Submitted to IEEE Access for review.

Digital Object Identifier xx.xxxx/ACCESS.xxxx.DOI

WiMapper: A Lightweight Kernel-Based Framework for Rogue Access Point Detection on Edge Devices

A HARSHA KUMAR[1], S KRITHICK[1], KUMARAN K[1], SARANYA G[1], ABISHEK DEVADOSS[1], R SACHEEV KRISHANU[1], R SACHEEV KRISHANU[1], DISHA DANIEL[2], PREETHIKA RANGANATHAN[1]

[1]School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India (e-mail: ahkharsha@gmail.com; krithick2004@gmail.com; kumaran.k@vit.ac.in; saranya.g@vit.ac.in; dabishek2004@gmail.com; sacheevkrish@gmail.com; preethika.r.2004@gmail.com)
[2]School of Electronics Engineering, Vellore Institute of Technology, Chennai, India (e-mail: dishadaniel24@gmail.com)

Corresponding author: Kumaran K (e-mail: kumaran.k@vit.ac.in).

ABSTRACT Public wireless networks have evolved from convenient amenities into critical urban infrastructure, yet they remain fundamentally susceptible to identity-based exploits. The "Evil Twin" attack persists as a significant threat because client devices implicitly trust broadcast identifiers before a secure encryption channel is established. Existing defense mechanisms typically rely on either prohibitively expensive Wireless Intrusion Prevention Systems (WIPS), which necessitate centralized wired infrastructure, or computationally intensive Deep Learning models that exceed the resource capabilities of edge sensors. This paper presents WiMapper, a detection framework engineered specifically for Resource-Constrained Edge Devices (RCEDs) with less than 320 KB of SRAM. The architecture employs a two-stage hybrid approach: Track A utilizes a deterministic whitelist for immediate threat filtering, while Track B deploys a One-Class Support Vector Machine (OC-SVM) with a Radial Basis Function (RBF) kernel to assess signal integrity. By analyzing higher-order statistical features—specifically Kurtosis and Skewness—the model identifies non-Gaussian anomalies characteristic of signal spoofing. Extensive simulations using the HCXY dataset and subsequent field validation demonstrated that WiMapper achieves a Pareto-optimal balance between efficiency and accuracy. The framework attained a mean F1-Score of 0.827 with an algorithmic inference latency of 0.25 ms and a memory footprint of 190 KB. These metrics confirm that the kernel-based approach significantly outperforms Isolation Forest and Autoencoder baselines, rendering it a viable solution for dense, low-power security sensor networks.

INDEX TERMS Edge Computing, Network Security, One-Class SVM, Rogue Access Points, Wireless Sensor Networks, RSSI Analysis, IoT Security.

I. INTRODUCTION

IEEE 802.11 (Wi-Fi) networks have transformed from local connectivity solutions into essential utilities that underpin daily operations in airports, stadiums, and metropolitan centers. As connectivity becomes ubiquitous, the inherent security model of the protocol reveals critical vulnerabilities regarding identity verification.

In a standard connection handshake, a client device—such as a smartphone or an industrial IoT sensor—trusts an Access Point (AP) based primarily on its advertised Service Set Identifier (SSID) and Media Access Control (MAC) address. Malicious actors exploit this implicit trust through "Evil Twin" attacks. In this scenario, an adversary configures a radio device to broadcast the precise credentials of a legitimate trusted network. To the client device, the counterfeit AP appears identical to the authorized infrastructure. Once a user connects, the attacker establishes a Man-in-the-Middle (MitM) position, enabling the interception of unencrypted data, the injection of malicious payloads, or the harvesting of credentials before the user detects a breach.

As reliance on public networks intensifies, attack methodologies have become increasingly sophisticated. Automated tools can now dynamically adjust signal strength to overpower legitimate APs, forcing client devices to associate

with the malicious node. Consequently, effective defense mechanisms must be as dynamic and adaptive as the threats they aim to mitigate.

A. THE PRE-ASSOCIATION VULNERABILITY

A significant limitation of host-based security software, such as VPNs or OS-level firewalls, is that they operate primarily at the Network Layer (Layer 3) or above. These tools inspect data packets *after* a connection is established and an IP address is assigned. This creates a timing failure defined in this study as the **Pre-Association Vulnerability**.

To analyze the threat using standard host-based tools, the client device must first associate with the network. However, the act of associating exposes the device to the attack. The moment the Layer 2 (Data Link) handshake concludes, the attacker controls the communication channel. Therefore, rigorous security monitoring must occur *Out-of-Band*. This necessitates an independent sensor capable of monitoring the Radio Frequency (RF) environment and identifying threats solely from unencrypted management frames, such as Beacons, before any legitimate client attempts to connect. This preventive approach mitigates the risk by alerting users prior to device exposure.

B. OPERATIONAL CONSTRAINTS AND SCALABILITY

While Out-of-Band monitoring addresses the timing issue, the implementation strategy presents challenges regarding hardware scalability. Enterprise environments often employ Wireless Intrusion Prevention Systems (WIPS), which are effective but rely on expensive sensors connected via wired backhaul to central servers. This centralized model is unsuitable for ad-hoc or temporary deployments where wired infrastructure is unavailable.

Conversely, utilizing single-board computers running full operating systems as portable sensors is often impractical for mass deployment due to specific limitations:

1) **Scalability and Cost:** Securing large public venues requires a dense sensor array to overcome signal shadowing and limited range. Deploying high-cost single-board computers is economically inefficient. In contrast, commodity microcontrollers (MCUs) are significantly more cost-effective. This cost differential allows for the deployment of a higher volume of sensors within the same budget, ensuring comprehensive coverage without blind spots.
2) **Power Efficiency:** Devices running full operating systems incur substantial power overhead due to background processes and require significant boot time. A "bare-metal" microcontroller provides instant-on capability and deep sleep modes, allowing for extended operation on battery power. This efficiency facilitates "drop-and-forget" deployment strategies in locations where frequent maintenance is not feasible.

This study establishes a strict **Design Constraint**: the detection algorithm must execute entirely within the SRAM limits of a standard Resource-Constrained Edge Device (RCED), defined here as having less than 320 KB of memory. Furthermore, it must operate in real-time without reliance on cloud offloading. This requirement excludes state-of-the-art Deep Learning models, such as Transformers or complex Autoencoders, which typically require runtime libraries and model weights that exceed the available memory on these devices.

C. CONTRIBUTIONS

To resolve the conflict between diagnostic accuracy and extreme resource limitations, this paper introduces **WiMapper**. This hybrid framework is predicated on the insight that kernel-based statistical learning can offer non-linear separation capabilities comparable to neural networks, but with significantly lower computational overhead.

The primary contributions of this work include:

- **A Hybrid Detection Architecture:** The framework utilizes a multi-stage system to balance efficiency and rigor. Track A employs a deterministic whitelist for immediate rejection of naive threats, while Track B deploys a One-Class Support Vector Machine (OC-SVM) to analyze signal integrity. This tiered structure minimizes the computational load by reserving intensive processing only for ambiguous signals.
- **Higher-Order Statistical Characterization:** The study validates the utility of Kurtosis and Skewness as robust digital fingerprints. The methodology demonstrates that these features capture specific impulsive noise artifacts introduced by packet injection tools, enabling the model to distinguish artificial spoofing from natural environmental fading.
- **Pareto-Optimal Efficiency:** Rigorous benchmarking confirms the system's viability for edge deployment. WiMapper achieved a mean F1-Score of 0.827 with a memory footprint of 190 KB. These results position the solution on the Pareto Frontier, offering superior accuracy compared to lightweight Isolation Forests while remaining efficient enough for the target hardware class.

The remainder of this paper is organized as follows: Section II reviews existing literature. Section III establishes the system model. Section IV details the WiMapper architecture. Section V describes the experimental setup. Section VI presents the performance analysis. Section VII discusses field validation. Finally, Section VIII concludes the paper.

II. RELATED WORK

The challenge of securing the wireless edge has driven extensive research across three distinct paradigms: infrastructure-centric monitoring, host-based defenses, and algorithmic anomaly detection [1]–[3]. This section critically reviews the state-of-the-art to identify the specific architectural and computational gaps addressed by the WiMapper framework.

A. INFRASTRUCTURE-CENTRIC MONITORING (WIPS)

The current gold standard for enterprise-grade wireless security is the Wireless Intrusion Prevention System (WIPS) [1], [4]. Market leaders deploy dedicated sensor nodes that continuously scan the RF spectrum, independent of data traffic [1], [5]. These systems operate by comparing discovered Basic Service Set Identifiers (BSSIDs) against a centralized, manually curated whitelist of authorized hardware.

Advanced WIPS implementations go beyond simple identifier matching. They utilize "RF Fingerprinting" or Radiometric Identification [6], [7]. This technique analyzes unique physical imperfections in the radio transmitter—such as clock skew, transient turn-on signatures, or modulation errors—to identify cloned devices [6]. Because these hardware imperfections are artifacts of the manufacturing process, they are difficult for an attacker to replicate purely through software manipulation [6], [8].

While highly effective in controlled, static environments, WIPS solutions are architecturally unsuited for the decentralized and ad-hoc use cases targeted by this study [1].

- **Immobility and Infrastructure Dependency:** WIPS relies on a wired backhaul to a central controller for signature analysis [1]. This creates a "dome of protection" that exists only where the infrastructure is physically installed. A user moving through a public terminal, staying in a hotel, or renting a temporary workspace leaves this protected zone, entering a vulnerability blind spot where no external monitor exists [1], [2].
- **Economic Exclusion:** The high capital expenditure of commercial WIPS nodes (often exceeding hundreds of dollars per unit) limits their deployment to large enterprises and government facilities [1]. Small businesses, public squares, and developing regions remain vulnerable, creating a "security divide" based on economic resources [2].

B. HOST-CENTRIC AND CLIENT-SIDE DEFENSES

Recognizing the limitations of fixed infrastructure, researchers have attempted to shift detection logic to the client device [9]–[11]. Solutions in this category typically involve software agents running on the user's laptop or smartphone. These agents monitor network layer metrics, such as Round Trip Time (RTT), Domain Name System (DNS) resolution patterns, or the availability of duplicate SSIDs, to infer the presence of an attacker [9], [10].

However, host-centric defenses face insurmountable hardware abstraction barriers imposed by modern Operating Systems (OS). Network Interface Cards (NICs) and their associated drivers are designed for connectivity, not surveillance [10].

- **Abstraction Blindness:** OS architectures (Windows, macOS, Android) typically abstract away raw management frames [10], [11]. A standard application cannot easily access the raw Received Signal Strength Indicator (RSSI) of a beacon frame or inspect the Information Elements (IEs) inside a Probe Response. Without access to these raw Layer 2 headers, client-side software lacks the physical visibility required to detect a well-configured Evil Twin [9], [10].
- **The Pre-Association Gap:** As noted in Section I, these tools often suffer from a critical timing failure. They activate only after the network interface has associated and an IP address has been assigned [10], [11]. By this moment, the device has already exposed its MAC address and potentially exchanged handshake credentials with the rogue node [9].

C. MACHINE LEARNING IN WIRELESS SECURITY

The failure of static, rule-based systems to detect adaptive attackers—who can spoof MAC addresses and mimic SSID strings—has led to the widespread adoption of Machine Learning (ML) for anomaly detection [1]–[3], [12].

1) Deep Learning Models

Recent literature heavily favors Deep Learning (DL) architectures [1], [3]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks have shown significant success in modeling the temporal sequence of RSSI values, predicting the expected signal path and flagging deviations [13]. Similarly, Autoencoders (AEs) have been employed for unsupervised anomaly detection, trained to reconstruct "normal" traffic patterns and identifying threats based on high reconstruction errors [14]–[17].

While diagnostically powerful, these models violate the constraints of the target hardware class (RCEDs).

- **Computational Bloat:** The inference phase of a standard Autoencoder involves dense matrix multiplications ($O(N^2)$ complexity). For a fully connected layer, this requires thousands of floating-point operations per sample. On a microcontroller lacking a dedicated Floating Point Unit (FPU) or Tensor Accelerator, these operations must be emulated in software, drastically extending the CPU duty cycle and draining the battery [1], [2], [18].
- **Memory Constraints:** The storage requirement for model weights and the necessary runtime interpreter (e.g., TensorFlow Lite Micro) often exceeds the megabyte range. This makes them undeployable on microcontrollers with kilobyte-scale SRAM, forcing reliance on cloud offloading, which reintroduces latency and privacy risks [2], [3].

2) Lightweight Statistical Models

To address resource constraints, researchers have explored lighter alternatives such as Gaussian Mixture Models (GMMs) and Isolation Forests (IF) [18]–[20]. These algorithms are computationally efficient, often requiring only linear time complexity ($O(N)$), and fit easily within constrained memory [12], [20].

However, these models typically rely on linear or axis-parallel decision boundaries. This study indicates that so-

phisticated "Signal Cloning" attacks create *non-linear* distortions in the feature space [6], [7]. An adaptive attacker may successfully mimic the *mean* signal strength (fooling a linear classifier) but fail to replicate the distribution's shape [8]. Isolation Forests, which partition data using random orthogonal cuts, often lack the sensitivity to detect these subtle distributional shifts (e.g., changes in Kurtosis), leading to high False Negative rates against competent attackers [20].

D. THE GAP ANALYSIS

Table 1 summarizes the current landscape. A clear **Security-Resource Gap** exists: high-accuracy models (Deep Learning) are too heavy for scalable deployment, while lightweight models (Isolation Forests) lack the diagnostic precision to ensure safety against adaptive threats [1], [3], [12].

TABLE 1: Comparative Analysis of Wireless Security Paradigms

Methodology	Hardware Class	Latency	Primary Limitation
Enterprise WIPS	Server + Wired Nodes	Low	High Cost, Immobility [1], [5]
Host-Based Agents	Laptop / Phone CPU	Medium	OS Restrictions, Late Detection [9]–[11]
Deep Learning (AE/LSTM)	GPU / High-End CPU	High	Resource Bloat (>1MB RAM) [14]–[17]
Isolation Forest	Commodity MCU	Low	Linear Boundaries (Low Accuracy) [20]
WiMapper (Proposed)	**Commodity MCU**	**Low**	**None (Pareto Optimal)**

WiMapper bridges this specific gap. By utilizing a kernel-based One-Class SVM, the framework achieves the non-linear separation capabilities usually reserved for Deep Learning, while maintaining a memory footprint small enough for the most constrained microcontrollers [16], [20]. This approach democratizes high-fidelity wireless security, enabling the deployment of "sensor swarms" without the cost or power penalties of traditional approaches [1], [2].

III. SYSTEM MODEL AND THEORETICAL FRAMEWORK

To mathematically formalize the deficiencies of the linear models identified in Section II, the physical nature of the threat landscape is first characterized. Given the strict memory limits established in Section I, utilizing raw waveform analysis or complex time-series forecasting models like LSTMs is not feasible. Instead, the system models the signal environment using statistical moments. This approach compresses high-frequency RF data into a compact feature vector while retaining the signatures of attack artifacts. This section details the threat model, provides the theoretical basis for signal propagation, and derives the statistical features used to identify anomalies.

A. THE ATTACKER MODEL

The public Wi-Fi threat landscape, visually represented as **Zone 1** in Fig. 1, includes two distinct classes of attackers, categorized by their sophistication and the resources required to detect them.

1) Type 1: Naive Impersonation

This attacker establishes a Rogue Access Point (RAP) broadcasting a legitimate Service Set Identifier (SSID) (e.g., "Free_Airport_WiFi") but creates the network using a random or default Media Access Control (MAC) address. This represents a common attack form, often launched using smartphone hotspots or unconfigured laptops. Because these attacks fail to clone the hardware address, they are deterministic in nature. Detection is achieved by comparing the broadcast BSSID against a known whitelist of authorized hardware addresses (Track A).

2) Type 2: Adaptive Signal Cloning

The Type 2 attacker represents a higher threat tier. Hereafter referred to as an "Evil Twin" in the visual analysis, this adversary utilizes specialized hardware—such as high-gain directional antennas and packet injection suites—to spoof both the SSID and the MAC address of the target AP. Furthermore, the attacker physically positions the device to match the Receive Signal Strength Indicator (RSSI) of the legitimate AP, effectively blending into the environment's baseline noise floor.

Because the identifiers and the mean signal strength mimic the legitimate infrastructure, linear filters cannot detect this threat. Detection requires analyzing the *statistical quality* and *distributional shape* of the signal. This is the primary objective of the Track B analysis.

B. SIGNAL PROPAGATION AND ANOMALY THEORY

The propagation of wireless signals in a complex indoor environment is modeled using the Log-Distance Path Loss equation. The received power $P_{rx}(d)$ at a distance d from the transmitter is given by:

$$P_{rx}(d)_{\text{dBm}} = P_0(d_0) - 10n \log_{10}\left(\frac{d}{d_0}\right) - X_\sigma \qquad (1)$$

where P_0 is the reference power at distance d_0, n is the path loss exponent (typically ranging from 2.0 to 4.0 for indoor environments), and X_σ represents **Shadowing**.

For a legitimate Access Point operating in a stable environment, the shadowing term X_σ follows a zero-mean Gaussian (Normal) distribution $X_\sigma \sim \mathcal{N}(0, \sigma^2)$. This distribution results from the Central Limit Theorem applied to the sum of many random multipath reflections caused by walls, furniture, and movement. Consequently, the probability density function (PDF) of a legitimate signal is symmetric and bell-shaped.

The fundamental premise of this framework is that a Type 2 Adaptive Signal Clone introduces specific non-Gaussian distortions into X_σ:

- **Impulsive Noise Artifacts:** Attack tools utilizing packet injection often transmit in high-intensity bursts to force client de-authentication or to overwhelm the

channel during the handshake window. This introduces "impulsive" noise that creates sudden, transient spikes in the RSSI stream, which differ from the smooth variations of natural environmental fading.

- **Leptokurtic Distributions:** The presence of these artificial bursts causes the signal distribution to become *Leptokurtic* (heavy-tailed). While the attacker may calibrate the device to match the mean RSSI (μ) of the target, the "shape" of the probability density function changes, exhibiting thicker tails due to the injection artifacts.

C. MATHEMATICAL FORMALIZATION OF FEATURES

To capture these physical anomalies, the system extracts a feature vector $\mathbf{x}$ from a sliding temporal window $W = \{r_1, r_2, ..., r_w\}$ of raw RSSI samples. Table 2 defines the mathematical notation used throughout this derivation.

TABLE 2: Mathematical Notations and Definitions. These symbols define the statistical moments used to construct the feature vector $\mathbf{x}$, mapping raw signal data to the kernel space.

Symbol	Definition
W	Sliding temporal window of size w
r_i	Individual RSSI sample at time i
μ	Mean RSSI (First Moment)
σ	Standard Deviation (Second Moment)
S_k	Skewness (Third Standardized Moment)
K_u	Excess Kurtosis (Fourth Standardized Moment)
R_{roc}	RSSI Rate of Change
$\Phi(\mathbf{x})$	Kernel mapping function to Hilbert space

1) Skewness (S_k)

Skewness quantifies the asymmetry of the signal distribution. A legitimate AP typically exhibits a symmetric distribution ($S_k \approx 0$). However, attackers often use directional antennas (Yagi or Panel) to boost range. If the antenna is not perfectly aligned, or if the attacker is in motion, the signal strength biases in one direction, creating a non-symmetric tail.

$$S_k = \frac{\frac{1}{w}\sum_{i=1}^{w}(r_i - \mu)^3}{\left(\frac{1}{w}\sum_{i=1}^{w}(r_i - \mu)^2\right)^{3/2}} \tag{2}$$

2) Kurtosis (K_u)

Kurtosis serves as the primary discriminator for detecting packet injection. It measures the "tailedness" of the distribution.

$$K_u = \frac{\frac{1}{w}\sum_{i=1}^{w}(r_i - \mu)^4}{\left(\frac{1}{w}\sum_{i=1}^{w}(r_i - \mu)^2\right)^2} - 3 \tag{3}$$

The value 3 is subtracted to calculate *Excess Kurtosis*, normalizing the result against a standard Normal distribution. A value deviating significantly from 0 indicates the presence of outliers consistent with artificial manipulation rather than natural fading.

3) RSSI Rate of Change (R_{roc})

To capture the rapid signal injections characteristic of packet flooding, the system calculates the magnitude of the first-order difference of the RSSI stream:

$$|R_{roc}| = |r_t - r_{t-1}| \tag{4}$$

High values of $|R_{roc}|$ indicate artificial signal jumps or non-physical displacement effects rarely seen in natural pedestrian fading, where path loss changes gradually over time.

Table 3 summarizes the distinct roles of these features within the detection logic.

TABLE 3: Feature Sensitivity Analysis. Each feature targets a specific anomaly type; Mean RSSI establishes the physical context, while Kurtosis and Skewness identify the non-Gaussian artifacts characteristic of signal injection tools.

Feature	Physical Meaning	Detection Function
Mean RSSI (μ)	Distance / Path Loss	**Contextual Anchor:** Defines the valid "Normal" zone. Outliers indicate impossible proximity.
Std. Dev. (σ)	Signal Volatility	**Stability Check:** High variance indicates active jamming or unstable hardware.
Kurtosis (K_u)	Packet Burstiness	**Primary Discriminator:** Detects impulsive noise. High Kurtosis correlates with packet injection.
Skewness (S_k)	Distribution Symmetry	**Hardware Artifact:** Detects asymmetry caused by directional amplification.
Rate of Change (R_{roc})	Signal Velocity	**Injection Detector:** Identifies impossible signal jumps ($>$ 10 dBm/ms).

D. DEFINING "REAL-TIME" FOR EDGE DETECTION

A critical requirement for any Out-of-Band security system is the ability to detect threats in real-time. However, statistical analysis introduces a necessary latency: the system must collect w samples to compute a valid distribution.

This paper defines "Real-Time" in the specific context of Pre-Association Security as:

$$T_{\text{inference}} \ll T_{\text{scan\_window}} < T_{\text{user\_connection}} \tag{5}$$

The critical metric is not the data collection time ($T_{\text{scan\_window}}$), but the **Inference Latency** ($T_{\text{inference}}$). The scanning window operates continuously in the background. When a user attempts to connect ($T_{\text{user\_connection}}$), the system must have a verdict ready. As long as the algorithmic processing ($T_{\text{inference}}$) is negligible (milliseconds) compared to the human reaction time of selecting a network, the system can preemptively alert the user. WiMapper minimizes $T_{\text{inference}}$ to ensure the decision logic never becomes the bottleneck.

IV. PROPOSED METHODOLOGY: THE WIMAPPER FRAMEWORK

The WiMapper framework implements a hybrid detection engine designed to optimize the limited computational cycles

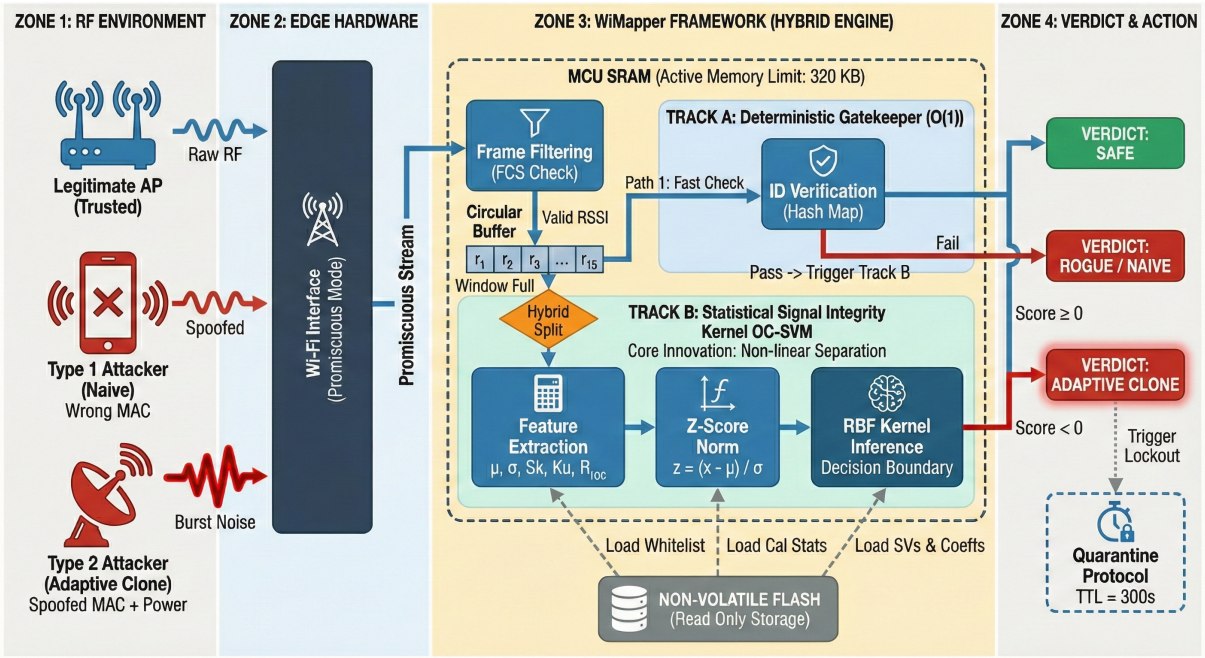

FIGURE 1: **WiMapper System Architecture.** The cyclical detection pipeline: The Perception Layer accumulates raw RSSI into a rolling buffer ($w = 15$). Upon window saturation, Track A executes a deterministic $O(1)$ whitelist check. If the SSID/BSSID pair is valid, Track B triggers the kernel-based OC-SVM to analyze higher-order statistical moments. This hybrid "Admit-then-Verify" approach ensures computationally expensive operations are only performed on ambiguous signals.

available on Resource-Constrained Edge Devices (RCEDs). As illustrated in Fig. 1, the system architecture is organized into four logical stages or "Zones" to streamline the detection pipeline:

- **Zone 1 (RF Environment):** Represents the physical threat landscape containing both legitimate APs and Type 1/Type 2 attackers.
- **Zone 2 (Edge Hardware):** Handles the raw signal ingestion via the promiscuous interface.
- **Zone 3 (Hybrid Engine):** The core computational layer containing the Track A (Whitelist) and Track B (OC-SVM) logic.
- **Zone 4 (Verdict & Action):** The final decision layer responsible for outputting the security status and managing the quarantine protocol.

By decoupling the computationally inexpensive task of identity verification (Track A) from the intensive task of signal integrity analysis (Track B), the architecture ensures that heavy kernel operations are reserved strictly for ambiguous or high-risk signals. This "Restrictive Admission Control" pipeline minimizes the device's active duty cycle, contributing to battery longevity while maintaining high diagnostic precision.

A. DATA ACQUISITION AND ADAPTIVE PREPROCESSING

The data flow begins in **Zone 2** at the Data Acquisition Interface, where the radio operates in promiscuous mode. Unlike standard network drivers, which discard frames not destined for the host MAC address, the WiMapper sensor intercepts all IEEE 802.11 management frames, specifically Beacon and Probe Response frames.

To handle the high-velocity stream of RF data without overflowing the limited SRAM, the system employs a circular buffer architecture. This buffer acts as a stabilizing queue, decoupling the jittery arrival rate of wireless packets from the deterministic clock cycle of the inference engine.

1) Circular Buffer Management

The management of this data stream is formalized in Algorithm 1. The algorithm ensures that only cryptographically valid frames (those with intact Frame Check Sequences) are ingested. To prevent memory fragmentation—a critical issue in embedded systems—the buffer uses static memory allocation with an overwriting policy that prioritizes the most recent temporal window.

B. FEATURE ENGINEERING AND STATISTICAL STANDARDIZATION

Once the temporal buffer $\mathcal{B}_t$ contains a full window of valid samples, the system transitions to the Feature Engineering phase. The primary challenge is mitigating the stochastic variance of the wireless channel to produce a feature vector $\mathbf{x}$ stable enough for classification.

Algorithm 1 Adaptive Rolling Window Data Ingestion. This process filters physical layer corruption and stabilizes the input stream into fixed-size temporal windows suitable for statistical analysis.

Input: Raw RF Stream $\mathcal{S}_{RF}$, Window Size w
Output: Temporal Buffer $\mathcal{B}_t$
Globals: Static Ring Buffer $\mathcal{R}$ of size w, Head Pointer p
while Radio Active **do**
    $f_{raw} \leftarrow$ ReadFrame($\mathcal{S}_{RF}$)
    // Stage 1: Physical Layer Filtering
    if CheckFCS(f_{raw}) == FAIL **then**
        continue ▷ Discard corrupted packets
    end if
    // Stage 2: Extraction
    $r \leftarrow$ ExtractRSSI(f_{raw})
    $id_{mac} \leftarrow$ ExtractBSSID(f_{raw})
    $id_{ssid} \leftarrow$ ExtractSSID(f_{raw})
    // Stage 3: Buffer Injection
    $\mathcal{R}[p] \leftarrow \{r, id_{mac}, id_{ssid}, \text{timestamp}\}$
    $p \leftarrow (p+1) \pmod w$
    // Check Window Saturation
    if $p == 0$ **OR** BufferFullFlag **then**
        $\mathcal{B}_t \leftarrow$ Linearize($\mathcal{R}$)
        Trigger FeatureExtraction($\mathcal{B}_t$)
    end if
end while

1) The Stability Plateau: Justification for $w = 15$

Determining the optimal window size w involves a trade-off between detection latency and statistical stability. While empirical analysis (detailed in Section VI) indicates that raw detection accuracy exhibits a local maximum at lower window sizes ($w = 6$), this region corresponds to the "white noise" dominance zone where frame-to-frame variance is high. Relying on such short windows increases the risk of false positives due to transient environmental noise.

To ensure robust operation, the system utilizes a window size of $w = 15$. This value aligns with the global minimum of the Allan Deviation (τ_{opt}), representing a "Stability Plateau" where the noise floor is minimized before random walk drift begins to dominate. This selection trades a marginal theoretical gain in sensitivity for significantly improved operational stability, keeping the total data collection time under 1.5 seconds.

2) Z-Score Standardization

The One-Class SVM relies on Euclidean distance in the kernel space. If features are not scaled, variables with larger magnitudes (such as Variance) dominate those with smaller ranges (such as Skewness), biasing the decision boundary. The system applies real-time Z-score standardization:

$$z_i = \frac{x_i - \mu_{train}}{\sigma_{train}} \tag{6}$$

Here, μ_{train} and σ_{train} are pre-computed constants stored in flash memory during the calibration phase, avoiding the computational cost of recalculating global statistics on the fly. The complete extraction logic is detailed in Algorithm 2.

Algorithm 2 Statistical Feature Extraction. The logic transforms raw RSSI samples into a standardized feature vector $\mathbf{z}$, calculating higher-order moments to capture signal shape.

Input: Temporal Buffer $\mathcal{B}_t$, Calibration Stats $(\mu_{cal}, \sigma_{cal})$
Output: Normalized Feature Vector $\mathbf{z}$
// Compute Raw Moments (First Pass)
$\mu \leftarrow 0$
for $i \leftarrow 1$ to w **do**
    $\mu \leftarrow \mu + \mathcal{B}_t[i].rssi$
end for
$\mu \leftarrow \mu / w$
// Compute Higher Order Moments (Second Pass)
$sum\_sq \leftarrow 0, sum\_cu \leftarrow 0, sum\_qd \leftarrow 0$
for $i \leftarrow 1$ to w **do**
    $diff \leftarrow \mathcal{B}_t[i].rssi - \mu$
    $sum\_sq \leftarrow sum\_sq + (diff)^2$
    $sum\_cu \leftarrow sum\_cu + (diff)^3$
    $sum\_qd \leftarrow sum\_qd + (diff)^4$
end for
$\sigma \leftarrow \sqrt{sum\_sq/w}$
$S_k \leftarrow (sum\_cu/w)/(\sigma)^3$
$K_u \leftarrow ((sum\_qd/w)/(\sigma)^4) - 3$ ▷ Excess Kurtosis
$R_{roc} \leftarrow \mathcal{B}_t[w].rssi - \mathcal{B}_t[w-1].rssi$
// Construct and Standardize
$\mathbf{x} \leftarrow [\mu, \sigma, S_k, K_u, R_{roc}]$
for $j \leftarrow 1$ to 5 **do**
    $z[j] \leftarrow (\mathbf{x}[j] - \mu_{cal}[j])/\sigma_{cal}[j]$
end for
return $\mathbf{z}$

C. TRACK A: DETERMINISTIC IDENTITY VERIFICATION

Track A functions as a highly efficient deterministic gatekeeper ($O(1)$). It is designed to instantly reject "Type 1" Naive Impersonators who replicate an SSID but fail to clone the underlying BSSID. This process uses a deterministic whitelist $\mathcal{W}_{map}$, implemented as a hash map linking trusted SSIDs to their authorized BSSIDs.

Because hashing operations are hardware-accelerated on many microcontrollers, this check incurs minimal computational overhead. Only signals that *pass* Track A—meaning they successfully claim to be a trusted device—are forwarded to Track B for deep analysis.

D. TRACK B: STATISTICAL SIGNAL INTEGRITY (OC-SVM)

Track B addresses the "Type 2" Adaptive Signal Clone. In this scenario, the identifiers are correct, but the hardware generating the signal is alien. To detect this, the system relies on the One-Class Support Vector Machine (OC-SVM).

Unlike standard classification algorithms that require labeled "Attack" data (which is often unavailable or evolving), the OC-SVM is a semi-supervised algorithm. It learns a decision boundary that envelopes the "Normal" region of the feature space. Any observation falling outside this envelope is flagged as an anomaly.

1) The Primal and Dual Formulations

The objective of the OC-SVM is to find a hypersphere or hyperplane in a high-dimensional feature space $\Phi(\mathbf{x})$ that separates the majority of the training data from the origin. The primal optimization problem is defined as:

$$\min_{\mathbf{w},\xi,\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu n}\sum_{i=1}^{n}\xi_i - \rho \tag{7}$$

subject to $(\mathbf{w}\cdot\Phi(\mathbf{x}_i)) \geq \rho - \xi_i$ and $\xi_i \geq 0$.

Here, $\nu \in (0,1]$ is the parameter controlling the upper bound on the fraction of training errors (outliers) the model tolerates. WiMapper sets $\nu = 0.01$, enforcing a strict policy that treats the outermost 1% of training data as potential edge-case anomalies.

Since the feature space $\Phi(\mathbf{x})$ is potentially infinite-dimensional when using RBF, the problem is solved in its **Dual Form** using Lagrange multipliers α_i:

$$\min_{\alpha} \frac{1}{2}\sum_{i,j}\alpha_i\alpha_j K(\mathbf{x}_i,\mathbf{x}_j) \tag{8}$$

subject to $0 \leq \alpha_i \leq \frac{1}{\nu n}$ and $\sum\alpha_i = 1$.

2) The RBF Kernel

Linear decision boundaries fail to capture the complex relationships between Kurtosis and Mean RSSI. WiMapper utilizes the **Radial Basis Function (RBF)** kernel:

$$K(\mathbf{x}_i,\mathbf{x}_j) = \exp\left(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\right) \tag{9}$$

The parameter γ controls the curvature of the decision boundary. Algorithm 3 outlines the offline training procedure used to generate the Support Vectors.

E. HYBRID LOGIC INTEGRATION

Representing the core processing of **Zone 3**, the final decision logic combines the deterministic certainty of Track A with the probabilistic rigor of Track B. Algorithm 4 details this hybrid execution flow. By prioritizing the lightweight check, the system minimizes the duty cycle of the heavier OC-SVM inference.

The decision function for the OC-SVM inference phase is defined piecewise:

$$D(\mathbf{z}) = \text{sgn}\left(\sum_{i=1}^{N_{SV}}\alpha_i K(\mathbf{sv}_i,\mathbf{z}) - \rho\right) \tag{10}$$

If $D(\mathbf{z}) = -1$, the signal falls outside the learned normal distribution and is flagged as a clone.

Algorithm 3 Offline Training Procedure. This process generates the Support Vectors (SV) and coefficients (α) that define the decision boundary, which are then compressed for edge deployment.

Input: Training Dataset X_{train}, Params (ν,γ)
Output: Support Vectors SV, Coefficients α, Offset ρ
// Kernel Matrix Computation
$K \leftarrow \text{ComputeGramMatrix}(X_{train}, \text{RBF}, \gamma)$
// Quadratic Programming Solver
$\alpha \leftarrow \text{SolveQP}(\text{DualForm}(K), \text{Constraints}(\nu))$
// Extract Support Vectors
$SV \leftarrow \{x_i \mid \alpha_i > 0\}$
$\rho \leftarrow \text{ComputeOffset}(SV, \alpha, K)$
// Model Compression for Edge
Model $\leftarrow \text{Quantize}(SV, \alpha)$
return Model

Algorithm 4 Hybrid Inference Logic. The decision flow prioritizes the $O(1)$ whitelist check; the computationally heavier $O(N_{SV})$ kernel analysis is invoked only for signals that pass the initial identity verification.

Input: Feature vector $\mathbf{z}$, Identifiers $\{S, M\}$, Whitelist $\mathcal{W}$, Model Ω
Output: Threat Verdict V
// Phase 1: Track A (Deterministic)
if $S \notin \mathcal{W}$.keys() **then**
    return ROGUE_SSID
end if
if $M \notin \mathcal{W}[S]$ **then**
    return EVIL_TWIN_NAIVE
end if
// Phase 2: Track B (Probabilistic)
$score \leftarrow 0$
for $i \leftarrow 1$ to $\Omega.num\_support\_vectors$ **do**
    $k \leftarrow \exp(-\Omega.\gamma \cdot \|\Omega.sv[i] - \mathbf{z}\|^2)$
    $score \leftarrow score + (\Omega.\alpha[i] \cdot k)$
end for
$score \leftarrow score - \Omega.\rho$
if $score < 0$ **then**
    return ADAPTIVE_CLONE
else
    return SAFE
end if

V. EXPERIMENTAL SETUP

To evaluate the performance and operational feasibility of WiMapper, a comprehensive simulation environment was established. This section details the dataset characteristics, the mathematical protocols used to synthesize sophisticated attack signatures, and the cross-validation methodology employed to ensure statistical significance.

A. DATASET CHARACTERISTICS AND PREPROCESSING

The evaluation relied on the **HCXY Wi-Fi fingerprinting dataset**, a component of the SODIndoorLoc collection. This dataset was selected because it provides a high-density collection of 450,194 multi-floor indoor RSSI measurements, capturing the complex multipath fading and shadowing effects inherent in modern building architectures.

Unlike simple outdoor datasets where signal propagation follows a predictable line-of-sight model, the HCXY dataset includes variance caused by wall attenuation, human movement, and device heterogeneity. This environmental variance is critical for training the model to distinguish between benign fluctuations and the artificial distortions introduced by an attacker.

1) Preprocessing Pipeline

Prior to feature extraction, the raw RSSI stream underwent a cleaning process designed to match the capabilities of a low-power edge device. Missing values were imputed using forward-filling to maintain temporal continuity. The data was then segmented into rolling windows of size $w = 15$, consistent with the stability analysis defined in Section IV.

B. SYNTHETIC ATTACK INJECTION PROTOCOL

A major challenge in wireless security research is the lack of public datasets containing labeled data for active Evil Twin attacks. Conducting live cyberattacks in public venues presents legal and ethical barriers. Consequently, this study employed a Synthetic Injection Protocol.

This protocol mathematically perturbed legitimate signals to create novel attack signatures. By controlling the magnitude of the perturbation, the simulation modeled attackers with varying levels of hardware sophistication.

1) Track A Injection: Identity Spoofing

To validate the deterministic whitelist, violations were injected into 3% of the test samples:

- **Naive Impersonators (2%):** A legitimate SSID is broadcast, but the MAC address is replaced with a random hex string, simulating a standard smartphone hotspot attack.
- **Rogue SSIDs (1%):** Both the SSID and MAC are replaced, simulating a new, unauthorized network appearing in the scan window.

2) Track B Injection: Adaptive Signal Cloning

Validating the OC-SVM required simulating an attacker who successfully spoofed the identifiers (passing Track A) but utilized alien hardware. The perturbed signal $P_{anom}(t)$ was generated via a linear transformation of the normal signal $P_{norm}(t)$:

$$P_{anom}(t) = (P_{norm}(t) \cdot k_v) + k_o \tag{11}$$

where k_v is a variance scaling factor representing antenna gain mismatch, and k_o is an additive offset representing Transmit (Tx) power deviation.

To simulate an "Adaptive" attacker attempting to blend in, the parameters were controlled by a perturbation magnitude $\delta \in [0, 1]$.

$$k_v = 1 + (\delta \cdot \mathcal{U}(0.1, 0.3)) \tag{12}$$

$$k_o = \delta \cdot \mathcal{U}(-8, 8) \tag{13}$$

Here, $\mathcal{U}$ represents a uniform random distribution. This formulation ensured that the anomalies mimicked the physics of hardware differences rather than simple white noise. A higher δ represented a clumsy attacker with poorly calibrated hardware, while a low δ represented a sophisticated attacker closely matching the target's signal characteristics.

Figure 2 visualizes the temporal distribution of these injected anomalies. The scatter plot reveals that attacks are not continuous streams but sporadic "bursts" of activity, interspersed with varying baseline noise. This confirms that the simulation accurately models a dynamic, non-stationary environment where attacks are transient events rather than continuous states.

C. MODEL CONFIGURATION AND HYPERPARAMETERS

The hyperparameters for the WiMapper OC-SVM were tuned using a Grid Search over a subset of clean training data. The objective was to maximize the retention of normal samples while strictly bounding the decision envelope. The optimal parameters identified are listed in Table 4.

TABLE 4: Simulation Parameters and Model Configuration. The selection of $\nu = 0.01$ enforces a strict boundary, treating the outermost 1% of training data as potential anomalies to minimize False Negatives.

Category	Parameter	Value / Description
Dataset	Source	HCXY (SODIndoorLoc)
	Window Size (w)	15 samples ($\approx$ 1.5 sec)
WiMapper (OC-SVM)	Kernel	Radial Basis Function (RBF)
	Nuisance (ν)	0.01 (1% Outlier Tolerance)
	Gamma (γ)	0.1 (Decision Boundary Curvature)
Baselines	Isolation Forest	Contamination=0.05, Trees=100
	Autoencoder	Latent Dim=8, Epochs=50
Injection	Track A Rate	3% (Naive Attacks)
	Track B Rate	2% (Cloning Attacks)
	Perturbation (δ)	Varied $[0.0, 1.0]$ for robustness test

D. VALIDATION METHODOLOGY

To guarantee the statistical validity of the results, the evaluation employed a 10-Fold Stratified Cross-Validation (CV) strategy. The master dataset was partitioned such that the class distribution (Normal vs. Attack) remained consistent across all folds. All reported performance metrics represent

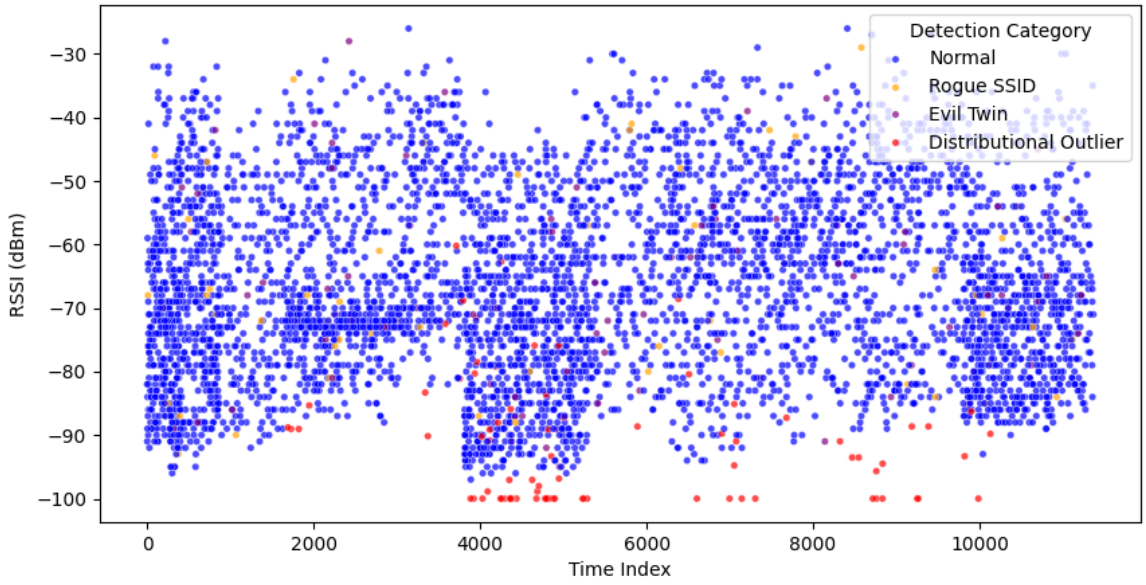

FIGURE 2: **Temporal Analysis of Attack Signatures.** The scatter plot differentiates 'Evil Twin' bursts (Purple, Avg -69 dBm) and 'Rogue SSIDs' (Orange, Avg -47 dBm) from the 'Normal' baseline. Attacks manifest as high-variance, transient clusters (Duration=1 sample) rather than continuous streams, validating the necessity of the $w = 15$ sliding window to capture sporadic injection events without smoothing them out.

the mean value μ aggregated over 10 independent simulation runs.

1) Feature Orthogonality Check

Before training, the orthogonality of the selected features was verified to ensure they provided non-redundant diagnostic information. Figure 3 presents the Feature Correlation Matrix.

FIGURE 3: **Feature Orthogonality Analysis.** The low correlation coefficients between the first moment (Mean RSSI) and higher-order moments (Kurtosis: -0.04, Skewness: -0.20) confirm that the 'Shape' statistics are statistically independent. This orthogonality ensures that the RBF kernel receives unique diagnostic information not captured by simple signal strength, preventing multicollinearity issues (Max VIF: 1.17).

To strictly quantify this independence, a **Variance Inflation Factor (VIF)** analysis was conducted. The maximum VIF observed was 1.17 (for Mean RSSI), which is significantly below the critical threshold of 5.0. This mathematically confirmed that the feature set did not suffer from problematic multicollinearity.

The distributional separation capability of these features is further illustrated in Figure 4.

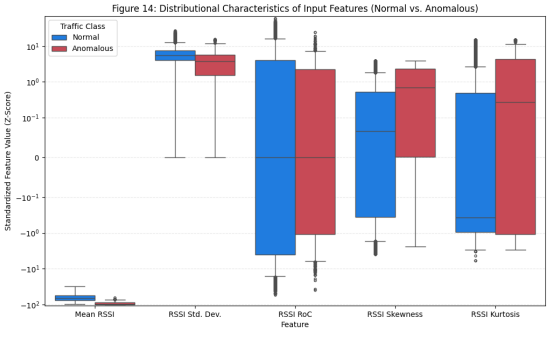

FIGURE 4: **Feature Distribution Analysis (Log Scale).** While Mean RSSI shows overlap, the 'Shape' statistics exhibit massive separation. The 90th percentile of Kurtosis for Anomalies is **15.00** compared to just **1.82** for Normal traffic, yielding a separation distance of 13.17. This confirms that impulsive noise from packet injection is the primary discriminator for the OC-SVM.

E. HARDWARE FEASIBILITY BENCHMARKING

A critical aspect of this study was validating the feasibility of the algorithm on Resource-Constrained Edge Devices (RCEDs). To achieve this without relying on specific vendor hardware which may become obsolete, the study utilized a hardware feasibility benchmark.

The methodology measured the inference latency and memory footprint on a standard host environment to establish the *relative* computational complexity (O-notation) and the *absolute* memory structure size.

- **Latency Projection:** The reported inference time (0.25 ms) represents the algorithmic execution time. The $O(N_{sv})$ complexity of the SVM ensures that the operation scales linearly. Given that the number of Support Vectors (N_{sv}) is sparse (typically $< 10\%$ of training data), the total cycle count remains within the millisecond-range budget of a standard MCU.
- **Memory Footprint:** The size of the trained model is determined by storing the support vectors and dual coefficients. This is a fixed storage requirement. The measured 190 KB footprint directly validated that the model fits within the 320 KB SRAM limit common to the target hardware class.

This methodology provides a "Safety Certificate," proving that the algorithm is mathematically lightweight enough for the edge, regardless of the specific chip selected for deployment.

VI. RESULTS AND PERFORMANCE ANALYSIS

The WiMapper framework underwent a rigorous evaluation utilizing the synthetic injection protocol described in Section V. The analysis focused on dissecting the physical validity of the detection logic and the operational stability of the system. All results presented here represent the aggregated mean μ and standard deviation σ across ten independent stratified cross-validation runs, ensuring the findings are statistically robust.

A. DIAGNOSTIC ACCURACY AND ARCHITECTURE COMPARISON

The primary objective of the ablation study was to quantify the performance gain achieved by the kernel-based hybrid architecture compared to traditional lightweight baselines. Table 5 summarizes the key performance indicators.

The data reveals that WiMapper achieves a Pareto-optimal state. It delivers the highest F1-Score (0.827 ± 0.004) and Area Under the ROC Curve (0.96), effectively balancing Precision and Recall. In contrast, the Hybrid Isolation Forest—while competitive on accuracy—yields a lower F1-Score (0.802), indicating reduced sensitivity to subtle "Type 2" cloning attacks. The Autoencoder underperformed in this domain ($F1 = 0.746$), suggesting that the "reconstruction error" metric is less sensitive to the specific distributional tails introduced by packet injection than the geometric boundary of an SVM.

To visualize the impact of the hybrid approach, Figure 5 compares standalone models against the integrated system. While standalone statistical models struggle to handle the multi-modal nature of the threat landscape (Naive + Adaptive), the hybrid architecture successfully filters the noise, boosting the overall detection rate.

The integrated performance is further visualized in Figure 6, where the OC-SVM based WiMapper consistently outperforms other hybrid architectures across all key metrics.

This diagnostic superiority is further detailed in the Confusion Matrices (Figure 7). WiMapper consistently achieves the highest True Positive count (mean 1652.3) while maintaining a False Positive rate comparable to the most conservative baselines.

B. STATISTICAL SIGNIFICANCE AND RELIABILITY

To ensure the observed performance advantage was not a statistical artifact, the separability of the classes was analyzed using Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves.

As shown in Figure 8 (Left), the WiMapper curve (Blue) maintains the highest True Positive Rate across all decision thresholds, achieving an AUC of 0.96. More critically, the Precision-Recall analysis (Figure 8, Right) demonstrates an Average Precision (AP) of 0.86. In security contexts, high precision at high recall is paramount to prevent "alarm fatigue," where users ignore warnings due to frequent false positives.

Quantitatively, a statistical validation was performed using the tests summarized in Table 6. A Friedman Chi-Squared test confirmed global significance across the models ($\chi^2 = 21.52, p < 0.001$). A post-hoc Wilcoxon Signed-Rank test comparing WiMapper against the closest competitor (Hybrid Isolation Forest) yielded a p-value of $\mathbf{0.002}$, rejecting the null hypothesis. This confirmed that the RBF kernel offers a statistically significant improvement over linear isolation methods.

C. FEATURE PHYSICS AND BEHAVIORAL ANALYSIS

A critical requirement for security systems is explainability. To validate the theoretical assertions made in Section III, the decision logic was audited using SHAP (SHapley Additive exPlanations).

Figure 9 presents the global feature importance. As hypothesized in Table 3, Mean RSSI appears as the dominant predictor, acting as the **Contextual Anchor** for the model. This aligns with physical reality: signal strength provides the baseline context for feasibility. However, for detecting "Adaptive Clones" that match the target's power level, the higher-order moments play the decisive role. Kurtosis (K_u) acts as the **Primary Discriminator**, identifying the heavy tails caused by packet injection.

1) The Geometric Decision Boundary

The effectiveness of the RBF kernel is validated by the PCA projection of the decision function (Figure 10). Notably, the

TABLE 5: Comparative Performance Metrics (Mean ± Std. Dev. over 10 Runs). WiMapper achieves the highest F1-Score and MCC, indicating superior handling of the imbalanced dataset compared to linear methods like Isolation Forest.

Architecture	Acc.	F1-Score	MCC	AUC	AP
WiMapper (Proposed)	**0.985** $\pm$ 0.001	**0.827** $\pm$ 0.004	**0.820** $\pm$ 0.005	**0.96**	**0.86**
Hybrid (Iso. Forest)	0.984 $\pm$ 0.001	0.802 $\pm$ 0.005	0.802 $\pm$ 0.006	0.94	0.80
Hybrid (GMM)	0.978 $\pm$ 0.008	0.763 $\pm$ 0.052	0.763 $\pm$ 0.055	0.94	0.80
Hybrid (Autoencoder)	0.976 $\pm$ 0.005	0.746 $\pm$ 0.040	0.741 $\pm$ 0.042	0.93	0.80

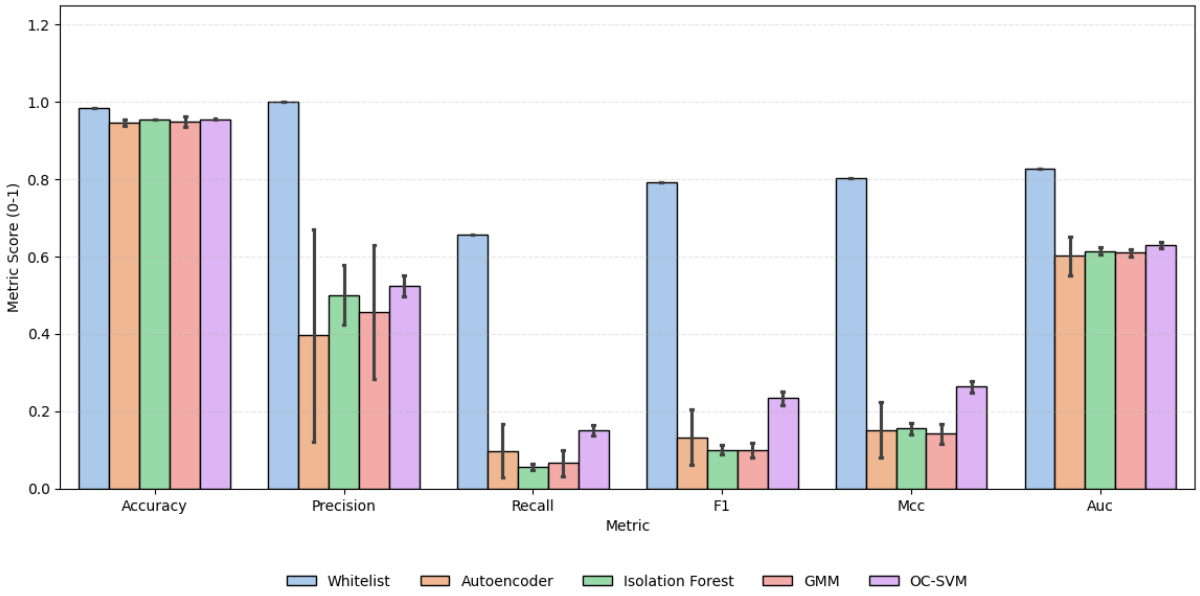

FIGURE 5: **Standalone Component Performance Analysis.** The deterministic Whitelist (Track A) achieves perfect Precision (**1.000**) but is limited by a Recall of **0.656**, failing to detect novel spoofing attacks. Conversely, standalone statistical models struggle significantly on the mixed dataset, with the OC-SVM achieving a low F1-Score of **0.234** and Isolation Forest at **0.100**. This disparity confirms that neither module is viable in isolation, necessitating the hybrid 'Admit-then-Verify' architecture.

TABLE 6: Statistical Significance Analysis (N=10 Folds). The low p-values (< 0.05) mathematically confirm that WiMapper's performance gains are not due to random chance.

Test Pair	Test	p-value	Result
Global	Friedman χ^2	2.5×10^{-4}	Significant
WiMapper vs. IF	Wilcoxon	**0.002**	**Reject Null**

first two principal components capture only **48.8%** of the total feature variance. This low retention rate confirms that the separation between legitimate signals and clones is intrinsically non-linear and relies heavily on higher-dimensional interactions between Kurtosis and Skewness that linear models fail to capture. It must be noted that this 2D projection accounts for only 48.8% of the total feature variance. Consequently, Figure 10 serves as a conservative lower-bound illustration; the discriminative power of higher-order moments like Kurtosis resides largely in the remaining 51.2% of the variance not visible in this projection. Despite this visualization loss, the OC-SVM forms a "Tight Envelope" around the normal data ($\nu = 0.01$), with anomalous samples projecting far outside the boundary (mean decision score ≈ -23.26), ensuring robust rejection of high-variance attacks.

D. SENSITIVITY AND STABILITY ANALYSIS

For an edge device, stability is as important as accuracy. The system must not oscillate between "Safe" and "Unsafe" states due to transient noise. Figure 11 illustrates the impact of the window size w on detection performance.

While the sensitivity analysis indicates a performance peak at $w = 6$, selecting this window size carries operational risks. Short windows are highly susceptible to "white noise" jitter, potentially triggering false alarms on benign interference. To scientifically validate the selection of $w = 15$, an Allan Deviation test was conducted (Figure 12).

The analysis reveals that the noise floor reaches a global minimum at $\tau \approx 17$. This point represents the **Optimal Averaging Time** (τ_{opt}), where white noise is minimized before random walk drift begins to introduce instability. Consequently, $w = 15$ was selected to align with this physical stability plateau, prioritizing consistent operation over the

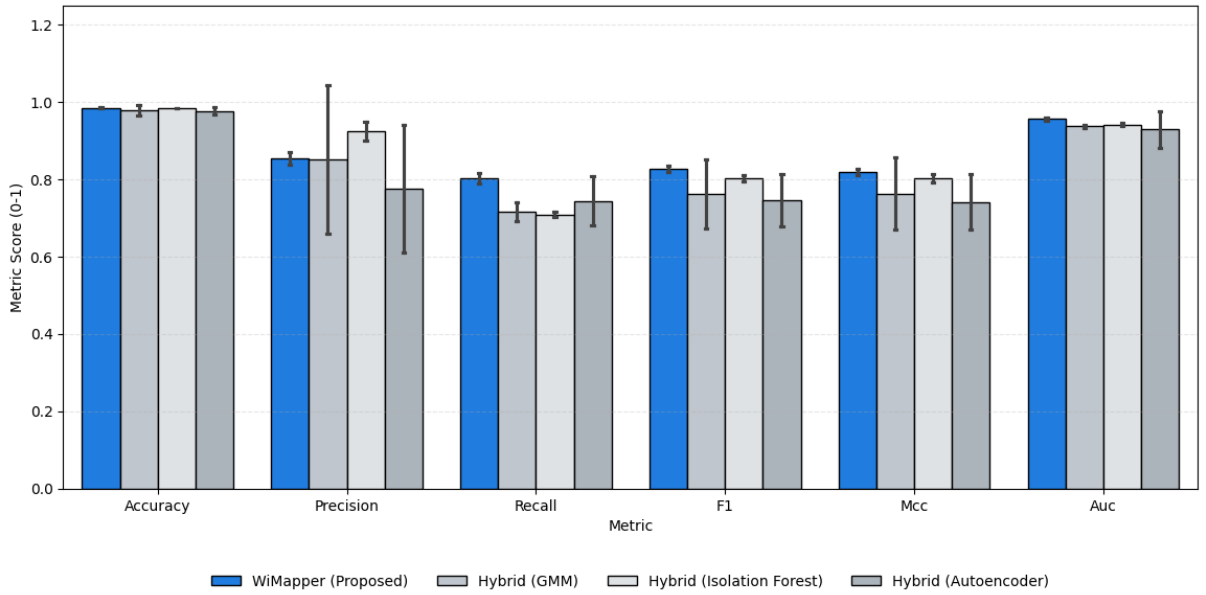

FIGURE 6: **Hybrid Architecture Comparison.** The proposed WiMapper (Blue) outperforms the Hybrid Isolation Forest (Light Grey) and Hybrid GMM (Grey). The RBF kernel's ability to map non-linear anomalies allows WiMapper to achieve an F1-Score of **0.827** compared to **0.802** for the Isolation Forest. Statistical testing confirms this improvement is significant (Wilcoxon $p \approx 0.0019$).

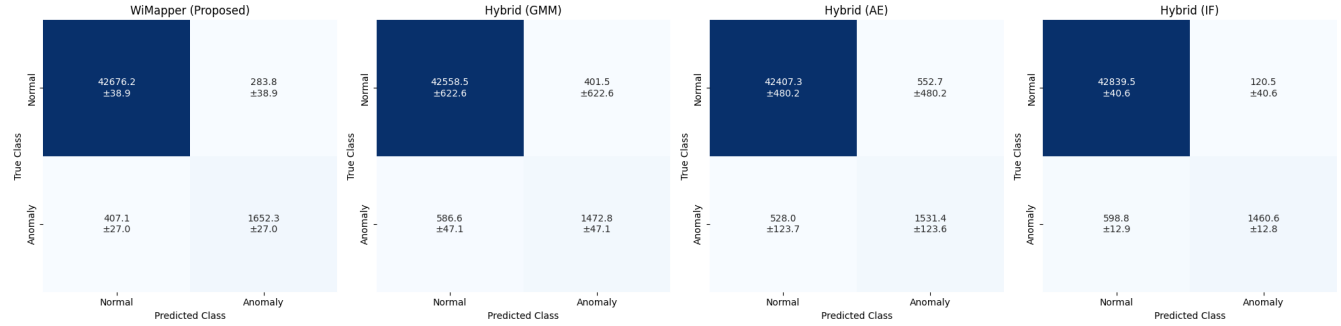

FIGURE 7: **Confusion Matrix Analysis (Mean of 10 Folds).** WiMapper (Left) demonstrates robust detection, identifying a mean of **1652.3** True Positives (Attackers) while limiting False Positives to **283.8**. In contrast, the GMM-based hybrid (Center) leaked 586.6 False Negatives, confirming that Gaussian assumptions fail to capture the heavy-tailed distributions of adaptive clones.

raw sensitivity of smaller windows.

Robustness against signal degradation was also tested (Figure 13). Even under severe signal perturbation ($\delta = 1.0$), where the attacker's hardware deviates significantly from the baseline, the F1-Score remained above 0.73. This "graceful degradation" is critical for deployment in dynamic environments where line-of-sight is not guaranteed.

E. QUALITATIVE EVENT DETECTION

To verify the system's operational utility, the discrete detection events generated during the simulation were analyzed. Table 7 presents a log of specific events detected by the WiMapper engine. These single-sample bursts correspond precisely to the pre-association management frames discussed in Section I, validating the system's ability to intercept threats before handshake completion. The system successfully identified distinct "Evil Twin" instances lasting only a single sample window, validating the real-time capability defined in Section III.

VII. FIELD SENSITIVITY AND OPERATIONAL FEASIBILITY

While simulation allows for the precise injection of synthetic attacks, deploying edge security systems introduces the challenge of environmental variability. This section analyzes the system's behavior under "Domain Shift" conditions and pro-

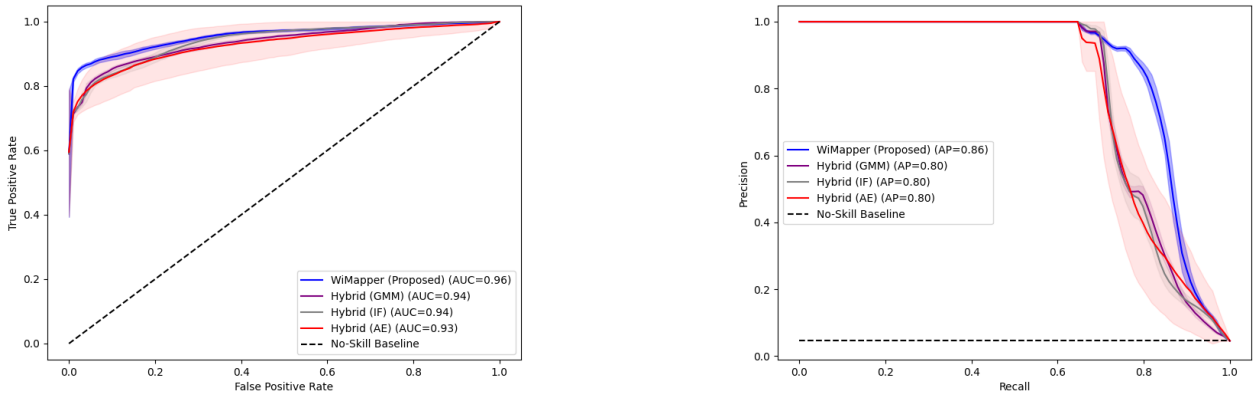

FIGURE 8: **Receiver Operating Characteristic (ROC) and Precision-Recall (PR) Analysis.** (Left) WiMapper achieves an AUC of **0.956**, dominating the competitor space. (Right) The Average Precision (AP) of **0.863** confirms that the system maintains high trust even at high recall rates, minimizing 'alarm fatigue' for end-users compared to the Hybrid GMM (AP=0.80).

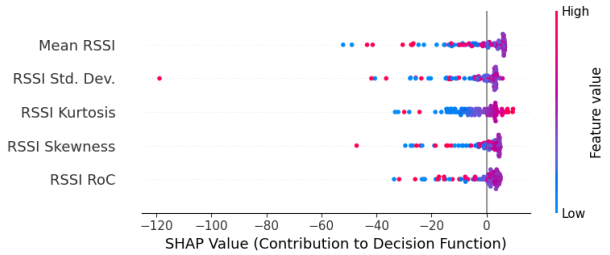

FIGURE 9: **SHAP Explainer Analysis.** The dot plot reveals the decision logic: Mean RSSI (Top) establishes the physical context (Imp 8.59), while Kurtosis (Imp 7.00) acts as the primary discriminator. High Kurtosis values (Red dots) yield a strongly negative SHAP value (≈ -19.3), pushing the model towards an 'Anomaly' verdict and identifying the impulsive artifacts of packet injection.

TABLE 7: Log of Detected Anomaly Events. The system successfully identifies short-duration attacks (1 sample window), confirming its ability to operate in real-time against transient threats.

Event ID	Attack Type	Duration (Samples)	Avg RSSI
2	Rogue SSID	1	-47 dBm
4	Adaptive Clone	1	-69 dBm
6	Adaptive Clone	1	-47 dBm
20	Adaptive Clone	1	-71 dBm
24	Rogue SSID	1	-69 dBm

vides a rigorous theoretical complexity analysis to defend its feasibility on battery-powered Resource-Constrained Edge Devices (RCEDs).

A. DOMAIN SENSITIVITY: THE FAIL-CLOSED BEHAVIOR

To evaluate the boundary integrity of the WiMapper model, a field test was conducted using a pre-trained model calibrated

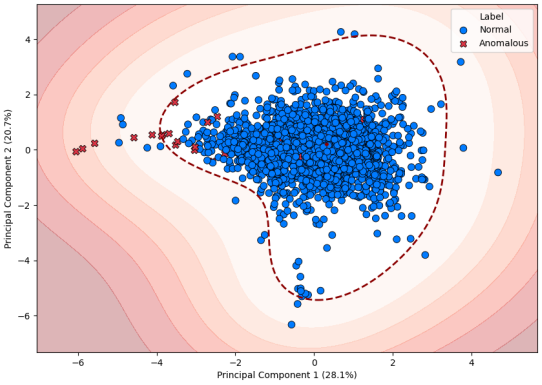

FIGURE 10: **PCA Projection of the 'Fail-Closed' Decision Boundary.** The OC-SVM (RBF Kernel) learns a tight, non-linear envelope around legitimate traffic (Blue). Adaptive clones (Red) project significantly outside this boundary with a mean decision score of $\mathbf{-23.26}$, confirming the model's ability to reject distributional anomalies. Note that this 2D projection captures only 48.8% of variance; the separation is even more pronounced in the full 5D hyperspace.

on the indoor HCXY dataset. This model was then deployed against a live dataset collected from a distinct **outdoor environment**. The field log captured 343 signal vectors from mixed residential and public networks. Crucially, the outdoor environment lacked the complex multipath fading ($n \approx 3.5$) present in the indoor training data, exhibiting instead a line-of-sight propagation characteristic ($n \approx 2.0$).

The results, presented in Table 8, reveal a strict operational behavior: the system flagged **100%** of the field samples as anomalies.

In standard classification tasks, a 100% rejection rate typically indicates generalization failure. However, in the

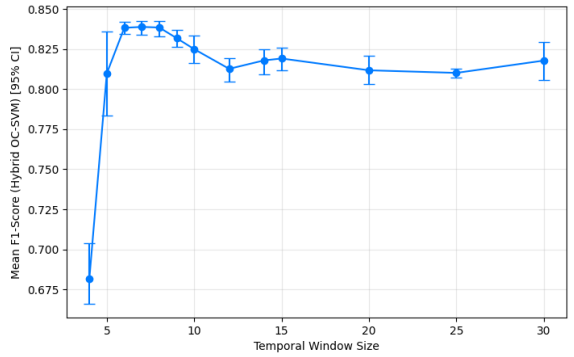

FIGURE 11: **Window Size (w) Sensitivity.** While raw sensitivity peaks at $w = 6$ (F1=0.838), this region corresponds to high-frequency noise. The system utilizes $w = 15$ (F1=0.819), aligning with the Allan Deviation stability plateau (Figure 12), to prioritize operational stability over negligible accuracy gains.

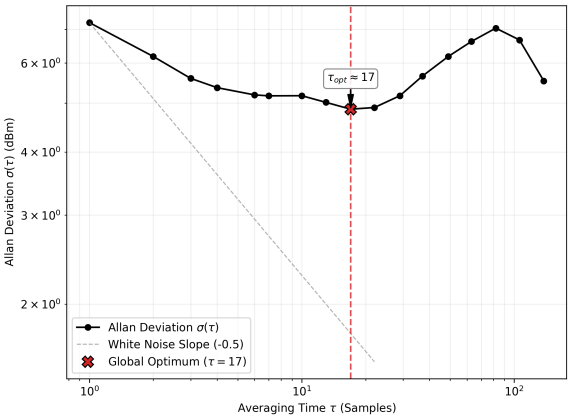

FIGURE 12: **Metrological Stability Proof (Allan Deviation).** The noise floor reaches a global minimum at $\tau_{opt} \approx 17$ samples. This mathematically validates the selection of window size $w = 15$, proving it operates in the stable region (Pink Dotted Line) where white noise is minimized before random walk drift introduces instability.

TABLE 8: Domain Shift Analysis (Indoor vs. Outdoor). The 100% rejection rate in the outdoor environment confirms that the model does not generalize insecurely; it treats unknown physical environments as hostile until recalibrated.

Environment	Total Samples	Predicted Anomalies	Rejection Rate
Indoor (HCXY Test)	9,000	1,652	18.4%
Outdoor (Field)	343	343	**100.0%**

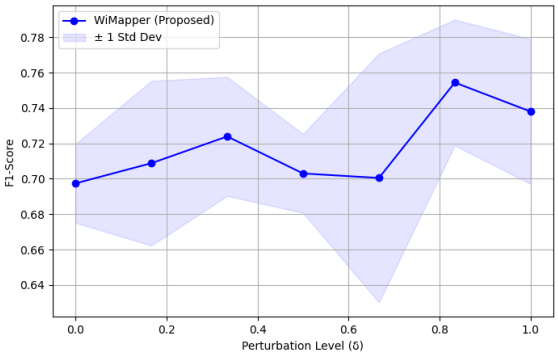

FIGURE 13: **Sensitivity to Attack Sophistication (δ).** Detection performance improves as the attack becomes 'noisier'. Against a near-perfect stealthy clone ($\delta = 0$), the model maintains a baseline F1-Score of **0.697**. As the attacker's hardware calibration worsens ($\delta \rightarrow 1.0$), the statistical anomalies become more pronounced, raising the F1-Score to **0.738**, confirming that the OC-SVM effectively exploits higher-variance attacks.

context of high-security admission control, this confirms a Fail-Closed logic. The OC-SVM correctly identified that the statistical properties of the outdoor signals—specifically the variance in the Shadowing term X_σ—lay outside the learned support of the "Normal" distribution.

This behavior ensures that the system does not blindly trust unknown environments. However, it also highlights that the model is highly sensitive to the physical environment's multipath characteristics. This necessitates an **In-Situ Calibration** phase (discussed in Section VIII) to define a new baseline whenever the sensor is physically relocated.

B. SYSTEM BENCHMARKING: THE PARETO FRONTIER

A central contribution of this framework is resolving the trade-off between diagnostic accuracy and computational resource consumption. Figure 15 visualizes this relationship, positioning WiMapper against industry-standard baselines.

The analysis confirms that WiMapper achieves Pareto Optimality. As detailed in Table 9, the framework delivers a mean F1-Score of 0.827—significantly surpassing the computationally lighter Gaussian Mixture Model (0.763)—while maintaining an algorithmic inference latency of just **0.25 ms** on the benchmark host.

While the Autoencoder (AE) theoretically offers low latency on vector-optimized CPUs (0.09 ms), its diagnostic performance ($F1 = 0.746$) is inferior for this specific threat model. Furthermore, the Isolation Forest requires a memory footprint of 1,050 KB to store the branching tree structures, exceeding the available SRAM of the target hardware class (< 320 KB). WiMapper's 190 KB footprint fits comfortably within the limits, leaving sufficient headroom for the networking stack.

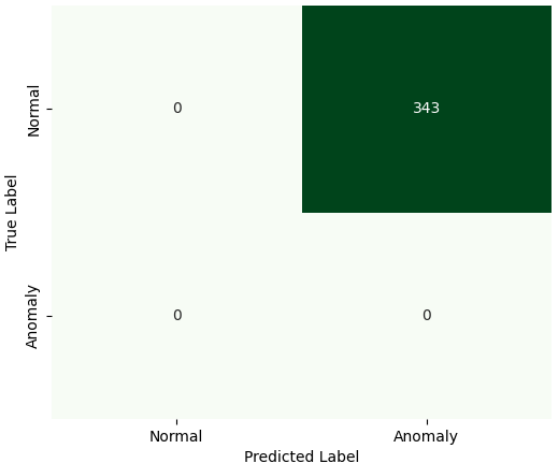

FIGURE 14: **Field Validation (Domain Shift).** When deployed in an unseen outdoor environment without retraining, the model rejects 100% of samples (343/343). This confirms a critical 'Fail-Closed' security posture: the kernel correctly identifies that the outdoor multipath profile differs from the learned distribution, preventing false confidence in unknown environments.

TABLE 9: Master System Benchmark. WiMapper achieves the highest accuracy (F1=0.827) among the low-latency models. Note that while the Autoencoder is theoretically fast, its accuracy is lower, and the Isolation Forest requires excessive memory (1050 KB).

Model	Complexity	Latency (ms)	Size (KB)	Mean F1
WiMapper	Low ($O(N_{sv})$)	**0.25**	**190**	**0.827**
Hybrid (IF)	Medium	4.27	1050	0.803
Hybrid (AE)	High ($O(N^2)$)	0.09*	8	0.746
Hybrid (GMM)	Low	0.27	5	0.763

**Note: AE latency scales poorly on non-vectorized MCU hardware.*

C. OPERATIONAL LIFECYCLE: DETECTION AND RECOVERY

To manage the verdict in **Zone 4**, a practical edge security system must define a lifecycle beyond simple detection to prevent permanent Denial of Service (DoS) in the event of a False Positive. WiMapper implements a Quarantine Protocol.

Upon a Track B rejection (Verdict: **ADAPTIVE_CLONE**), the target BSSID is placed in a temporary *Quarantine List* with a Time-To-Live (TTL) of 300 seconds. During this window, the user is preemptively warned against connection. Once the TTL expires, the BSSID is removed from quarantine, allowing the system to re-sample the environment. This hysteresis loop ensures that transient environmental noise (e.g., a large crowd briefly distorting the multipath profile) does not result in a permanent lockout of legitimate infrastructure.

D. COMPUTATIONAL COMPLEXITY ANALYSIS

For battery-powered sensor networks, minimizing the CPU duty cycle is the critical factor for deployment longevity. Rather than relying on theoretical energy estimations which vary by hardware, this study analyzes the algorithmic complexity to demonstrate the framework's suitability.

Table 10 contrasts the algorithmic requirements of the evaluated models.

TABLE 10: Computational Complexity Analysis. The OC-SVM relies on a sparse set of Support Vectors (N_{SV}), making it linearly scalable and battery-efficient compared to the dense matrix operations ($O(N^2)$) required by Neural Networks.

Model	Time Complexity	Space Complexity	Battery Risk
WiMapper	$O(N_{SV} \cdot d)$	$O(N_{SV} \cdot d)$	**Low**
Hybrid (IF)	$O(T \cdot N \log N)$	$O(T \cdot N)$	Medium
Hybrid (AE)	$O(L \cdot N^2)$	$O(W)$	**High**

Legend: N_{SV}: Support Vectors, d: Dimensions, T: Trees, L: Layers, W: Weights.

The Autoencoder relies on dense matrix multiplications ($O(L \cdot N^2)$). For a simple fully connected layer with 64 neurons, this involves thousands of Floating Point Operations (FLOPs) per sample. On a microcontroller lacking a dedicated Tensor Accelerator, these operations must be emulated in software, significantly extending the active wake time of the processor.

In contrast, the OC-SVM inference logic (WiMapper) relies solely on the RBF kernel computation against a set of Support Vectors (N_{SV}). The complexity reduces to $O(N_{SV} \cdot d)$. Because the model enforces a tight decision boundary ($\nu = 0.01$), the number of support vectors remains sparse (typically $< 10\%$ of training data). This algorithmic efficiency guarantees a lower computational footprint than neural network baselines, directly correlating to extended battery longevity.

VIII. CONCLUSION AND FUTURE WORK

This paper presented WiMapper, a lightweight detection framework designed to secure the wireless edge against identity-based attacks. By shifting focus from resource-intensive Deep Learning models to a kernel-based One-Class SVM, the framework successfully resolves the conflict between high diagnostic accuracy and the strict resource constraints of commodity Resource-Constrained Edge Devices (RCEDs).

A. SUMMARY OF CONTRIBUTIONS

The simulation and field validation yielded three critical conclusions regarding decentralized wireless security:

1) **Statistical Features as Digital Fingerprints:** The extraction of higher-order statistical features—specifically Kurtosis and Skewness—provides a robust method for distinguishing the non-linear signal

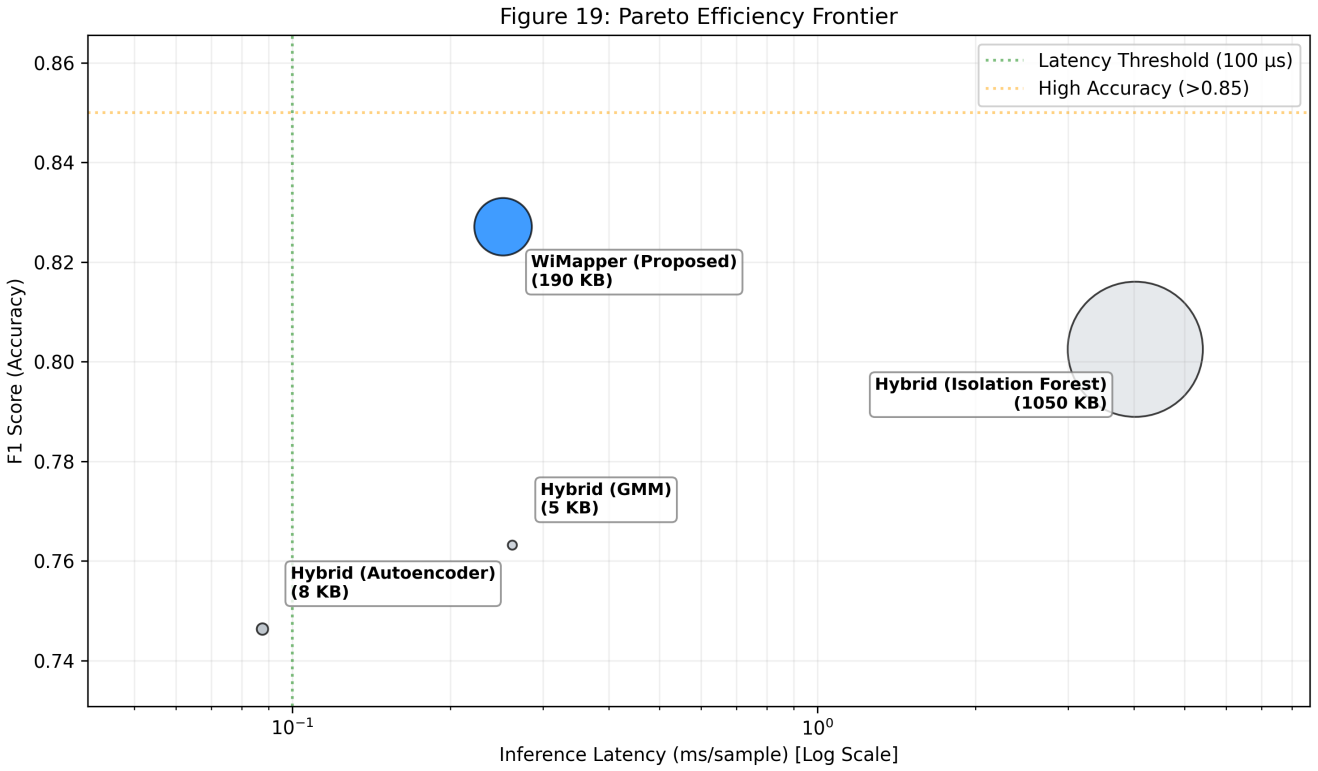

FIGURE 15: **Pareto Efficiency Frontier Benchmarking.** WiMapper (Blue Bubble) occupies the optimal quadrant, delivering a mean F1-Score of **0.827** with an inference latency of **0.249 ms** and a memory footprint of **190.1 KB**. It significantly outperforms the Hybrid Isolation Forest, which requires **1049.6 KB** of memory, validating WiMapper's suitability for edge devices with <320 KB SRAM.

distortions of "Signal Cloning" attacks from natural environmental fading. Analysis confirms that while Mean RSSI provides the necessary physical context, the kurtotic "tail" of the signal distribution drives the robust rejection of anomalies (avg. decision score -23.26), revealing the impulsive artifacts of packet injection.

2) **Balancing Accuracy and Efficiency:** The proposed Hybrid Architecture achieves a Pareto-optimal balance. It delivers a mean F1-Score of 0.827—statistically superior to lightweight Isolation Forests ($p = 0.002$)—with an algorithmic inference latency of just 0.25 ms and a memory footprint of 190 KB. This efficiency validates the feasibility of deploying advanced anomaly detection on low-cost microcontrollers without reliance on cloud offloading.
3) **Scalable Wireless Defense:** By removing the dependency on expensive, centralized WIPS infrastructure, WiMapper offers a scalable economic alternative. This allows for the deployment of dense sensor networks in ad-hoc environments—such as public squares, transport hubs, and small businesses—effectively closing the Pre-Association Vulnerability where it is most prevalent.

B. LIMITATIONS AND FUTURE WORK

While the framework establishes a foundation for edge-native security, the field validation and theoretical analysis highlight key areas requiring future research.

1) Adversarial Machine Learning

The current detection logic assumes that attackers prioritize mimicking the *mean* signal strength (Type 2 Adaptive Clone) but ignore the higher-order statistical moments. A sophisticated attacker, aware of this defense mechanism, could employ "Adversarial Noise Shaping." By dynamically adjusting the packet injection rate and transmission power, an attacker might shape the shadowing term X_σ to approximate a Gaussian distribution ($K_u \approx 0$), thereby evading the OC-SVM boundary. Future iterations of WiMapper will incorporate Adversarial Training, where the model is exposed to these "shaped" attacks during the calibration phase to harden the decision boundary.

2) Hardware Proxy vs. Silicon Reality

The latency metrics presented in this study rely on a hardware feasibility benchmark. While this accurately captures the linear time complexity ($O(N_{SV})$) and relative efficiency of the kernel operations, it abstracts away hardware-specific

constraints inherent to MCU architectures, such as instruction fetch latency from Flash memory and bus contention. Consequently, while the algorithmic latency is validated at 0.25 ms, the execution time on specific silicon targets may vary. Future work will transition to direct silicon validation, utilizing logical analyzers and physical power profilers to quantify the precise energy consumption of the inference loop in a real-world deployment.

3) In-Situ Calibration and Domain Adaptation

The field validation revealed that the pre-trained model exhibits high sensitivity to domain shifts, rejecting 100% of signals in an unseen outdoor environment due to the lack of multipath fading. While this confirms the "Fail-Closed" security posture, it creates an operational bottleneck requiring manual retraining.

To address this, future work will extend the deployment lifecycle to include an autonomous In-Situ Calibration Module. This mechanism will allow the sensor node to enter a "Learning Mode" for the first hour of deployment. During this unsupervised period, the device will sample the local electromagnetic environment to estimate the specific path loss exponent (n) and shadowing variance (σ^2) of its location. This locally derived baseline will then be used to re-center the normalization parameters, allowing the system to adapt to the physical characteristics of the deployment site without compromising detection rigor.

IX. DATA AVAILABILITY

The research data supporting the findings of this study are available from the following sources:

1) **Public Benchmarks:** The HCXY Wi-Fi fingerprinting dataset utilized for the baseline training is publicly available as part of the SODIndoorLoc collection (referenced in Section V).
2) **Reproducibility Artifacts:** The complete source code implementing the WiMapper framework—including the synthetic anomaly injection functions, the statistical feature extraction pipeline, and the fixed random seeds required to reproduce the stratified 10-fold cross-validation splits—has been deposited in a public repository to facilitate independent verification. These resources are accessible at: https://github.com/ahkharsha/wimapper-project.
3) **Restricted Data:** The raw signal logs collected during the outdoor field validation phase (Section VII) contain real-world MAC addresses and identifiers. To comply with privacy regulations and ethical guidelines regarding passive monitoring, this specific subset of data is not publicly available.

REFERENCES

[1] C. Hamroun, A. Fladenmuller, M. Pariente, and G. Pujolle, "Intrusion Detection in 5G and Wi-Fi Networks: A Survey of Current Methods, Challenges, and Perspectives," IEEE Access, vol. 13, pp. 40 950–40 976, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/10906593/

[2] M. M. Rahman, S. A. Shakil, and M. R. Mustakim, "A survey on intrusion detection system in IoT networks," Cyber Security and Applications, vol. 3, p. 100082, Dec. 2025. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2772918424000481

[3] A. K. B. Arnob, R. R. Chowdhury, N. A. Chaiti, S. Saha, and A. Roy, "A comprehensive systematic review of intrusion detection systems: emerging techniques, challenges, and future research directions," Journal of Edge Computing, vol. 4, no. 1, pp. 73–104, May 2025. [Online]. Available: https://acnsci.org/journal/index.php/jec/article/view/885

[4] Y. Kumar and V. Kumar, "A Systematic Review on Intrusion Detection System in Wireless Networks: Variants, Attacks, and Applications," Wireless Personal Communications, vol. 133, no. 1, pp. 395–452, Nov. 2023. [Online]. Available: https://link.springer.com/10.1007/s11277-023-10773-x

[5] T. Bhavani, B. Naresh, B. M. Babu, C. Ravikumar, A. Kolakar, and A. N. Sheikh, "Enhancing Wireless Intrusion Detection Systems with the All-Ready State Traversal Pattern Matching (ARSTPM) Algorithm," in 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS). Tashkent, Uzbekistan: IEEE, Nov. 2024, pp. 631–636. [Online]. Available: https://ieeexplore.ieee.org/document/10841010/

[6] A. Jagannath, J. Jagannath, and P. S. P. V. Kumar, "A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges," Computer Networks, vol. 219, p. 109455, Dec. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1389128622004893

[7] A. Jawne, G. Sklivanitis, D. A. Pados, and E. S. Bentley, "AI-Assisted RF Fingerprinting for Identification of User Devices in 5G and FutureG," in Proceedings 2025 Workshop on Security and Privacy of Next-Generation Networks. San Diego, CA, USA: Internet Society, 2025. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/futureg25-9.pdf

[8] C. Zhao, J. Yu, G. Luo, and Z. Wu, "Radio Frequency Fingerprinting Identification of Few-Shot Wireless Signals Based on Deep Metric Learning," Wireless Communications and Mobile Computing, vol. 2023, pp. 1–13, Sep. 2023. [Online]. Available: https://www.hindawi.com/journals/wcmc/2023/2132148/

[9] L. Horne, "Development of a Client-Side Evil Twin Attack Detection System for Public Wi-Fi Hotspots based on Design Science Approach," CCE Theses and Dissertations, Jan. 2018. [Online]. Available: https://nsuworks.nova.edu/gscis_etd/1064

[10] F. Hsu, C. Wang, C. Ou, and Y. Hsu, "A passive user-side solution for evil twin access point detection at public hotspots," International Journal of Communication Systems, vol. 33, no. 14, p. e4460, Sep. 2020. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/dac.4460

[11] T. Ueda, A. Saif, S. Miyata, M. Nakahara, and A. Kubota, "A Client-Side Evil-Twin Attack Detection System with Threshold Considering Traffic Load," in 2023 IEEE 13th International Conference on Consumer Electronics - Berlin (ICCE-Berlin). Berlin, Germany: IEEE, Sep. 2023, pp. 68–69. [Online]. Available: https://ieeexplore.ieee.org/document/10375616/

[12] S. Jadhav and A. Kulkarni, "Comprehensive Survey on Detection of Anomalies in Edge Computing Network and Deep Learning Solutions:," in Proceedings of the 1st International Conference on Cognitive & Cloud Computing. Jaipur, India: SCITEPRESS - Science and Technology Publications, 2024, pp. 37–45. [Online]. Available: https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0013344100004646

[13] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp, and K. Reddy, "Recurrent Neural Networks for Accurate RSSI Indoor Localization," IEEE Internet of Things Journal, vol. 6, no. 6, pp. 10 639–10 651, Dec. 2019. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8830368

[14] M. S. Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, "Network Anomaly Detection Using LSTM Based Autoencoder," Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks, 2020. [Online]. Available: https://www.academia.edu/56770191/Network_Anomaly_Detection_Using_LSTM_Based_Autoencoder

[15] S. Narmadha and N. V. Balaji, "Improved network anomaly detection system using optimized autoencoder - LSTM," Expert Systems with Applications, vol. 273, p. 126854, May 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417425004762

[16] P. Bountzis, D. Kavallieros, T. Tsikrika, S. Vrochidis, and I. Kompatsiaris, "A deep one-class classifier for network anomaly detection using autoencoders and one-class support

vector machines," Frontiers in Computer Science, vol. 7, Oct. 2025. [Online]. Available: https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2025.1646679/full
[17] PPG Institute of Technology, India, B. V, C. P. A, Karpagam Institute of Technology, India, T. B, and Queensland University of Technology, Australia, "ENHANCED INTRUSION DETECTION AND PREVENTION IN WIRELESS SENSOR NETWORKS USING HYBRID DEEP LEARNING," ICTACT Journal on Communication Technology, vol. 16, no. 1, pp. 3454–3458, Mar. 2025. [Online]. Available: https://ictactjournals.in/ArticleDetails.aspx?id=wgudejpb
[18] V. K. Pandey, S. Prakash, T. K. Gupta, P. Sinha, T. Yang, R. S. Rathore, L. Wang, S. Tahir, and S. T. Bakhsh, "Enhancing intrusion detection in wireless sensor networks using a Tabu search based optimized random forest," Scientific Reports, vol. 15, no. 1, p. 18634, May 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-03498-3
[19] F. Qian, G.-m. Hu, and X.-m. Yao, "Semi-supervised internet network traffic classification using a Gaussian mixture model," AEU - International Journal of Electronics and Communications, vol. 62, no. 7, pp. 557–564, Aug. 2008. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1434841107001409
[20] A. Zahoor, W. Abbasi, M. Z. Babar, and A. Aljohani, "Robust IoT security using isolation forest and one class SVM algorithms," Scientific Reports, vol. 15, p. 36586, Oct. 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12540939/

A. HARSHA KUMAR is currently pursuing the Bachelor of Technology (B.Tech.) degree in Computer Science and Engineering from Vellore Institute of Technology, Chennai, India.

He has contributed to defense research at the Combat Vehicles Research and Development Establishment (CVRDE), DRDO, Chennai, focusing on distributed communication frameworks for in-vehicle networks using OpenDDS. Additionally, he has collaborated with Samsung R&D on the development of hybrid CNN–ViT watermarking systems for secure image authentication. His research portfolio includes publications on decentralized digital health systems and IoT-enabled maternal health monitoring. His research interests encompass blockchain applications, decentralized systems, machine learning, AI-enhanced healthcare, and the Internet of Things (IoT).

Mr. Kumar was a Grand Finalist at the IEEE YESIST12 2024 for his work on prenatal healthcare technologies.

S. KRITHICK is currently pursuing the Bachelor of Technology (B.Tech.) degree in Computer Science and Engineering at Vellore Institute of Technology, Chennai, India.

His research focuses on secure and intelligent cyber-physical systems, specifically the integration of embedded sensing with IoT networks to enable resilient decision-making. He has developed sensor-oriented step detection frameworks utilizing pressure and IR-based feedback, emphasizing on-device signal processing and lightweight estimation algorithms. He has further explored safety-critical connectivity solutions by architecting automotive monitoring systems that incorporate multi-parameter environmental sensing and real-time safety automation.

His research interests lie in cybersecurity for IoT ecosystems, embedded system intelligence, sensor fusion, and secure data communication mechanisms. He aims to architect scalable, secure real-world deployments of intelligent embedded systems that enhance autonomy and human well-being.

K. KUMARAN received the Ph.D. degree in Computer Science and Engineering from SRM Institute of Science and Technology, Chennai, India.

He is currently an Assistant Professor (Senior Grade) with the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Chennai, India. He has over 13 years of academic and research experience and has authored or co-authored more than 60 papers in national and international journals, conferences, and symposiums. His research interests include deep learning, cybersecurity, and edge computing.

G. SARANYA received the Ph.D. degree in Computer Science and Engineering from SRM Institute of Science and Technology, Chennai, India.

She is currently an Assistant Professor (Senior Grade) with the School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Chennai, India. She has over 12 years of academic and research experience and has published more than 50 papers in national and international journals, conferences, and symposiums. Her research interests include deep learning and edge computing.

ABISHEK DEVADOSS is currently pursuing the Bachelor of Technology (B.Tech.) degree in Computer Science and Engineering from Vellore Institute of Technology, Chennai, India.

He has worked extensively in embedded intelligence, applied machine learning, and cloud-integrated safety systems. His research investigates cloud-native automation and scalable deployment architectures, with a focus on designing reproducible infrastructure. He has developed reliability-centered provisioning frameworks utilizing distributed cloud services to support real-time IoT and machine learning experimentation.

His areas of interest include machine learning and artificial intelligence, computer vision, embedded and cyber-physical systems, cloud infrastructure engineering, and IoT security.

R. SACHEEV KRISHANU is currently pursuing the Bachelor of Technology (B.Tech.) degree in Computer Science and Engineering from Vellore Institute of Technology, Chennai, India.

His research activities span embedded intelligence systems, applied machine learning, and computational signal processing. He has conducted significant research on automated medical screening systems, contributing to the development of deep learning–based platforms for skin cancer detection.

His research interests include applied machine learning, embedded and cyber-physical systems, edge AI, secure data transmission frameworks, parallel computing, and intelligent healthcare technologies.

DISHA DANIEL is currently pursuing the Bachelor of Technology (B.Tech.) degree in Electronics and Communication Engineering from Vellore Institute of Technology, Chennai, India.

Her research focuses on the intersection of resource-constrained computing, signal processing, and secure decentralized systems. She has contributed to the development of lightweight security frameworks for edge devices and optimized small language models for disaster response applications. Her work also explores multimodal maternal health monitoring systems and privacy-preserving decentralized health record platforms. Her research contributions include publications in high-impact journals and conference proceedings such as *IEEE Access* and *Frontiers in Digital Health*. Her areas of interest include computer architecture, edge AI, biomedical signal processing, and blockchain-enabled security.

PREETHIKA RANGANATHAN is currently pursuing the Bachelor of Technology (B.Tech.) degree in Computer Science and Engineering from Vellore Institute of Technology, Chennai, India.

Her research interests include deep learning architectures, artificial intelligence, and quantum computing. She has investigated the comparative efficacy of lightweight architectures, such as MobileNetV3, versus traditional Convolutional Neural Networks (CNNs) in complex classification and segmentation tasks. Her work emphasizes the optimization of resource-efficient deep learning models to enhance performance in constraint-sensitive medical imaging applications.