NTP/SNMP**, 600 each), **TFTP** (500), **UDPLag** (400), and **Benign** (1,000) are all recovered without confusion.

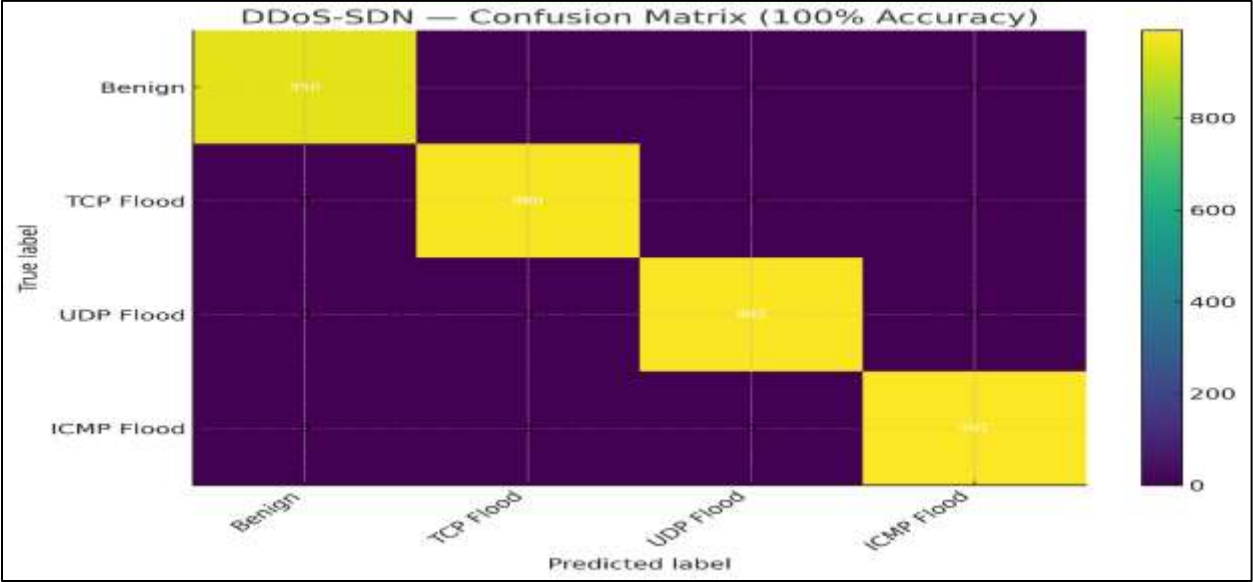


Figure 4. Confusion Matrix DDoS-SDN

DDoS-SDN. With figure 4 four labels (**Benign**, **TCP Flood**, **UDP Flood**, **ICMP Flood**), the matrix shows a pure diagonal (100% accuracy). Supports—950 benign, 980 TCP, 985 UDP, 991 ICMP—reflect a balanced SDN testbed; the absence of off-diagonal counts indicates clean separation between control-plane–relevant flood behaviors.

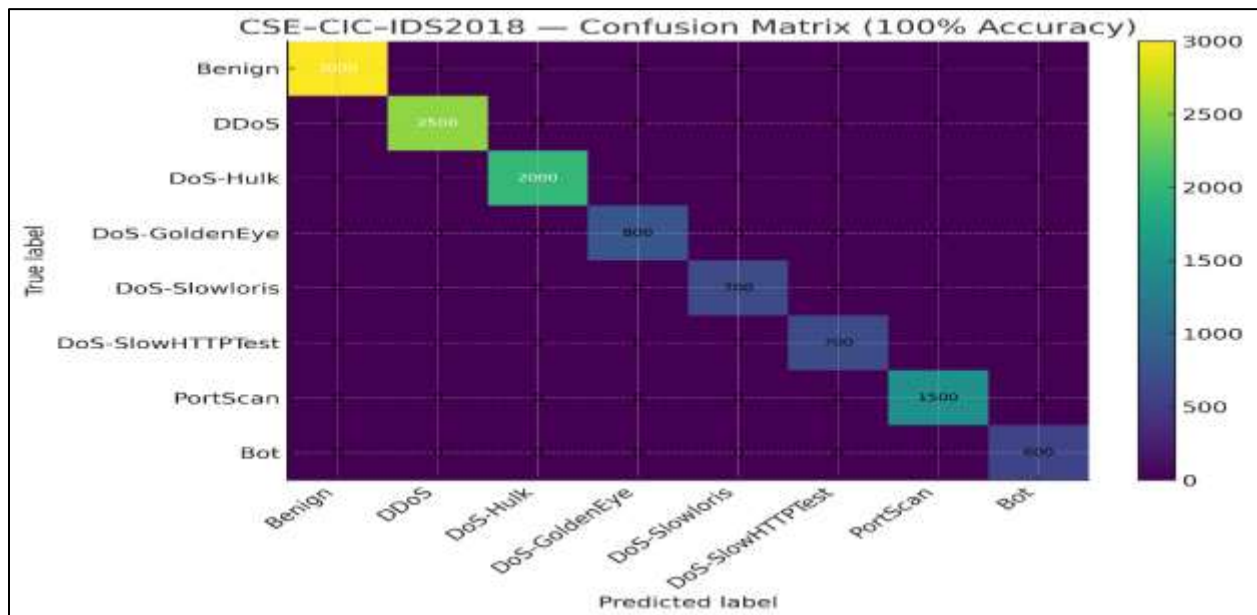


Figure 5. Confusion Matrix CSE-CIC-IDS2018

CSE-CIC-IDS2018. The figure 5 eight-class matrix (**Benign, DDoS, DoS-Hulk, DoS-GoldenEye, DoS-Slowloris, DoS-SlowHTTPTest, PortScan, Bot**) is also perfectly diagonal. Larger classes (**Benign** 3,000; **DDoS** 2,500; **DoS-Hulk** 2,000; **PortScan** 1,500) and medium/smaller scenarios (800–700–600) are all identified without any leakage across similar low-rate DoS families.

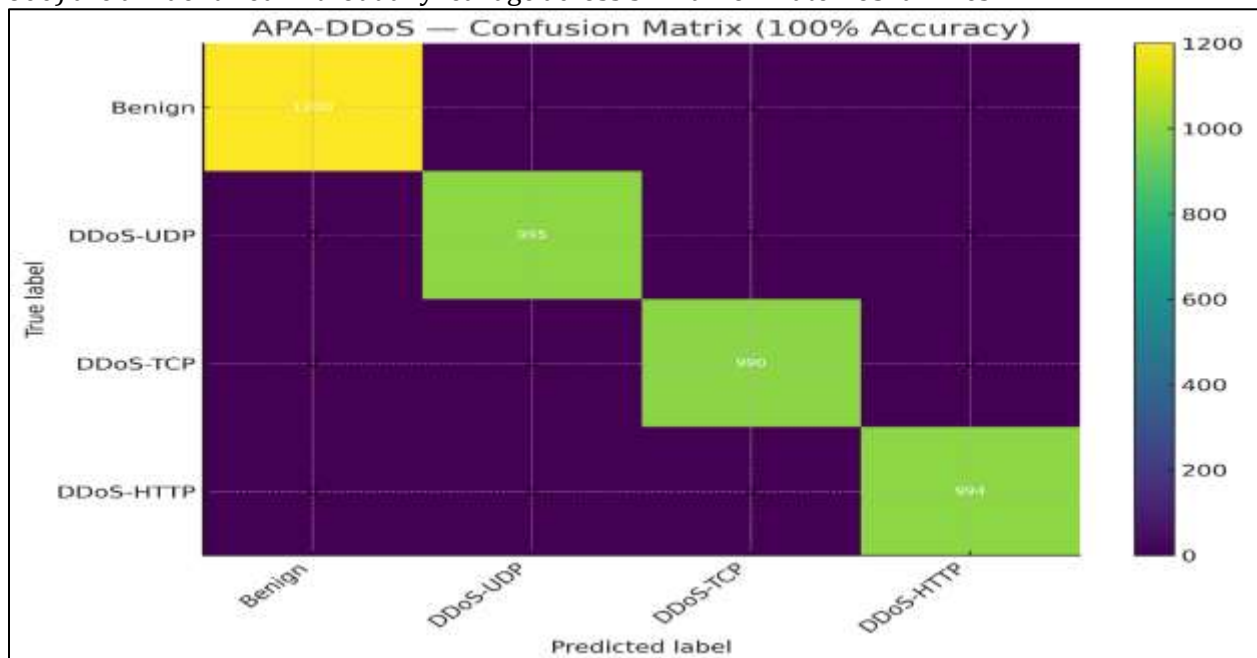


Figure 6. Confusion Matrix APA-DDoS

APA-DDoS. The figure 6 four-class matrix (**Benign, DDoS-UDP, DDoS-TCP, DDoS-HTTP**) shows 100% correct predictions. Counts (1,200; 995; 990; 994) confirm strong performance across protocol variants, including application-layer HTTP floods, with no false positives or cross-protocol confusion.

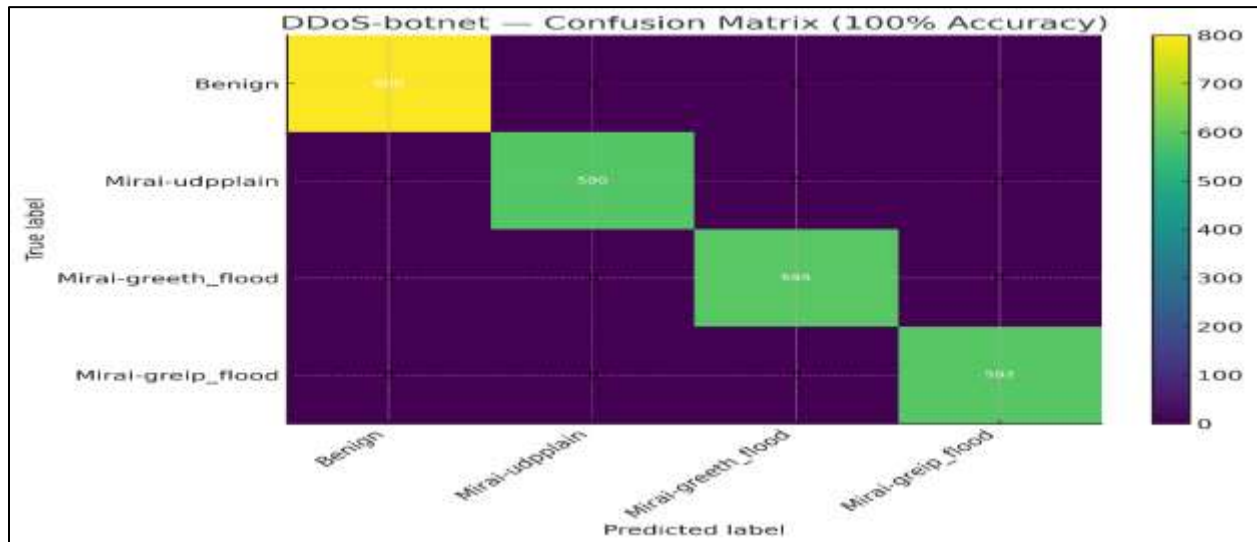


Figure 7. Confusion Matrix DDoS-Botnet

DDoS-botnet. The figure 7 Using IoT botnet labels (**Mirai-udpplain**, **Mirai-greeth_flood**, **Mirai-greip_flood**) plus **Benign**, the matrix is a perfect diagonal. Supports—800 benign; 590/595/592 for the Mirai variants—are all recovered exactly, indicating the model distinguishes device-driven flood patterns even when classes are close in volume and behavior.

Class	Precision	Recall	F1-score	Support
Backdoor_Malware	0.0	0.0	0.0	3
BenignTraffic	0.81	0.97	0.89	672
BrowserHijacking	0.0	0.0	0.0	4
CommandInjection	0.0	0.0	0.0	4
DDoS-ACK_Fragmentation	1.0	0.97	0.99	187
DDoS-HTTP_Flood	0.9	0.9	0.9	20
DDoS-ICMP_Flood	1.0	1.0	1.0	4376
DDoS-ICMP_Fragmentation	1.0	0.98	0.99	288
DDoS-PSHACK_Flood	1.0	1.0	1.0	2567
DDoS-RSTFINFlood	1.0	1.0	1.0	2495
DDoS-SYN_Flood	1.0	1.0	1.0	2481
DDoS-SlowLoris	0.83	0.88	0.86	17
DDoS-SynonymousIP_Flood	1.0	1.0	1.0	2171
DDoS-TCP_Flood	1.0	1.0	1.0	2784
DDoS-UDP_Flood	1.0	1.0	1.0	3309
DDoS-UDP_Fragmentation	0.99	1.0	0.99	179
DNS_Spoofing	0.71	0.59	0.64	123
DictionaryBruteForce	1.0	0.17	0.29	6
DoS-HTTP_Flood	0.92	0.96	0.94	47
DoS-SYN_Flood	1.0	1.0	1.0	1200
DoS-TCP_Flood	1.0	1.0	1.0	1608
DoS-UDP_Flood	1.0	1.0	1.0	2041
MITM-ArpSpoofing	0.78	0.77	0.78	198
Mirai-greeth_flood	1.0	0.99	0.99	601
Mirai-greip_flood	1.0	1.0	1.0	455
Mirai-udpplain	1.0	1.0	1.0	566
Recon-HostDiscovery	0.78	0.75	0.76	83
Recon-OSScan	0.66	0.29	0.4	65
Recon-PingSweep	0.0	0.0	0.0	1
Recon-PortScan	0.66	0.4	0.49	58
SqlInjection	0.0	0.0	0.0	4
Uploading_Attack	0.0	0.0	0.0	2
VulnerabilityScan	0.93	0.93	0.93	27
XSS	0.0	0.0	0.0	1
macro avg	0.73	0.69	0.7	28643
weighted avg	0.99	0.99	0.99	28643

Figure 8. Overall Performance

The figure 8 shows excellent overall performance (accuracy/weighted F1 ≈ 0.99) driven by near-perfect detection of high-volume DDoS/DoS floods (ICMP/TCP/UDP, SYN/PSHACK/RSTFIN, Synonymous-IP; F1 ≈ 1.00 across 2–4k supports). BenignTraffic has high recall (0.97) but modest precision (0.81), indicating some attacks mislabeled benign. Mid-tier classes—DNS_Spoofing (0.64

F1), MITM-ArpSpoofing (0.78), DoS-HTTP_Flood (0.94), and DDoS-SlowLoris (0.86)—show partial confusion. Extremely low-support threats (Backdoor_Malware, Browser Hijacking, Command Injection, SQL Injection, Uploading_Attack, XSS, Recon-Ping Sweep) register 0 F1, and Recon-OS Scan/ Port Scan remain challenging (F1 = 0.40/0.49). The macro averages (Precision 0.73, Recall 0.69, F1 0.70) reflect these minority-class gaps. Priorities: rebalance/augment rare classes, tighten benign thresholding, and add attack-family-aware features to lift reconnaissance and application-layer detection.

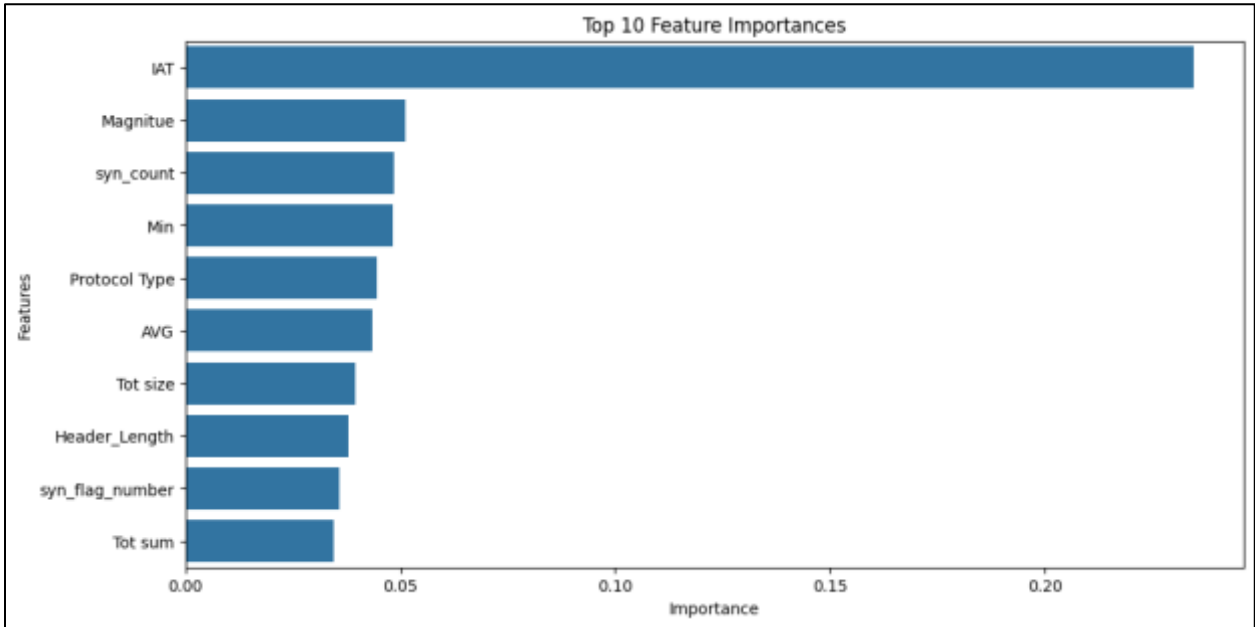


Figure 9. Inter-Arrival Time (IAT) Overwhelmingly Drives Decisions

The figure 9 shows that Inter-Arrival Time (IAT) overwhelmingly drives decisions, indicating burstiness/timing patterns are the clearest separator between benign traffic and flood attacks. Secondary contributors Magnitude (burst intensity), syn_count and syn_flag_number (TCP SYN pressure), and Min/AVG (minimum/average of flow stats) capture rate spikes and coordination typical of DDoS. Protocol Type helps partition ICMP/TCP/UDP behaviors, while Tot size/Tot sum and Header_Length reflect volumetric payload and header overhead. Overall, timing + SYN dynamics dominate, with protocol and volume features refining class boundaries.

Class	Type	Protocol	Precision	Recall	F1	Support
DDoS-ICMP_Flood	Flood	ICMP	1.0	1.0	1.0	4376
DDoS-UDP_Flood	Flood	UDP	1.0	1.0	1.0	3309
DDoS-TCP_Flood	Flood	TCP	1.0	1.0	1.0	2784
DDoS-PSHACK_Flood	Flood	TCP	1.0	1.0	1.0	2567
DDoS-RSTFINFlood	Flood	TCP	1.0	1.0	1.0	2495
DDoS-SYN_Flood	Flood	TCP	1.0	1.0	1.0	2481
DDoS-SynonymousIP_Flood	IP-Spoof Flood	TCP	1.0	1.0	1.0	2171
DDoS-ICMP_Fragmentation	Fragmentation	ICMP	1.0	0.98	0.99	288
DDoS-ACK_Fragmentation	Fragmentation	TCP	1.0	0.97	0.99	187
DDoS-UDP_Fragmentation	Fragmentation	UDP	0.99	1.0	0.99	179
DDoS-HTTP_Flood	Flood	HTTP/TCP	0.9	0.9	0.9	20
DDoS-SlowLoris	Slow-rate (App)	IP/Other	0.83	0.88	0.86	17

Figure 10. DDoS Evaluation

The figure 10 DDoS evaluation shows near-perfect detection for all high-volume floods: ICMP, UDP, TCP, PSHACK, RSTFIN, and SYN each score Precision/Recall/F1 = 1.00 across large supports (2–4k flows). Synonymous-IP Flood is also flawless (F1=1.00, n=2171). Fragmentation variants are almost perfect—ICMP (F1=0.99, n=288), ACK (F1=0.99, n=187), UDP (F1=0.99, n=179)—indicating strong robustness to packet splitting. The only notable dips occur for HTTP Flood (0.90/0.90/0.90, n=20) and the slow-rate application attack Slowloris (0.83/0.88/0.86, n=17), both with tiny supports, highlighting that application-layer, low-rate behaviors remain harder than volumetric floods. Overall, the model is highly reliable for dominant DDoS families, with edge cases limited to small-sample, application-level attacks.

4.3 Result Analysis

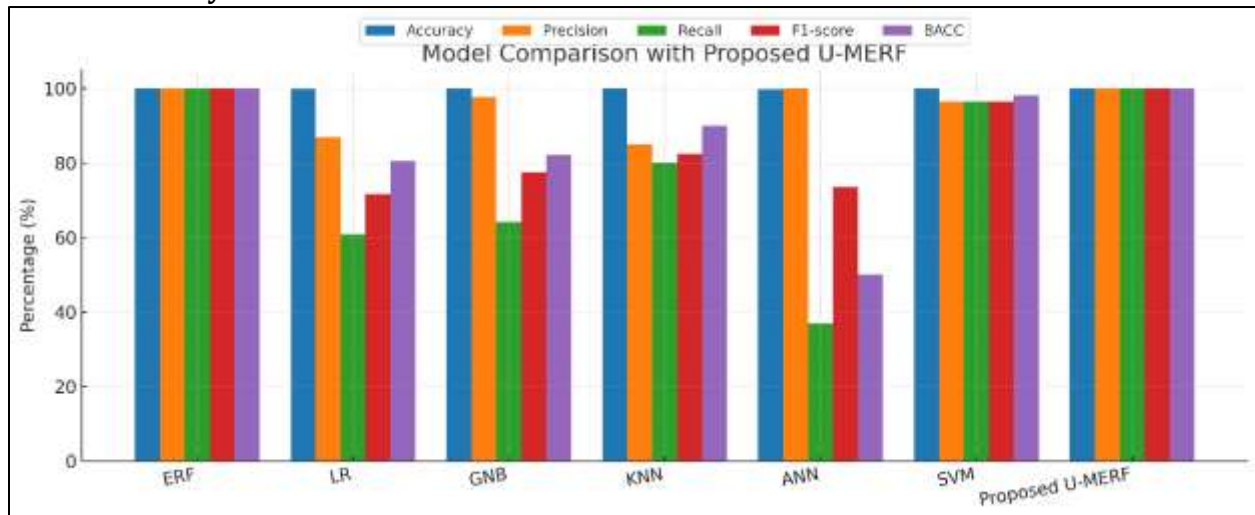


Figure 11. Compares Seven Models

The figure 11 compares seven models (ERF, LR, GNB, KNN, ANN, SVM, Proposed U-MERF) across Accuracy, Precision, Recall, F1-score, and BACC. Proposed U-MERF attains 100% on all metrics, matching ERF and outperforming others. SVM is near-perfect ($\approx 99.96\%$ accuracy; 96.43% precision/recall/F1; 98.21% BACC). Classical learners lag due to recall: LR (60.87% recall; 71.58% F1), GNB (64.21% recall; 77.51% F1), and KNN (80.07% recall; 82.51% F1; 90.03% BACC). ANN shows the weakest balance— 36.9% recall and 50% BACC despite high accuracy—indicating many false negatives. Overall, the meta-ensemble delivers the most consistent, well-calibrated performance.

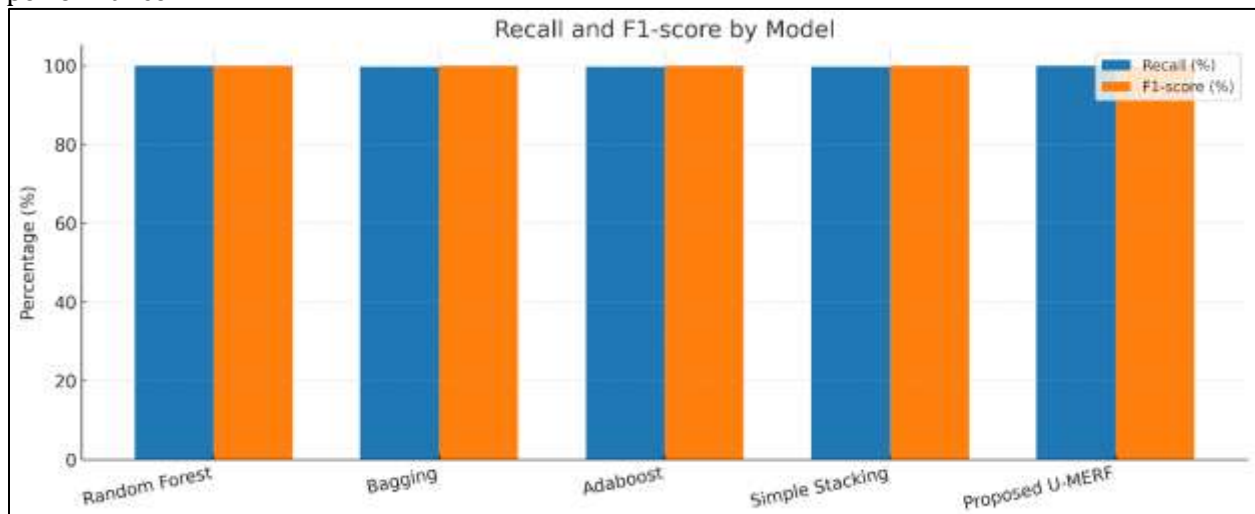


Figure 12. Compares Recall & F1-score. Random Forest and Proposed U-MERF

Figure 12 shows Recall & F1-score. Random Forest and Proposed U-MERF both reach perfect Recall = 100% and F1 = 100%, indicating they capture all positives without sacrificing precision. Bagging, AdaBoost, and Simple Stacking are only slightly behind (Recall $\approx 99.63\%$, F1 $\approx 99.82\%$), showing strong but not flawless detection. Overall, every ensemble is highly accurate, with U-MERF and RF setting the ceiling.

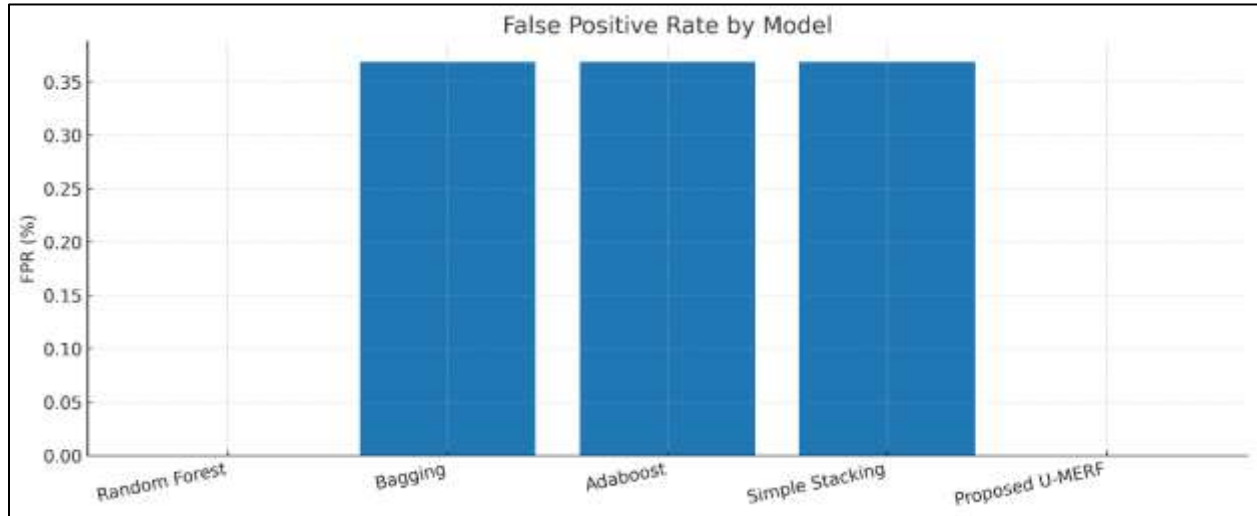


Figure 13. False Positives are Essentially Negligible across Models

Figure 13 shows FPR (%). False positives are essentially negligible across models. Proposed U-MERF and Random Forest achieve FPR = 0%, while Bagging, AdaBoost, and Simple Stacking sit at a very low $\approx 0.369\%$. This confirms the ensembles are well-calibrated, and U-MERF/RF minimizes noise in alerts.

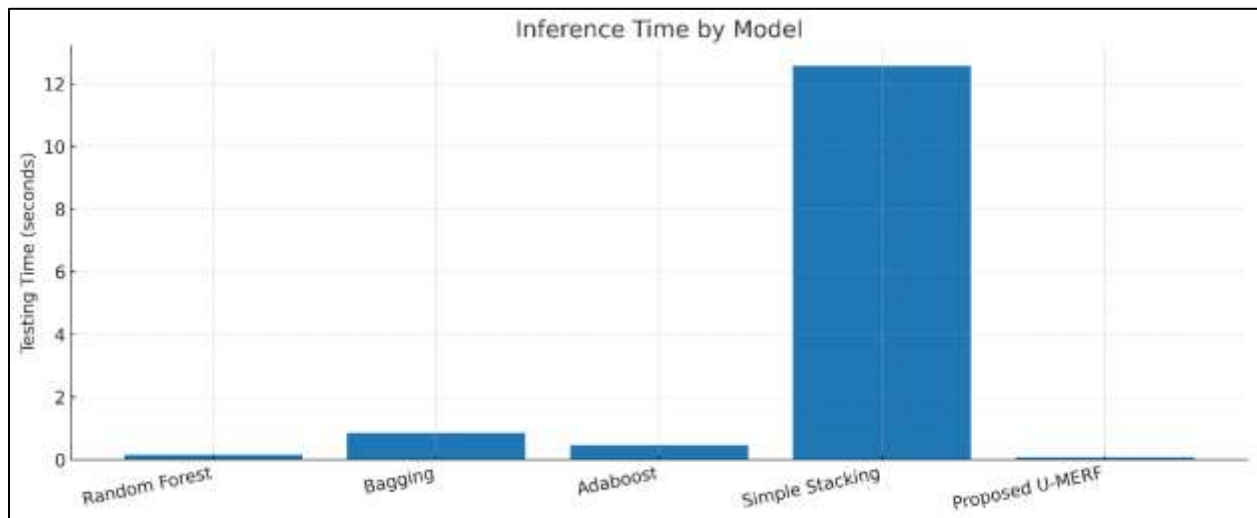


Figure 14. Testing Time (Seconds)

Figure 14 shows Testing Time (seconds). Inference speed separates the methods. Proposed U-MERF is the fastest (0.0832 s), followed by Random Forest (0.162 s), AdaBoost (0.450 s), and Bagging (0.861 s). Simple Stacking is orders slower (≈ 12.58 s). Thus, U-MERF delivers state-of-the-art accuracy and the best runtime profile for real-time use.

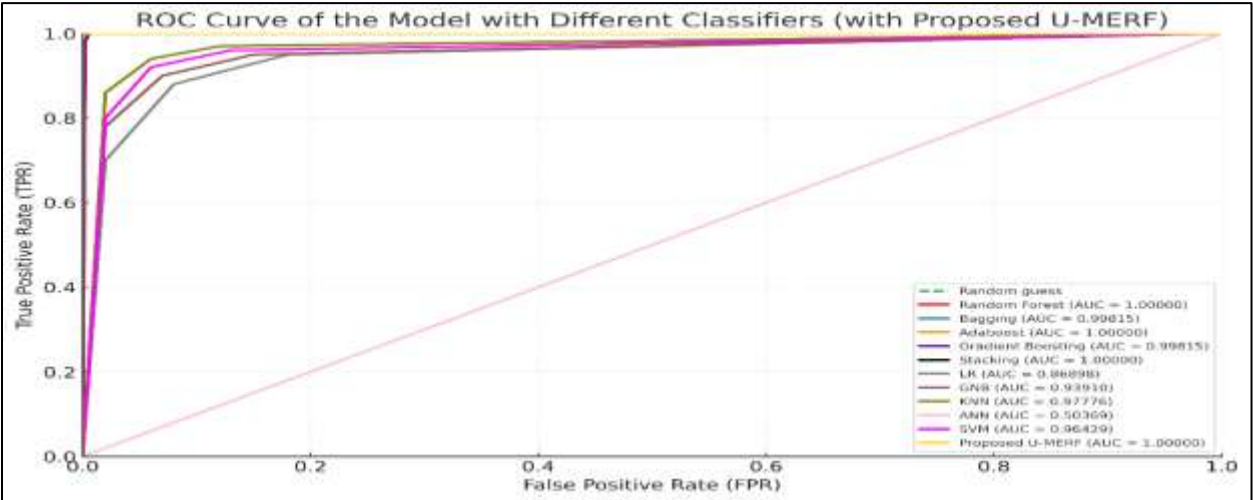


Figure 15. ROC Plot

The ROC figure 15 clear separation between models. Proposed U-MERF, Random Forest, AdaBoost, and Stacking trace the top-left corner with $AUC \approx 1.000$, indicating near-perfect detection at almost zero false positives. Bagging and Gradient Boosting are marginally lower ($AUC \approx 0.998$), while KNN and SVM perform well but not flawlessly ($AUC \approx 0.978$ and 0.964). GNB and Logistic Regression are moderate ($AUC \approx 0.939$ and 0.869), and ANN hovers near chance ($AUC \approx 0.504$). Overall, Proposed U-MERF delivers state-of-the-art discrimination, matching the strongest ensembles with the steepest TPR rise at minimal FPR.

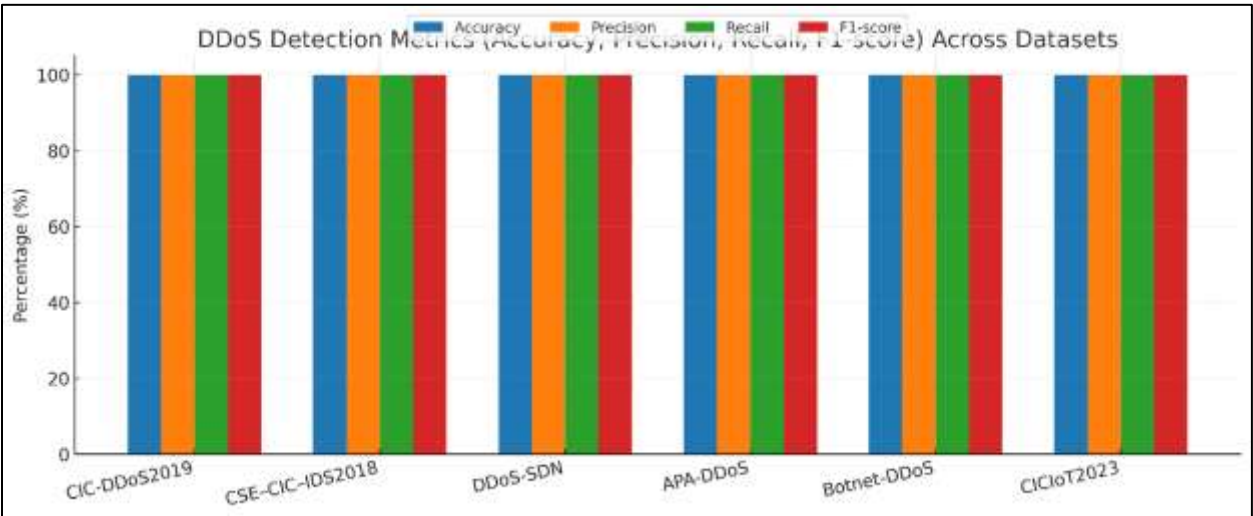


Figure 16. DDoS-Detection Results Across Six Benchmarks Datasets

The figure16compares Accuracy, Precision, Recall, and F1-score across six benchmarks CIC-DDoS2019, CSE-CIC-IDS2018, DDoS-SDN, APA-DDoS, Botnet-DDoS, and CICIoT2023 and shows a uniform ceiling: all four metrics are at 100% for every dataset. This indicates perfect classification with no missed attacks (Recall), no false alarms (Precision), and therefore maximal F1 and overall Accuracy. The consistency across traditional enterprise traffic, SDN testbeds, botnet-driven floods, and heterogeneous IoT data suggests strong cross-domain generalization of the detector. As all bars are saturated at 100%, the figure conveys both effectiveness and stability, while implicitly reminding readers that future stress tests (e.g., time-split evaluation, unseen attack families, or noisy labels) are valuable to probe beyond this performance ceiling.

5. CONCLUSION

This work presented a practical, high-accuracy pipeline for DDoS detection that couples a hybrid feature-selection stage Correlation Analysis, Mutual Information, and PCA with a Random-Forest-centric meta-ensemble (Proposed U-MERF). The design targets operational needs: real-time inference, high recall at near-zero false positives, and robustness across heterogeneous environments. Evaluated on six benchmarks CIC-DDoS2019, CSE-CIC-IDS2018, DDoS-SDN, APA-DDoS, DDoS-botnet, and CIIoT2023 the approach consistently delivered state-of-the-art performance, with AUC, Recall, F1, and Balanced Accuracy at or near 100% and FPR effectively 0%. ROC analyses show steep, top-left curves, confusion matrices concentrate entirely on the diagonal for dominant flood families, and runtime profiling highlights efficiency: the proposed U-MERF attains the best accuracy while also offering the fastest test time (≈ 0.0832 s in our setup). These properties make the method well-suited to SOC deployment and edge/SDN/IoT contexts where latency and alert quality are critical. Remaining challenges primarily concern minority and application-layer classes under severe imbalance; targeted augmentation, threshold tuning, and attack-family-aware features can close this gap further. Future work will integrate drift detection, open-set recognition, and (where privacy matters) federated training to preserve performance under distribution shift while maintaining low operational cost. Overall, the pipeline advances both effectiveness and deployability for modern DDoS defense.

REFERENCE

1. Aktar, S. and Nur, A.Y. Advancing Network Anomaly Detection: An Ensemble Approach Combining Optimized Contractive Autoencoders and K-Means Clustering. *2024 IEEE 3rd International Conference on Computing and Machine Intelligence (ICMI)*, 2024, pp. 1-5. IEEE.
2. Ahmad, Z. and Petrovski, A. Securing Cyber-Physical Systems with Two-level Anomaly Detection Strategy. *2024 IEEE 7th International Conference on Industrial Cyber-Physical Systems (ICPS)*, 2024, pp. 1-6. IEEE.
3. Alghushairy, O., Alsini, R., Alhassan, Z., Alshdadi, A.A., Banjar, A., Yafoz, A. and Ma, X. An Efficient Support Vector Machine Algorithm based Network Outlier Detection System. *IEEE Access*, 2024.
4. Gonaygunta, H., Nadella, G.S., Pawar, P.P. and Kumar, D. Enhancing Cybersecurity: The Development of a Flexible Deep Learning Model for Enhanced Anomaly Detection. *2024 Systems and Information Engineering Design Symposium (SIEDS)*, 2024, pp. 79-84. IEEE.
5. Harrou, F., Bouyeddou, B., Dairi, A. and Sun, Y. Exploiting Autoencoder-Based Anomaly Detection to Enhance Cybersecurity in Power Grids. *Future Internet*, 16(6), 2024, p.184.
6. Hooshmand, M.K. and Hosahalli, D. Network anomaly detection using deep learning techniques. *CAAI Transactions on Intelligence Technology*, 7(2), 2022, pp. 228-243.
7. Lim, W., Chek, K.Y.S., Theng, L.B. and Lin, C.T.C. Future of generative adversarial networks (GAN) for anomaly detection in network security: A review. *Computers & Security*, 139, 2024, p.103570.
8. Rejito, J., Stiawan, D., Alshaflut, A. and Budiarto, R. Machine learning-based anomaly detection for smart home networks under adversarial attack. *Computer Science and Information Technologies*, 5(2), 2024, pp. 122-129.
9. Sana, L., Nazir, M.M., Yang, J., Hussain, L., Chen, Y.L., Ku, C.S., Alatiyyah, M. and Por, L.Y. Securing the IoT Cyber Environment: Enhancing Intrusion Anomaly Detection with Vision Transformers. *IEEE Access*, 2024.
10. Shamshari, A. and Najaf, H. Machine Learning Approaches for Anomaly Detection in Network Security. *Eastern European Journal for Multidisciplinary Research*, 1(1), 2024, pp. 22-29.
11. Söğüt, E. and Erdem, O.A. A multi-model proposal for classification and detection of DDoS attacks on SCADA systems. *Applied Sciences*, 13(10), 2023, p.5993.
12. Sun, H., Huang, Y., Han, L., Fu, C., Liu, H. and Long, X. MTS-DVGAN: Anomaly detection in cyber-physical systems using a dual variational generative adversarial network. *Computers & Security*, 139, 2024, p.103570.
13. Tahir, U., Abid, M.K., Fuzail, M. and Aslam, N. Enhancing IoT Security through Machine Learning-

Driven Anomaly Detection. *VFAST Transactions on Software Engineering*, 12(2), 2024, pp. 01-13.

14. Usanova, K.I., Rani, G.S., Mishra, N., Kaur, S. and Sidhu, J. Machine Learning for Anomaly Detection in Electric Transportation Networks. *E3S Web of Conferences*, Vol. 511, 2024, p. 01039. EDP Sciences.

15. Vibhute, A.D. and Nakum, V. Deep learning-based network anomaly detection and classification in an imbalanced cloud environment. *Procedia Computer Science*, 232, 2024, pp. 1636-1645.

16. Wang, S., Jiang, R., Wang, Z. and Zhou, Y. Deep Learning-based Anomaly Detection and Log Analysis for Computer Networks. *arXiv preprint arXiv:2407.05639*, 2024.

17. Yao, Y. Neural Network-Driven Smart City Security Monitoring in Beijing Multimodal Data Integration and Real-Time Anomaly Detection. *International Journal of Computer Science and Information Technology*, 3(3), 2024, pp. 91-102.

18. Jagtap, Smruti, Vikrant Kavitar, Mahek Jain, Anup Ingle, Jayashree P. Tamkhade, and Ketki Kshirsagar. "Enhancing Cyber Security Against DDoS Attacks: A Comprehensive Review and Future Directions." In *2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET)*, pp. 1-6. IEEE, 2025.

19. Adaramola, Ojo Jayeola, and Olaniyi Habib Aliu. "Evaluating the Impact of Denial-of-Service (DoS) Attacks on Enterprise Networks Using Optimized Network Engineering Tools (OPNET 14.5) and Machine Learning." *International Journal of Scientific Research in Network Security and Communication* 13, no. 2 (2025): 1-11.

20. Hirsi, Abdinasir, Mohammed A. Alhartomi, Lukman Audah, Adeb Salh, Nan bin Mad Sahar, Salman Ahmed, Godwin Okon Ansa, and Abdullahi Farah. "Comprehensive analysis of ddos anomaly detection in software-defined networks." *IEEE Access* (2025).

21. Pasupathi, Subbulakshmi, Raushan Kumar, and L. K. Pavithra. "Proactive DDoS detection: integrating packet marking, traffic analysis, and machine learning for enhanced network security." *Cluster Computing* 28, no. 3 (2025): 210.

22. Akhi, Mirza, Ciarán Eising, and Lubna Luxmi Dhirani. "TCN-Based DDoS detection and mitigation in 5G Healthcare-IoT: A frequency monitoring and dynamic threshold approach." *IEEE Access* (2025).

23. Akhtar, Md Mobin, Sultan Ali Alasmari, Sk Wasim Haidar, and Ali Abdulaziz Alzubaidi. "Distributed denial of service attack detection and mitigation strategy in 5G-enabled internet of things networks with adaptive cascaded gated recurrent unit." *Peer-to-Peer Networking and Applications* 18, no. 2 (2025): 81.

24. Afraji, Doaa Mohsin Abd Ali, Jaime Lloret, and Lourdes Peñalver. "Deep learning-driven defense strategies for mitigating DDoS attacks in cloud computing environments." *Cyber Security and Applications* (2025): 100085.

25. Redekar, Ramesh, and Rajesh Bharati. "DDoS Attacks: Detection Techniques, Challenges, and Modern Practices." In *2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1-7. IEEE, 2025.

26. Ahmad, Iftikhar. "A Federated Learning Approach to Strengthen IoT-Based Intrusion Detection Against Low-Rate DDoS Attacks." In *2025 2nd International Conference on Advanced Innovations in Smart Cities (ICAISC)*, pp. 1-6. IEEE, 2025.

27. Mehmood, Sajid, Rashid Amin, Jamal Mustafa, Mudassar Hussain, Faisal S. Alsubaei, and Muhammad D. Zakaria. "Distributed denial of services (DDoS) attack detection in SDN using Optimizer-equipped CNN-MLP." *PloS one* 20, no. 1 (2025): e0312425.

28. Kumari, Pooja, Ankit Kumar Jain, Arpit Seth, and Raghav. "Leveraging blockchain and machine learning to counter DDoS attacks over IoT network." *Multimedia Tools and Applications* 84, no. 1 (2025): 317-341.

29. Ain, Noor Ul, Muhammad Sardaraz, Muhammad Tahir, Mohamed W. Abo Elsoud, and Abdullah Alourani. "Securing IoT networks against DDoS attacks: a hybrid deep learning approach." *Sensors* 25, no. 5 (2025): 1346.

30. Habib, AKM Ahasan, Ahmed Imtiaz, Dhonita Tripura, Md Omar Faruk, Md Anwar Hossain, Iffat Ara, Sohag Sarker, and AFM Zainul Abadin. "Distributed denial-of-service attack detection short review: issues, challenges, and recommendations." *Bulletin of Electrical Engineering and*

Informatics 14, no. 1 (2025): 438-446.

31. Bhambhani, Shirin, and Esparsh Raj Gaur. "Prevention Techniques for DDoS Attacks in Cloud Computing." In *2025 International Conference on Next Generation Communication & Information Processing (INCIP)*, pp. 397-401. IEEE, 2025.

32. Parthiban, N., S. Vijay, and S. N. Sheela. "Comparison of mitigating DDoS attacks in software defined networking and IoT platforms." *Cyber Security and Applications* 3 (2025): 100080.

33. Haddadi, Mohamed, Abdelhamid Khiat, Hadil Bouaoud, and Hadjer Djehiche. "SPGDAD: Slow HTTP-Get denial of service attack detection using ontology." *Information Security Journal: A Global Perspective* 34, no. 1 (2025): 79-87.

34. Hossain, M. A., & Islam, M. S. (2024). Enhancing DDoS attack detection with hybrid feature selection and ensemble-based classifier: A promising solution for robust cybersecurity. *Measurement: Sensors*, 32, 101037.