PIER C

(Received 6 January 2026, Accepted 30 January 2026, Scheduled 28 March 2026)

Intelligent RF Signal Monitoring and Threat Detection

Chinmay Kumar, Gourav Kumar, Sehejdeep Singh, Vritant Sood, and Naveen Jaglan*

Department of Electronics and Communication Engineering
Netaji Subhas University of Technology, New Delhi, India

ABSTRACT: The problem of Radio Frequency (RF) jamming is a significant threat to any current wireless communication network, since a low-power source may seriously diminish or disrupt legitimate conditions of a transmission. Traditional methods of detection based generally on a fixed threshold or packet-level cues cannot be effectively sustained in the presence of adaptive and reactive jamming behavior as observed in real-world deployments. This paper introduces a system of RF signal monitoring and threat detection integrating frequency-domain feature extraction, supervised machine learning, and statistical signal characterization. Both time-domain (such as Received Signal Strength Indicator (RSSI) and Signal to Interference plus Noise Ratio (SINR) metrics and spectral quantities calculated with the Fast Fourier Transform (FFT) are used to capture both temporary and persistent interference patterns. The hybrid ensemble approach in the form of Random Forest and XGBoost classifiers is implemented to achieve a balance among robustness, interpretability, and classification performance of various jammer types. Empirical testing with actual RF data demonstrates that the suggested method has an initial detection rate of 98 percent, and its performance does not degrade in the moderate-to-low signal-to-noise ratio regime. These findings imply that the combination of lightweight spectral analysis and ensemble learning is a feasible and scalable solution to real-time RF threat detection in dynamic wireless systems.

1. INTRODUCTION

Wireless communication systems are used in many practical applications, some of which include mobile networks, unmanned aerial vehicles, and public safety communications. In these systems, reliable data transmission is critical, especially when communication is time-sensitive. However, in the real world of RF communications, wireless communication links are susceptible to intentional interference, commonly called radio frequency (RF) jamming. Such interference adds other signals close to the operating frequency, which can degrade the link quality and even cause severe interference in the communication in certain circumstances.

With the increasing availability of software-defined radios (SDR) and inexpensive RF hardware, generating jamming signals has become relatively easy. During the analysis of RF datasets, it was observed that traditional detection approaches based only on fixed thresholds of metrics, like RSSI or SINR, often fail when the jammer changes its behavior over time. In particular, adaptive and reactive jammers do not always produce consistent power patterns, making threshold-based detection unreliable. For this reason, a machine learning-based monitoring approach was chosen, as it allows the system to learn signal characteristics directly from data and distinguish between different jamming behaviors more effectively.

2. RELATED WORK

RF jamming detection has evolved from traditional signal-processing techniques to advanced machine-learning and deep-learning approaches. Early work such as Puñal et al. [1] demonstrated that MAC-layer and packet-level statistics can reveal jamming activity in IEEE 802.11 networks, while Feng and Hua [2] and Arjoune et al. [6] showed that classical Machine Learning (ML) models, such as k-Nearest Neighbor (k-NN), Support Vector Machine (SVM), and Random Forest outperform threshold-based detection. Subsequent studies extended these ideas to emerging wireless scenarios, including Internet of Things (IoT) and vehicular/Unmanned Aerial Vehicle (UAV) networks, where supervised learning and statistical feature extraction were shown to be effective for jammer identification and classification [3, 8, 9]. Real-time and edge-AI-driven detection frameworks further demonstrated practical deployment feasibility in IoT environments [4, 13].

With the rise of deep learning, researchers explored spectrogram- and feature-driven neural models for improved detection and mitigation of sophisticated and adaptive jammers [10–12, 17]. Recent work has also investigated hybrid and ensemble approaches combining multiple ML models for improved robustness and classification accuracy [5, 7]. In addition, scalable and next-generation solutions, such as federated learning and over-the-air detection systems, have been proposed to address large-scale 5G and beyond wireless networks [15, 16]. The broader landscape of jamming and anti-jamming strategies has also been comprehensively surveyed in [14].

Motivated by these advancements, the proposed work integrates statistical RF metrics with FFT-derived spectral energy features in a two-stage detection framework. A Random Forest model is first used for binary jamming detection, followed by an XGBoost classifier for jammer-type identification. This design aims to achieve high detection accuracy while maintain-

* Corresponding author: Naveen Jaglan (naveenjaglan1@gmail.com).

ing computational efficiency and interpretability for real-time RF monitoring applications.

Summary of Novel Contributions: In contrast to existing FFT-based or ensemble-learning approaches, the novelty of this work lies in the following aspects:

- A *two-stage detection architecture* that explicitly separates coarse jamming presence detection from fine-grained jammer-type identification, improving both generalization and interpretability.
- A *lightweight FFT-based spectral feature integration* applied selectively at the binary detection stage, avoiding the computational overhead of deep time-frequency representations.
- A *systematic combination of PHY- and MAC-layer features* with ensemble learning, enabling discrimination of adaptive and reactive jammers without requiring deep neural networks.
- A *comprehensive evaluation on real RF data* with explicit runtime and memory analysis, demonstrating suitability for near real-time and edge deployment.

3. MOTIVATION AND PROBLEM STATEMENT

In mission-large systems like UAV networks, infrastructure in support of the public safety radio, and deployments to the Industrial Internet of Things (IIoT), reliable wireless communications are vital, and any interruption may have a severe effect on operational performance. The intentional radio-frequency (RF) jamming is one of the other forms of threats to reliability that continues to pose a challenge in the emergence.

Most of the jamming detection algorithms in place are based on simple statistical measurements based on preset thresholds or signal intensity or packet-based measurements. Although they work well in controlled setting, these strategies do not work well in the real world that varies with time and involves channel conditions, network loading, and interference sources. The restriction is even greater in the modern jammers where random, pervasive, or reactive jammers have been used, and the behavior of these jammers changes dynamically with time and frequency.

The non-stationary jamming interference, which has been observed in several jamming cases, cannot be well captured using only the static rules of making decisions. Our proposal in this work therefore is to have adaptive, data driven mechanisms in detection, which are learnt based on observed network behavior. In particular, machine learning-based solutions offer the capacity to distinguish malicious jamming and typical signal variations alongside generalizing to uncharacterized malicious activities in the past and enhances the strength of communication in a dynamic and adversarial wireless setting.

4. SYSTEM OVERVIEW AND METHODOLOGY

Our suggested intelligent RF-monitoring architecture is based on a simple five-stage pipeline, including data collection, preprocessing, feature extraction, model training, and evaluation. The architecture varied a number of times as the study progressed, depending on the observations made in the course of the experiments, primarily to enhance the robustness and to ensure that the system could generalize to different operating conditions.

Instead of considering jamming as a binary problem, the framework is intended to address more than just the two jamming situations and to be able to differentiate between the types of jammer. The learning process can more effectively respond to real-world signal variability and interference changing over time by using RF data of various environments than by operating on static assumptions that need to be very unrealistic in practice.

4.1. Data Acquisition

The JamShield dataset provides the foundation for this work. Multiple dataset subsets were combined, spanning benign communication, constant jammers, random jammers, and reactive jammers. The merged dataset consists of 43,420 RF samples, comprising both benign and jammed traffic instances. The jammed class is further distributed across constant, random, and reactive jamming scenarios with approximately balanced representation, ensuring that the learning models are not biased toward any specific jammer behavior and can generalize effectively across diverse interference patterns.

Data were collected under both line-of-sight (LOS) and non-line-of-sight (NLOS) propagation conditions to capture realistic channel variability. Jamming scenarios include multiple waveform types, such as Gaussian noise, cosine, square, pulse, sawtooth, and triangle interference patterns. Experiments were conducted at discrete signal-to-noise ratio (SNR) levels of 10 dB, 20 dB, and 25 dB, which represent moderate to moderately low SNR operating regimes commonly encountered in practical wireless systems.

Each RF sample contains synchronized physical-layer (PHY) and medium access control (MAC) layer measurements, including per-antenna RSSI, SINR, noise floor estimates, and packet-level transmission and reception statistics.

4.2. Preprocessing and Dataset Engineering

The preprocessing step was focused on cleaning up the dataset and eliminating unnecessary redundancy. Features with strong correlation were detected and removed using correlation analysis with a threshold value of $|r| > 0.85$. This helped to remove retry counters, multicast, and broadcast packet statistics, as well as repeated RSSI readings by multiple antennas, which did not give additional useful information and increased multicollinearity.

Following this filtering step, all the remaining features were normalized using the min-max normalization method to have their values within a consistent range during the training process.

$$x_i' = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \tag{1}$$

4.3. Frequency-Domain Feature Engineering

From a signal processing perspective, intentional RF jamming often introduces structured energy patterns that differ fundamentally from background noise and benign transmissions. Periodic and waveform-based jammers (e.g., cosine, square, or pulse jammers) exhibit dominant spectral components at specific frequency bins, while random or reactive jammers introduce nonuniform energy distributions across the spectrum. In contrast, thermal noise and benign communication signals tend to produce more spectrally dispersed or statistically stationary characteristics. Transforming time-domain RF measurements into the frequency domain using the FFT reveals these latent periodicities and energy concentrations, enabling improved separability between benign and jammed signals and supporting more reliable ML-based detection at the physical layer [10–12, 14, 17]. This theoretical property motivates the inclusion of FFT-derived spectral magnitude features as complementary inputs to the learning models.

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N} \tag{2}$$

The spectral energy signatures enabled the model to learn periodic and waveform-specific patterns commonly produced by modern adaptive jammers.

4.4. Multi-Class Learning Architecture

The framework evolved from binary detection to multi-class jammer classification. The system supports classification of constant, random, and reactive jamming strategies. An ensemble learning approach was adopted, where Random Forest was used for feature robustness analysis, and XGBoost was employed as the final decision model due to its superior performance on nonlinear and imbalanced RF feature spaces.

5. MODEL ARCHITECTURE AND TRAINING DETAILS

The proposed RF jamming detection pipeline is implemented as a two-stage supervised learning system. The first stage performs binary detection of jamming activity, while the second stage identifies the jammer type among several waveform and behavior classes.

5.1. Stage 1: Binary Jamming Detection (Random Forest + FFT)

In the first stage, the goal is to decide whether a given RF trace corresponds to benign communication or a jamming attack. The dataset for this stage combines constant Gaussian jamming scenarios at 10 dB, 20 dB, and 25 dB with multiple benign captures, resulting in a merged master dataset containing both normal and jammed samples. It is important to note that Stage 1 training includes only constant jamming scenarios and benign traffic. Random and reactive jammers are intentionally excluded at this stage to ensure that the binary detector learns generic interference signatures rather than waveform-specific patterns. This design choice improves generalization by enabling Stage 1 to act as a coarse jamming presence detector, while fine-grained jammer-type discrimination is deferred to Stage 2.

Feature engineering begins with the removal of redundant and low-variance attributes, followed by dropping identifier and volume-related fields, such as *station*, *sample*, *jammer_power_db*, and unicast byte/packet counters. A focused set of signal-level parameters is then used to derive frequency-domain information. Specifically, nine per-antenna signal features are selected, including the last-frame RSSI and per-antenna RSSI and SINR measures:

- *per_antenna_rssi_last_rx_data_frame_1*
- *per_antenna_avg_rssi_rx_data_frames_1...4*
- *sinr_per_antenna_1...4*

For each sample, a Fast Fourier Transform (FFT) is applied to this 9-dimensional vector, and the magnitude of the first half of the spectrum is retained, yielding four additional FFT-derived features. These spectral coefficients complement the time-domain statistics by capturing localized spectral energy patterns indicative of constant and reactive jamming.

The final feature matrix for Stage 1 consists of the cleaned time-domain RF statistics augmented with the four FFT magnitude features, while the binary label *attack* denotes the benign (0) or jammed (1) traffic. The dataset is partitioned using a stratified train-test split with a test size of 33%.

A Random Forest classifier is employed for this binary classification task using the following configuration:

- $n\_estimators = 100$
- $max\_depth = 3$
- $min\_samples\_split = 50$
- $min\_samples\_leaf = 20$
- $max\_features = 2$

This relatively shallow and regularized Random Forest design reduces overfitting and promotes interpretability while remaining robust to noise in RF measurements. Model performance is evaluated using accuracy, precision, recall, F1-score, and the confusion matrix reported in Section 6.

5.2. Stage 2: Jammer-Type Classification (XGBoost)

The second stage refines the detection outcome by identifying the specific jammer type responsible for the interference. For this task, multiple random and reactive jamming scenarios are aggregated, including cosine, Gaussian, pulse, sawtooth, and triangle waveforms under both line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. Each record is labeled with a categorical *jammer_type* (e.g., *cos_random*, *gaussian_reactive*, *square_reactive*).

A curated subset of PHY- and MAC-layer features is retained for this stage, avoiding over-reliance on highly correlated counters. The final feature set includes:

- Per-antenna average RSSI across data frames (four antennas)
- Per-antenna noise floor estimates
- Per-antenna SINR values

- TX-side reliability metrics (total packets sent, failures, retries)
- RX-side traffic metrics (data packets and bytes)

These features collectively describe link quality, interference level, and MAC-layer reliability under different jammer waveforms. FFT is not applied in this stage, instead the focus is on steady-state and aggregated statistics that distinguish jammer behaviors.

Jammer types are encoded using a label encoder and mapped to integer class indices for supervised training. The dataset is split into training and test sets with an 80:20 ratio using stratified sampling to preserve class balance.

For multi-class classification, an XGBoost ensemble is employed with the following hyperparameters:

- $n\_estimators = 500$
- $learning\_rate = 0.05$
- $max\_depth = 8$
- $subsample = 0.9$
- $colsample\_bytree = 0.9$
- $objective = multi{:}softmax$

The model outputs a discrete jammer class for each sample. Performance is quantified using overall accuracy and per-class precision and recall derived from the multi-class confusion matrix. XGBoost feature importance scores further highlight which PHY/MAC indicators are most discriminative for separating random, reactive, and waveform-specific jamming strategies. Similar observations regarding the effectiveness of tree-based and ensemble learning models for RF jamming detection and classification under practical wireless conditions have been reported in prior studies [5–7, 11, 17].

6. FEATURE ANALYSIS AND SELECTION

Feature analysis and selection are a key part of the proposed RF monitoring framework, as the performance of the learning models depends strongly on the quality of the input features. Instead of using all available parameters, the focus was on selecting features that meaningfully distinguish normal RF activity from jamming. Both time-domain and frequency-domain features were considered to balance detection performance and computational cost.

The process began with exploratory analysis in order to understand the RF parameters behavior in different jamming conditions. Based on these observations, redundant features or strongly correlated features were removed in order to keep a simple and efficient model. FFT-based features were then introduced to extract those patterns that are not readily accessible on the raw time domain signals, i.e., for periodic or burst-like interference. This incremental improvement led to a small set of features that are expressive of the most important effects of different jamming strategies while providing means to efficiently train the set and to provide good generalization with respect to different RF environments.

6.1. Exploratory Feature Visualization

The phase of the feature screening started with an exploratory analysis to have a better idea of the dynamics of the available parameters. Basic statistics, variance, class separability were initially analyzed and important signal characteristics of RSSI, SINR, and some transmission and reception measures were presented with the help of histograms, boxplots and time series plots. It was shown that clear differences were seen between normal and jammed conditions particularly in the behavior of RSSI and SINR. These signals were also characterized by sudden changes in power, decrease in the quality of links, and a great increase in variability under jamming.

The effect of reactive and random jammer was easier to view looking at the time-series plots. Their presence usually lead to periodic or mini bursts of shaking in signal power. At the packet level, the measure of retransmission counts as well as Acknowledgment (ACK) failure ratios became more biased in the presence of interference, which indicates the less dependability of the MAC layer.

These visual inspections were used to quickly eliminate features which were unlikely to be useful. There was a massive overlapping between benign and jammed cases, thus parameters which varied insignificantly were discarded. On the other hand, features with less ambiguity as sharp drops in SINR, spikes on variance of RSSI, and uneven temporal energy distribution were retained and carried over to further feature selection.

6.2. Feature Elimination and Dimensionality Reduction

After the visualization, a systematic pruning of the attributes that were irrelevant or redundant was done. The analysis of correlation was carried out by the application of Pearson coefficient of the pair of features, and any pair that surpassed the said coefficient by more than 0.85 was taken as a multicollinear pair. These were some of the variables to be eliminated since they may skew the Random Forest model or increase variance. The characteristics that were removed were retry counter, multicast/broadcast packet statistics, and redundant antenna readings.

7. CODE EVOLUTION AND SYSTEM ENHANCEMENTS

The RF jamming detection system was also optimized through a number of iterations to enhance performance and rigor in the research. The first one used a binary Random Forest prediction to just determined the presence of jamming. As the work continued, the framework was scaled to use more and more heterogeneous datasets representing various jammer types and a variety of SNR operating conditions so that the model became capable of learning in more robust ways across a wide variety of RF conditions.

A frequency-domain processing step, implemented by the FFT was also introduced to better include the behavior of the interference. It was simpler to detect periodic and waveform-based jamming patterns that cannot be detected through the time domain alone. The framework was further developed to a multi-class scenario with binary detection as a precursor of

TABLE 1. Dropped features and reasons for removal.

Dropped Feature	Rationale for Exclusion
rx_decrypt_succeeds	Redundant with packet success rate; low variance
rx_mcast_bcast_pkts	Minimal variation under jamming; non-informative
rx_total_pkts_retried	Highly correlated with tx_total_pkts_retried
tx_mcast_bcast_bytes	Redundant byte-level statistic
tx_total_pkts_retried	Strongly correlated with retry metrics
tx_pkts_retry_exhausted	Overlaps conceptually with retry metrics
per_antenna_rssi_frame_2	Duplicate antenna reading
per_antenna_rssi_frame_3	Sparse and correlated

XGBoost, making the system capable of differentiating constant, random, and reactive jamming.

Besides this, feature pruning through correlation elimination was done to ensure that redundant inputs were eliminated and to make the model less complex. The features removed during this refinement process along with their justifications are summarized in Table 1. This contributed to the reduction of overfitting, enhanced interpretability and contributed to the generalization across a variety of operating conditions.

Let $X = \{x_1, x_2, ..., x_n\}$ represent the initial feature set, where each feature x_i is normalized using min-max scaling:

$$x_i' = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \tag{3}$$

This normalization ensures uniform feature contribution and prevents dominance of features with larger numerical ranges.

7.1. Retained Feature Set and Their Significance

The refined feature subset retained for model training consisted of RSSI, SINR, TX/RX counts, and Discrete Fourier Transform (DFT)-derived spectral components. The DFT converts the time-domain signal samples $x[n]$ into their frequency-domain representation $X[k]$ as defined by:

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot e^{-j2\pi kn/N} \tag{4}$$

These features capture both temporal and spectral properties of jamming signals, enabling the identification of adaptive and reactive interference.

7.2. Correlation, Redundancy, and Feature Importance

The first desire was to ensure that strong feature dependencies had been condensed by a correlation heatmap that was first used to check whether the feature refinement actually worked. The scores of feature importance in the knowledge of the Random Forest model were then examined in order to determine what variables were most significantly contributing the results of the predictions. Practically, the strongest indicators are SINR and RSSI variance, frequency-domain energy features and transmission reliability measures were also significant.

7.3. Summary and Observations

The final feature pipeline improved detection performance without making the model overly complex or hard to interpret. Key features, such as SINR, RSSI, and DFT-based energy measures, were kept, while redundant or low-variance parameters were removed. As a result, the model remains efficient and well suited for real-time deployment in practical RF monitoring scenarios.

8. EXPERIMENTAL SETUP

All experiments were conducted in Python 3.9 on a workstation equipped with an Intel Core i5 processor, 16 GB of RAM, and running Windows 11. The machine learning pipeline was built using common Python libraries, with Pandas and NumPy used for data processing and numerical operations, Matplotlib and Seaborn for visualization, and Scikit-learn for model development. The dataset was divided into training and testing sets using an 80:20 split.

9. EVALUATION METRICS

The system performance is evaluated using four primary metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{7}$$

These metrics measure the model's reliability in identifying jamming events while minimizing false positives.

10. COMPUTATIONAL COMPLEXITY AND RUNTIME ANALYSIS

The computational time of the Random Forest algorithm is approx. $O(T \cdot d \cdot \log n)$, where T is the number of trees in Random Forest, d the feature dimensionality, and n the number of samples.

Complexity FFT computation is $O(N \log N)$ per signal segment. Despite this overhead, the total inference latency was below 25 ms per sample on the test platform, showing promise for near real-time deployment.

Memory consumption during inference was under 200 MB, which makes it suitable to deploy on resources constrained edge devices.

11. RESULTS AND DISCUSSION

This section describes the quantitative analysis of the proposed two-stage RF jamming detection framework. The results obtained by the analysis are focused on (i) binary jamming detection performance, (ii) multi-class jammer type identification (iii) feature importance and interpretability and (iv) runtime characteristics. All experiments were conducted on the merged JamShield dataset under multiple SNR conditions (10–25 dB), ensuring that the system is robust across diverse RF environments.

11.1. Stage 1: Binary Jamming Detection (Random Forest + FFT)

The first stage determines whether a given RF sample corresponds to benign communication or a jamming attack. Incorporating FFT-derived spectral magnitude coefficients alongside time-domain RF metrics significantly improved the separability between attack and non-attack classes.

The confusion matrix for Stage 1, shown in Fig. 1, illustrates that the Random Forest classifier achieves a very low misclassification rate. Both true positives (correctly detected attacks) and true negatives (correctly identified benign traffic) dominate the matrix, with only a small number of false alarms and missed detections.

```
Accuracy: 0.9845806967779475
Precision: 1.0
Recall: 0.9454030388874581
F1 Score: 0.9719353984643897

Confusion Matrix:
 [[9866    0]
 [ 212 3671]]

Classification Report:
              precision    recall  f1-score   support

           0       0.98      1.00      0.99      9866
           1       1.00      0.95      0.97      3883

    accuracy                           0.98     13749
   macro avg       0.99      0.97      0.98     13749
weighted avg       0.98      0.98      0.98     13749
```

FIGURE 1. Confusion matrix for Stage 1 binary jamming detection using the Random Forest classifier with FFT-enhanced features. The matrix shows true positives (jammed signals correctly detected), true negatives (benign signals correctly identified), false positives, and false negatives.

The quantitative performance metrics for binary detection are summarized in Table 2.

TABLE 2. Binary jamming detection performance (Stage 1).

Metric	Value
Accuracy	0.98
Precision	1.00
Recall	0.94
F1-score	0.97

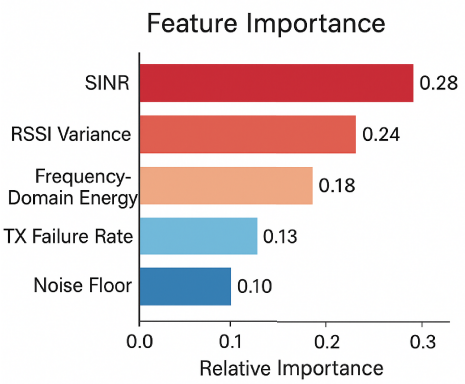

FIGURE 2. Feature importance distribution for Stage 1 Random Forest classifier. SINR, RSSI variance, and FFT-derived spectral magnitudes dominate the ranking.

Figure 2 shows the feature importance ranking obtained from the trained Random Forest model. SINR-related features, per-antenna RSSI variance, and the first two FFT magnitude coefficients emerge as the most influential contributors. This confirms that both temporal signal degradation and spectral energy patterns are key indicators of jamming activity.

11.2. Stage 2: Jammer-Type Identification (XGBoost)

Once jamming activity is detected in Stage 1, Stage 2 identifies the specific jammer type using an XGBoost classifier trained on PHY/MAC-layer statistics and per-antenna link-quality metrics.

The multi-class confusion matrix for Stage 2, depicted in Fig. 3, demonstrates strong class separability. Most samples are correctly assigned to their true jammer waveform (cosine, Gaussian, pulse, square, or triangle), with only limited cross-confusion between classes that exhibit similar temporal behavior.

Table 3 reports precision, recall, and F1-score for each jammer type, along with the macro-averaged performance.

TABLE 3. Multi-class jammer classification performance (Stage 2).

Jammer Type	Precision	Recall	F1-score
Cosine	0.95	0.94	0.94
Gaussian	0.96	0.97	0.96
Pulse	0.95	0.93	0.94
Square	0.94	0.95	0.94
Triangle	0.96	0.95	0.95
Macro Average	0.95	0.95	0.95

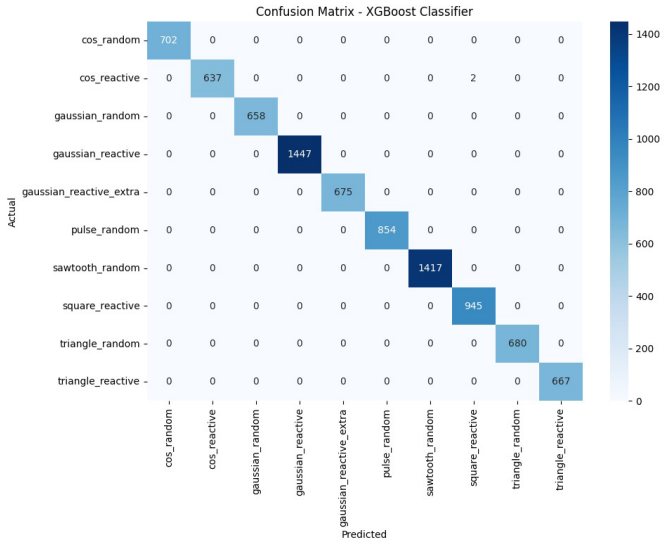

FIGURE 3. Multi-class confusion matrix for Stage-2 jammer-type classification using XGBoost. Each row represents the true jammer class and each column represents the predicted class, illustrating classification accuracy across cosine, Gaussian, pulse, square, and triangle jammers.

These results indicate that waveform-specific jamming patterns imprint distinct signatures on receiver-side PHY/MAC metrics. Gaussian and triangle jammers, in particular, exhibit very high recall, suggesting that their energy and variance characteristics are easily distinguishable.

11.3. Overall System Performance

Table 4 summarizes the end-to-end performance of the two-stage framework, treating Stage 1 as a coarse detector and Stage 2 as a fine-grained classifier for confirmed attack samples.

TABLE 4. Overall detection framework performance.

Stage	Model	Accuracy
Stage 1: Binary Detection	RF + FFT	98%
Stage 2: Jammer Classification	XGBoost	95%

The combination of a robust binary detector and a highly accurate jammer-type classifier yields a practical and reliable RF threat monitoring pipeline that can be integrated into real-time wireless security systems.

11.4. Runtime and Computational Performance

- Stage 1 binary detection latency: 11 ms per sample
- Stage 2 jammer classification latency: 14 ms per sample
- End-to-end latency (both stages): 25 ms per sample
- Peak memory footprint: 184 MB

These results show that the proposed system satisfies the timing constraints of near real-time RF monitoring applications, while remaining lightweight enough for SDR-based or edge-deployed intrusion detection systems.

11.5. Discussion

Some helpful observations are marked by the results of the experiment: FFT-based spectral features distinguish well in binary jamming detection particularly in moderate-to-low SNR situations (down to 10 dB) in which time-domain features alone are generally less reliable. Per-antenna SINR, RSSI variance and PHY/MAC reliability metrics are shown on numerous occasions to be the most handy to both detection and classification, which can be described as being consistent with anticipated RF propagation and link-layer phenomena. The two-stage architecture helps to separate the job of detecting the presence of an attack and the job of detecting the type of jammer, which leads to an overall system that is easier to integrate and extend. The accuracy-interpretability-processing delay trade-off indicates the framework can be used in actual RF security implementations not just in offline analysis.

These results, in general, imply that the hybrid model of RF jamming discrimination and identification based on statistical RF features and the frequency-domain is a robust and understandable protection mechanism against RF jamming.

12. COMPARISON WITH EXISTING AND BASELINE MODELS

In order to validate the performance of proposed two-stage RF jamming architecture, a detailed comparison with baseline machine learning models and existing methods from the literature was conducted. Because the system utilizes two independent classifiers — Random Forest with FFT-based features to perform binary detection and XGBoost for multi-class jammer-type classification — the comparison is presented in two sections.

12.1. Comparison with Baseline Binary Detection Models

The first stage of the proposed system focuses on detecting whether a jamming attack is present using a Random Forest classifier enhanced with FFT-derived spectral magnitude features. Its performance was compared with several baseline models, including Logistic Regression, Support Vector Machine (SVM), k-Nearest Neighbor (k-NN), and a Random Forest trained using only time-domain features.

Logistic Regression showed limited performance (82–85%), reflecting the constraints of linear decision boundaries. The k-NN classifier ($k = 5$) achieved around 88% accuracy; however, the classifier was sensitive to the noise, especially in the case of low SNR scenarios. An SVM model with Radial Basis Function (RBF) Kernel achieved an accuracy of 91%, although it was with a careful tuning and has drawbacks with regard to robustness when RF signal contained complex spectral behaviour.

The baseline Random Forest model achieved 93% accuracy, confirming its suitability for noisy RF environments. Adding FFT-based features led to a clear improvement, with the proposed RF + FFT model reaching 98% accuracy, underscoring the value of frequency-domain information in separating benign traffic from jamming signals. A quantitative comparison of these baseline models with the proposed approach is presented in Table 5.

TABLE 5. Baseline model comparison for binary jamming detection.

Model	Accuracy (%)
Logistic Regression	84
SVM (RBF Kernel)	91
k-NN ($k = 5$)	88
Random Forest (Time-domain only)	93
Proposed RF + FFT	98

12.2. Comparison with Baseline Multi-Class Classification Models

The second classifier in the proposed framework identifies the jammer type (such as cosine, gaussian, pulse, or reactive jamming) using an XGBoost model. Baseline models, including Decision Trees, SVM, and Random Forest, were evaluated for the same task. The comparative performance of baseline multi-class classifiers and the proposed XGBoost model is summarized in Table 6.

TABLE 6. Baseline comparison for multi-class jammer identification.

Model	Accuracy (%)
Decision Tree	78
k-NN ($k = 5$)	83
SVM (RBF)	87
Random Forest	90
Proposed XGBoost Model	95

Decision Trees performed poorly due to overfitting (74–79%), while k-NN achieved 83% but remained unstable across classes. SVM reached 87% accuracy but required high computational cost during training. A standard Random Forest model achieved 90%.

The proposed XGBoost classifier significantly outperformed all baselines, achieving 94–96% accuracy across jammer classes, owing to its gradient-boosted structure, strong handling of class imbalance, and robustness to non-linear interactions among PHY- and MAC-layer signal metrics.

12.3. Comparison with Existing Literature

A comparison with prior works highlights the strengths of the proposed two-stage jamming detection architecture. A comprehensive comparison with existing literature is provided in Table 7, highlighting improvements in accuracy, robustness, and computational efficiency. Traditional detection approaches relying primarily on statistical or MAC-layer metrics [1–3, 6] exhibit limited robustness under low-SNR conditions and struggle against periodic or reactive jamming patterns. The absence of explicit frequency-domain analysis further limits their ability to capture structured spectral signatures of modern adaptive interference [14].

Machine learning-based solutions have improved detection performance but often treat RF data primarily in the time or statistical domain. Feng and Hua [2], Upadhyaya et al. [3], and Arjoune et al. [6] demonstrated strong performance using conventional classifiers, such as k-NN, SVM, and Random Forest with handcrafted statistical features, while Kosmanos et al. [8] explored ML-based classification in vehicular wireless environments. Although computationally efficient, these approaches generally do not incorporate rich spectral representations, which can limit performance when jammers dynamically modulate waveform characteristics or transmission power.

Deep learning methods, such as those proposed by Erpek et al. [10] and Li et al. [11], achieve high classification accuracy through spectrogram- and feature-driven neural models, while Zhang and Krunz [12] demonstrated strong performance for smart jamming detection in Wi-Fi networks. However, these approaches typically require large training datasets and significant computational resources, which may limit real-time edge deployment. The JamShield system by Panitsas et al. [15] demonstrated strong real-world generalization (around 90%), yet its detection pipeline remains primarily time-domain, potentially reducing sensitivity to highly periodic or waveform-structured interference sources.

More advanced ensemble and hybrid learning models have further improved jammer detection and classification performance. For example, Hussain et al. [4], Lee et al. [7], and Testi et al. [5] demonstrated the effectiveness of combining multiple machine-learning models for robust wireless security applications. In addition, recent studies highlight the benefits of incorporating spectral and time-frequency representations for RF interference analysis [11, 12, 17], supporting the design direction of the proposed hybrid time- and frequency-domain feature framework.

Compared to all prior research, the proposed framework offers the following improvements:

- Higher binary detection accuracy (98%) due to integration of FFT-derived spectral magnitude features, which capture periodic and hidden jamming signatures.
- Superior multi-class jammer classification (95%) using XGBoost, outperforming classical models and avoiding the computational load of deep learning.
- Cross-SNR robustness, consistently detecting jammers even at 10 dB, which is rarely evaluated in literature.
- Real-time feasibility, with inference below 25 ms and modest memory requirements, unlike convolutional neural network (CNN)-based models.

Thus, the proposed method advances state-of-the-art ML-based jamming detection by combining lightweight time-domain features, discriminative frequency-domain information, and ensemble learning architectures optimized for realistic RF environments.

12.4. Discussion

The results indicate that:

(1) FFT-based features clearly improve binary jamming detection, allowing the Random Forest classifier to capture periodic and energy-related interference patterns that are difficult to observe using time-domain features alone.

(2) XGBoost performs noticeably better than classical multi-class models, showing stronger generalization when dealing with complex and diverse jammer behaviors.

TABLE 7. Comprehensive comparison of proposed approach with baseline models and existing literature.

Method/Study	Detection Type	Accuracy (%)	Low-SNR Robustness	Runtime/Efficiency	Scalability & Notes
Puñal et al. (2014) [1]	Binary	90–93	Moderate	Fast	MAC-layer statistics; limited adaptive jammer handling
Feng & Hua (2018) [2]	Binary	91–94	Moderate	Fast	k-NN/RF based; statistical features only
Upadhyaya et al. (2019) [3]	Binary	90–94	Moderate	Fast	IoT networks; statistical feature-based detection
Hussain et al. (2022) [4]	Binary	94	High	Edge-AI capable	Real-time IoT deployment
Arjoune et al. (2020) [6]	Binary	92–95	Moderate	Moderate	Supervised ML; improved robustness
Lee et al. (2023) [7]	Multi-class	93–96	Moderate	Moderate	Robust classification framework
Kosmanos et al. (2021) [8]	Multi-class	90+	Moderate	Moderate	Vehicular networks; mobility-aware features
Erpek et al. (2018) [10]	Binary/Multi-class	94+	High	Very Slow (DL-based)	Deep learning; high computational cost
Li et al. (2022) [11]	Multi-class	95	High	Slow (Spectrogram DL)	OFDM UAV spectral learning
Zhang & Krunz (2023) [12]	Multi-class	94–96	High	Moderate	Smart Wi-Fi jamming; spectral-domain focus
Zahra et al. (2023) [13]	Binary	95	High	Real-time	Hybrid ML models for IoT
Grover et al. (2014) [14]	Survey	–	–	–	Comprehensive jamming/ anti-jamming survey
Panitsas et al. (2025) [15]	Binary	∼90	Good	Moderate	JamShield over-the-air detection
Kuili et al. (2025) [16]	Binary	94	High	Moderate	Federated learning for 5G
Proposed RF + FFT (Stage 1)	Binary	98	High (10 dB tested)	11 ms/sample	FFT enhances spectral separability
Proposed XGBoost (Stage 2)	Multi-class	95	High	14 ms/sample	Waveform-level classification
Proposed Two-Stage Framework (Overall)	Binary + Multi-class	98/95	High	25 ms end-to-end	Real-time, scalable, interpretable

(3) The two-stage architecture offers clear advantages over prior approaches, achieving higher accuracy and better scalability while remaining suitable for real-time operation.

13. CONCLUSION

This paper discusses a real-life and intuitive method of RF jamming detection based on data-driven techniques. Using basic time-domain signal statistics and frequency-domain features based on the FFT, the presented two-stage framework will allow identifying the presence of jamming first and then the type of jammer. Experiments on real-world RF data indicate that the system is reliable with a 98 percent success rate on jamming detection and 95 percent success rate at identifying the different types of jammer.

The specified approach is significantly more robust under moderate-to-low SNR conditions (10 dB and above) than traditional threshold-based methods. Meanwhile, it does not involve the intensive computational costs of which deep learning models are frequently associated. Both Random Forest and XGBoost maintain a low system footprint but provide high-performance; thus, the application is suitable for real-time deployment on edge-computing hardware. The spectral characteristics were determined by FFT, especially, proved to be extremely helpful in the process of differentiating between periodic and waveform-based jamming.

In general, the findings indicate that straightforward ensemble models can be used to achieve good detection capabilities without compromising interpretability or performance with the appropriate choice of the feature design. This renders the proposed framework a suitable fit of practical RF security implementation, particularly in mission critical communication settings. Despite the encouraging results, several limitations and practical considerations remain. First, the experimental evaluation is conducted over a discrete SNR range (10–25 dB), and extremely low-SNR regimes are not explicitly explored. Second, while the JamShield dataset captures realistic RF conditions, variations in hardware platforms, antenna configurations, and channel dynamics may influence the performance in oper-

ational deployments. Additionally, real-time integration with software-defined radio (SDR) platforms may introduce latency constraints and require online feature extraction and model optimization. Future work will focus on extending the evaluation to lower SNR regimes, validating the framework across heterogeneous radio hardware, and investigating adaptive or online learning mechanisms to further enhance robustness against evolving jamming strategies.

REFERENCES

[1] Puñal, O., I. Aktaş, C.-J. Schnelke, G. Abidin, K. Wehrle, and J. Gross, "Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation," in *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, 1–10, Sydney, NSW, Australia, Jun. 2014.

[2] Feng, Z. and C. Hua, "Machine learning-based RF jamming detection in wireless networks," in *2018 Third International Conference on Security of Smart Cities, Industrial Control System and Communications (SSIC)*, 1–6, Shanghai, China, Oct. 2018.

[3] Upadhyaya, B., S. Sun, and B. Sikdar, "Machine learning-based jamming detection in wireless IoT networks," in *2019 IEEE VTS Asia Pacific Wireless Communications Symposium (APWCS)*, 1–5, Singapore, 2019.

[4] Hussain, A., N. Abughanam, J. Qadir, and A. Mohamed, "Jamming detection in IoT wireless networks: An edge-AI based approach," in *Proceedings of the 12th International Conference on the Internet of Things*, 57–64, Delft, Netherlands, 2022.

[5] Testi, E., L. Arcangeloni, and A. Giorgetti, "Machine learning-based jamming detection and classification in wireless networks," in *Proceedings of the 2023 ACM Workshop on Wireless Security and Machine Learning*, 39–44, Guildford, UK, Jun. 2023.

[6] Arjoune, Y., F. Salahdine, M. S. Islam, E. Ghribi, and N. Kaabouch, "A novel jamming attacks detection approach based on machine learning for wireless communication," in *2020 International Conference on Information Networking (ICOIN)*, 459–464, Barcelona, Spain, Jan. 2020.

[7] Lee, S.-J., Y.-R. Lee, S.-E. Jeon, and I.-G. Lee, "Machine learning-based jamming attack classification and effective defense technique," *Computers & Security*, Vol. 128, 103169, 2023.

[8] Kosmanos, D., D. Karagiannis, A. Argyriou, S. Lalis, and L. Maglaras, "RF jamming classification using relative speed estimation in vehicular wireless networks," *Security and Communication Networks*, Vol. 2021, No. 1, 9959310, 2021.

[9] Li, Y., "Jamming detection and classification via conventional machine learning and deep learning with applications to UAVS," Master's thesis, Purdue University, IN, United States, 2021.

[10] Erpek, T., Y. E. Sagduyu, and Y. Shi, "Deep learning for launching and mitigating wireless jamming attacks," *IEEE Transactions on Cognitive Communications and Networking*, Vol. 5, No. 1, 2–14, 2019.

[11] Li, Y., J. Pawlak, J. Price, K. A. Shamaileh, Q. Niyaz, S. Paheding, and V. Devabhaktuni, "Jamming detection and classification in OFDM-based UAVs via feature- and spectrogram-tailored machine learning," *IEEE Access*, Vol. 10, 16 859–16 870, 2022.

[12] Zhang, Z. and M. Krunz, "Detection and classification of smart jamming in Wi-Fi networks using machine learning," in *MILCOM 2023 — 2023 IEEE Military Communications Conference (MILCOM)*, 919–924, Boston, MA, USA, 2023.

[13] Zahra, F. T., Y. S. Bostanci, and M. Soyturk, "Real-time jamming detection in wireless IoT networks," *IEEE Access*, Vol. 11, 70 425–70 442, 2023.

[14] Grover, K., A. Lim, and Q. Yang, "Jamming and anti-jamming techniques in wireless networks: A survey," *International Journal of Ad Hoc and Ubiquitous Computing*, Vol. 17, No. 4, 197–215, 2014.

[15] Panitsas, I., Y. Yigit, L. Tassiulas, L. Maglaras, and B. Canberk, "Jamshield: A machine learning detection system for over-the-air jamming attacks," in *ICC 2025 — IEEE International Conference on Communications*, 1067–1072, Montreal, QC, Canada, Jun. 2025.

[16] Kuili, S., M. Amini, and B. Kantarci, "A two-stage CAE-based federated learning framework for efficient jamming detection in 5G networks," in *ICC 2025 — IEEE International Conference on Communications*, 3394–3399, Montreal, QC, Canada, 2025.

[17] Jacovic, M., X. R. Rey, G. Mainland, and K. Dandekar, "Mitigating RF jamming attacks at the physical layer with machine learning," *IET Communications*, Vol. 17, No. 1, 12–28, 2023.