Finding a Needle in a (Spectrum) Haystack: Multi-Band Multi-Device Radio Fingerprinting

Ildi Alla[a], Milin Zhang[b], Jonathan Ashdown[c], Valeria Loscri[a], Francesco Restuccia[b]

[a] *Inria Lille-Nord Europe, France*
[b] *Institute for the Wireless Internet of Things at Northeastern University, United States*
[c] *Air Force Research Laboratory, United States*

Abstract

As the spectrum becomes increasingly crowded, quick and reliable authentication of wireless devices is critical to avoid harmful interference to incumbents of the spectrum. *Radio fingerprinting* achieves fast waveform-level authentication by distinguishing devices based on unique hardware imperfections in the radio circuitry. However, existing approaches can fingerprint only one signal in a specific band, making them inapplicable in real-world scenarios where multiple signals coexist in spectrum bands. This paper introduces *Multi-band Multi-device Radio Fingerprinting* (`M2RF`) to address this challenge. Specifically, we propose a learning-driven segmentation algorithm to directly process in-phase/quadrature (I/Q) samples coming from the receiver and assign each I/Q sample to a specific radio. In contrast to existing approaches, `M2RF` simultaneously identifies and locates in the spectrum multiple devices that emit overlapping signals and avoids the burden of processing data, making the overall approach with reduced overhead and faster. Our approach can be generalized to different channels and signal bandwidths without retraining, making it scalable. Experiments in three different spectrum scenarios under 2 transmission conditions and with 15 radio transmitters demonstrate the effectiveness of `M2RF`, achieving up to 99.56% of F1-score, and 92.44% detection rate of malicious users with only a 2.72% mean Miss

Email addresses: `ildi.alla@inria.fr` (Ildi Alla), `zhang.mil@northeastern.edu` (Milin Zhang), `jonathan.ashdown@us.af.mil` (Jonathan Ashdown), `valeria.loscri@inria.fr` (Valeria Loscri), `f.restuccia@northeastern.edu` (Francesco Restuccia)

Rate (MR). Dataset and code will be shared for reproducibility and a demo video is available (M2RF – Demo Video).

Keywords: RF Fingerprinting, Multi-Device Authentication, Semantic Spectrum Segmentation, Dynamic Spectrum Sharing, Anomaly Detection

1. Introduction

The sheer growth of the Internet of Things (IoT) is quickly saturating unlicensed spectrum bands [1]. As unlicensed bands become saturated, *spectrum sharing* will become one of the very few options to sustain the IoT growth in the years to come [2, 3, 4, 5]. The key issue is that today, IoT operators that want to share spectrum with licensed users – also called *incumbents* – must contact database systems located in the cloud, which determine if the spectrum is available based on geographical coordinates [6]. *This centralized manual approach lacks scalability and does not allow for fine-grained real-time spectrum management.* Conversely, a scalable and effective solution would be to let IoT devices opportunistically discover which spectrum sub-bands are currently available among ongoing licensed transmissions, provided they do not cause harmful interference to incumbents [7].

It is easy to observe that dynamic spectrum access systems will create fundamentally new security challenges where incumbents must be protected by secondary users not abiding by spectrum rules. To prevent such issues, spectrum must be *continuously monitored* to make sure only authorized devices are using the spectrum. Traditional wireless authentication systems such as WPA for Wi-Fi [8] or 5G-AKA for cellular networks [9] are based on cryptography or password-based authentication. As such, they operate primarily on the network or application layers, failing to meet the real-time requirements for spectrum sharing [10], [11]. In addition, these methods are proven insufficient against various attacks, such as spoofing, replay, and impersonation attacks [12, 13, 14].

In recent years, *radio fingerprinting* has emerged as a viable approach to spectrum-level authentication. Specifically, radio fingerprinting leverages the inherent hardware imperfections present in every radio circuitry [15, 16, 17, 18] to form a unique and unforgeable "fingerprint" that can authenticate devices [19]. By exploiting these characteristics, radio fingerprinting offers a security solution that is resistant to attacks such as MAC address spoofing and identity cloning [20].

Existing work – discussed in details in Section 2 – has a series of core limitations that make it unable to perform real-time spectrum-level authentication. Specifically, Figure 1 shows the fundamental difference between prior work and our proposed approach. **First**, current approaches only classify one signal in a given channel of interest. Conversely, multiple signals are usually overlapping in adjacent bands making the classification problem harder. **Second**, conventional methods assume prior knowledge of operating frequency of transmitters and only classify signals in that specific frequency band. However, signals may be partially observed by the receiver, e.g., because they are partially outside the operating bandwidth.

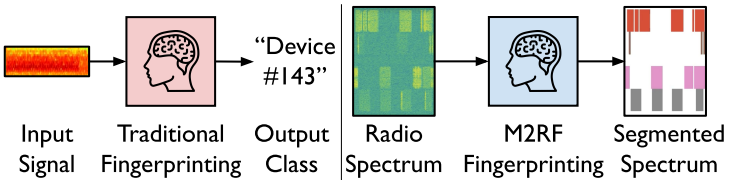


Figure 1: Traditional Radio Fingerprinting vs `M2RF`.

This paper changes the current state of the art by proposing the first ever spectrum-level authentication system named *Multi-band Multi-device Radio Fingerprinting* (M2RF), where multiple devices are located and identified in the same spectrum band using *spectrum segmentation*. The right side of Figure 1 shows at a very high level the main objective of `M2RF`. The proposed approach directly operates on unprocessed in-phase/quadrature (I/Q) inputs coming from the radio receiver front-end, thus eliminating pre-processing steps. The proposed spectrum segmentation model, based on a Deep Neural Network (DNN), has been specifically designed to handle dynamic signal and channel bandwidths through the integration of a non-local block, which captures long-range dependencies across frequency and distinguishes subtle differences in RF signals via a self-attention mechanism. In addition, `M2RF` incorporates a combined loss function that integrates both local-level and region-level features, further enhancing its ability to learn intricate signal features while maintaining consistent accuracy. An aggregation block is built to support wideband classification by combining predictions across overlapping frequency bands, allowing the model to span and accurately identify signals across different frequency segments.

The unique features of our solutions are very promising on different open problems. From one side, it is possible to detect in real-time the *intrusion*

tentative from malicious nodes. Another really interesting perspective provided by our approach is the capacity to "locate" in the spectrum the activity of malicious devices, posing the fundamentals of advanced anti-jamming solutions, targeting with high precision the "malicious" operating frequencies. From another perspective, our approach permits to better manage the shared resources, due to the real-time information of who (device) is where (in the spectrum).

Summary of Novel Contributions

• We propose a real-time technology-independent radio fingerprinting approach named `M2RF` that can simultaneously fingerprint multiple devices coexisting in the shared spectrum. `M2RF` includes (i) a scalable dataset generation pipeline that can represent real-world spectrum conditions such as overlapping signals, and (ii) an energy-efficient DNN optimized for resource-constrained devices. To the best of our knowledge, this is the first work proposing a simultaneous multi-device radio fingerprinting system;
• We introduce a new anomaly detection mechanism to detect adversary and interference in spectrum sharing scenarios. We leverage Total Variation (TV) analysis to identify attacks by detecting irregularities in the DNN output. Specifically, it exploits the fact that the DNN produces noisy and randomized outputs when fed with an unseen signal. This means that real-time detection is realized without prior knowledge of the specific attacks strategy;
• We evaluate the performance of `M2RF` using a comprehensive 82 GB dataset of over-the-air (OTA) data from 15 identical Wi-Fi cards, which represents the worst case for radio fingerprinting as identical devices may have closer fingerprints [21]. In addition, we have collected data via wired connection to have data unaffected by the wireless channel [22]. To simulate real-world threats, we consider both informed and uninformed adversaries. For *informed adversary*, we collected data from additional identical Wi-Fi devices having full knowledge of the authentication approach. For *uninformed adversary*, we collected data from different Wi-Fi cards. Moreover, we collect other wireless technologies as interference (e.g., BLE, LTE, Zigbee) to evaluate the `M2RF` performance in congested multi-technology environments;
• Our experimental results show that `M2RF` achieves F1-score of 94.99% and Intersection over Union (IoU) of 90.54% with over-the-air non-overlapping signals. In the challenging scenario of overlapping signals, `M2RF` achieves F1-score of 77.06% and IoU of 63.39% without retraining and/or fine-tuning. Moreover, `M2RF` detects adversaries with an accuracy of 92.44%, demonstrat-

ing resilience against both informed and uninformed attacks. When other technologies are present, M2RF achieves an overall accuracy of 81.52%. A demo video of M2RF is available (M2RF – Demo Video).

2. Background and Motivation

Radio fingerprinting is a technology that authenticates wireless devices based on unique characteristics inherent in their transmitted radio signals [23]. The key idea is based on the fact that each radio device has its unique hardware imperfections in its circuitry, which manifest as subtle yet measurable differences in signal transmission. Compared to conventional cryptography-based methods, radio fingerprinting offers a more robust authentication mechanism since these physical properties are inherently unclonable. Furthermore, by operating directly at the physical layer, radio fingerprinting provides greater agility and computational efficiency without requiring full-stack protocol operations.

With the rapid development of IoT, there is an increasing need of wireless communication service to connect massive devices to network. To maximize spectrum efficiency in massive connectivity scenarios, dynamic spectrum management has been proposed to enable opportunistic signal transmission in available sub-bands [24]. To this end, a scalable and rapid authentication method is required to identify massive IoT devices in real time. The computational agility inherent in waveform-level operations makes radio fingerprinting a promising candidate for this purpose.

However, existing radio fingerprinting approaches fail to address the following challenges in spectrum sharing:

• **Dynamic Operating Frequency**. Existing approaches involve band filtering and pre-processing to remove interference and channel effect before classification [25], which assumes prior knowledge of the bandwidth and operating frequency of transmitted signal. However, in the spectrum sharing system, devices can dynamically select their operating frequencies based on spectrum availability. Therefore, the target signal may not operate at the same frequency as assumed by the radio fingerprinting method, or may even fall partially outside the filter bandwidth, causing the algorithm to fail in identifying the target device.

• **Simultaneous Transmissions**. In spectrum sharing, multiple devices can transmit simultaneously within the same spectrum. This creates a significant

scalability challenge for existing radio fingerprinting methods, as current approaches can only authenticate one device at a time [22]. To classify multiple signals in real time, these methods must iteratively process multiple signal instances across different operating frequencies, which introduces considerable computational overhead and latency.

• **Uncontrolled Interference**. Another significant challenge in spectrum sharing is that the uncontrolled spectrum environment can have considerable noise and interference which can compromise the accuracy of fingerprinting. As the spectrum is an open resource, interference may exist intentionally or unintentionally. Current fingerprinting approaches designed to work in controlled environments and with minimal interference can fail to generalize to the complex and varying environment [26].

These limitations motivate us to create a brand-new radio fingerprinting design that can simultaneously identify multiple signals in the spectrum in real time. Specifically, we aim to address the following research questions:

• **RQ1 – How to properly model and process the complex spectrum environment?**

A typical data pre-processing pipeline in conventional radio fingerprinting involves shifting signals to the operating frequency, removing noise outside the band of interest, and processing the signal within the band for feature extraction [22]. This approach ensures that the data is controlled and purified to improve classification performance but requires prior knowledge of operating frequency and can only identify one radio at a time, which does not meet the requirements of spectrum sharing environments. To overcome this limitation, a new pipeline is needed to model and process complex spectrum environments where multiple radio transmissions do not match the expected operating frequency or are only partially observable by the receiver.

To address **RQ1**, we create a new data pre-processing pipeline – described in Section 4.1 – that can simulate complex spectrum conditions with controlled data transmission and pre-processing. This pipeline, similar to conventional radio fingerprinting approaches, creates a purified signal repository that enables the algorithm to effectively learn useful features in the signals of interest. However, it is distinct from other radio fingerprinting methods by augmenting and stitching multiple purified signals to simulate real spectrum conditions, where multiple signals coexist, overlap, or are partially observable.

• **RQ2 – How can simultaneous radio device authentication be achieved in this spectrum environment?**

A naive approach to achieving multiple device authentication would be to extend current radio fingerprinting methods to iteratively process each signal in the spectrum. However, this approach requires additional complexity to identify the center frequency of each radio transmission, as signals are transmitted dynamically throughout the spectrum. Moreover, the computational burden increases significantly when massive transmissions occur within the same spectrum, resulting in substantial latency. Therefore, developing a new fingerprinting paradigm that can achieve simultaneous multi-device authentication is a critical research challenge in dynamic spectrum sharing.

In Section 4.2, we address **RQ2** by leveraging a novel DNN solution based on "semantic spectrum segmentation". The neural network is trained to take I/Q samples as input, and directly segment waveforms in the frequency domain. This way, it removes the complexity of per-signal processing and hence achieving a real-time multi-device authentication.

• RQ3 – Is the new framework generalizable to different frequency and environment?

Spectrum sharing requires real-time monitoring of an ultra-wide spectrum band which is typically larger than the observable bandwidth of the spectrum sensor. For example, [27] considered a scenario covering the frequency range from 400 MHz to 6 GHz. It is infeasible to create a single fingerprinting algorithm to monitor the entire 6 GHz spectrum due to hardware constraints. A viable approach is to divide the entire spectrum into multiple channels and leverage fingerprinting to rapidly scan these channels. However, existing radio fingerprinting methods developed in controlled environments often struggle to generalize effectively to different environments and scenarios. For instance, [28] reported an 82% accuracy drop when testing in real-world scenarios. To ensure the approach is practical in real-world applications, it is critical to validate whether the proposed framework is generalizable.

To answer this question, we design adaptive bandwidth processing by sweeping and aggregating fingerprinting results across multiple channels. This approach enables the fingerprinting algorithm to be tested across different channel bandwidths and signal bandwidths. We describe this adaptive processing in Section 4.3.

• RQ4 – Can the algorithm detect interference and adversaries in the spectrum?

As the spectrum is an open resource, interference may occur intentionally or unintentionally. For example, an unauthorized device may attempt to occupy a channel by intentionally mimicking the behavior of authorized ra-

dios through spoofing or replay attacks. In addition, other unknown signals may transmit within the band and cause unintentional interference. Such interference may cause misclassification of the fingerprinting system, hence compromising the spectrum management. Therefore, detecting interference and adversaries in the spectrum is as important as identifying the authorized devices.

To address **RQ4**, we propose a post-processing method based on total variation that can effectively detect anomaly in the spectrum. The key idea is that interference will have lower certainty in the inference results, which can be detected by checking the consistency of inference. A detailed discussion of this method is provided in Section 4.4.

3. Threat Model and System Overview

As spectrum is an open resource, malicious traffic poses significant threats to legitimate users. For example, adversarial devices may attempt to authenticate by cloning legitimate user behavior, while unintentional interference can occur when signals are transmitted on the same channel. Both adversaries and interference can severely degrade the quality of service for authorized users. Therefore, an effective spectrum management requires the algorithm not only to authenticate legitimate users but also to detect malicious traffic in the spectrum. In this section, we outline these potential threats in the spectrum, as well as how `M2RF` is structured to defend against these challenges in a high level.

3.1. Threat Model

Figure 2 overviews the threat model in the dynamic spectrum management. Alice, the authorized user, will access the network through a specific channel with its unique hardware characteristics. On the other hand, Eve, the adversary tries to access the network through the same channel by cloning Alice's behavior. Bob, the authenticator, continuously monitors the spectrum to authenticate Alice and detect Eve using radio fingerprinting techniques. In this scenario, Eve can perform different attack strategies:
① *Spoofing.* This scenario involves an attacker (Eve) emulating the credentials of an authorized device (Alice) by cloning identifiers such as MAC addresses. Here, Eve's goal is to deceive the authentication system by masquerading as a legitimate device without replicating the hardware-specific imperfections that are unique to Alice.

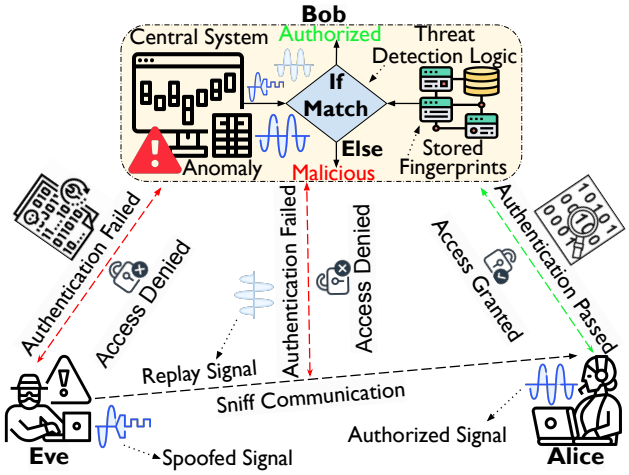


Figure 2: Overview of interactions among Alice (authorized device), Eve (attacker), and Bob (authenticator) to detect and prevent spoofing and replay attacks through radio fingerprinting.

② *Replay Attacks.* Eve intercepts transmissions from Alice and replays them, aiming to deceive the system and gain unauthorized access without the need to directly imitate Alice's radio signal characteristics. Replay attacks exploit captured communication sessions, assuming they will appear legitimate upon retransmission.
③ *Device Impersonation.* Eve manipulates signal characteristics to closely imitate Alice's fingerprint. By using similar devices or attempting software-based modifications [29], Eve aims to create a sufficiently close match to bypass RF fingerprint detection. This approach assumes that Eve has knowledge of Alice's signal characteristics and attempts to mimic them. Still, the unique hardware imperfections inherent to Alice's device cannot be acquired by Eve.

3.2. Type of Malicious Traffic

In our system, we consider three different type of malicious traffic based on its knowledge model for the legitimate device:
Informed Adversary. Eve possesses detailed knowledge of Alice's hardware features, the radio fingerprinting algorithm and how the authentication is performed by Bob. This knowledge allows Eve to adopt more sophisticated techniques to approximate Alice's signal characteristics.

Uninformed Adversary. Eve has basic knowledge about the system such as the wireless technology (e.g. Wi-Fi) but lacks specific knowledge about Alice's hardware imperfections and Bob's detection mechanisms. Eve may attempt standard spoofing or basic replay methods without insight into the physical layer defense, relying on generic attack methods.
Interference. Eve has no knowledge about the system. It unintentionally occupies the channel and creates malicious interference to the legitimate user. This scenario represents a common shared spectrum situation where different wireless technologies operate in the same frequency band (e.g. Bluetooth vs Wi-Fi).

3.3. Defense Mechanism

In a high level, the proposed framework `M2RF` achieves robust authentication with following strategies:
① *Real-Time Monitoring.* `M2RF` continuously monitors spectrum of interest, adapts to different frequency bands and bandwidths with scanning and aggregation as described in Section 4.3 for authentication of legitimate users and detection of malicious traffic.
② *Multi-Device Authentication.* As discussed in Section 4.2, `M2RF` leverages a Deep Learning (DL)-driven semantic segmentation algorithm to directly label each waveform data (I/Q samples) in the frequency domain based on their waveform features, which results in a simultaneous labeling of all waveforms in the channel.
③ *Anomaly Detection.* `M2RF` detects malicious traffic across multiple frequency bands based on fingerprint consistency check. Compared to legitimate signals, malicious signals show increased randomness in the inference results, which can be detected by total variation as detailed in Section 4.4.

4. The `M2RF` Framework

To address the research questions outlined in Section 2, we propose `M2RF`, a new radio fingerprinting framework for spectrum sharing. Figure 3 overviews the main components of `M2RF`. The process begins with acquiring I/Q data, followed by a novel pre-processing pipeline to address **RQ1**. These signals are then used to train a DNN for "semantic spectrum segmentation", a new approach which effectively address **RQ2**. During inference phase, adaptive

bandwidth processing and anomaly detection modules are proposed to address **RQ3** and **RQ4** respectively, ensuring the proposed framework is practical in real-world scenarios. We explain each component of `M2RF` in the following subsections.

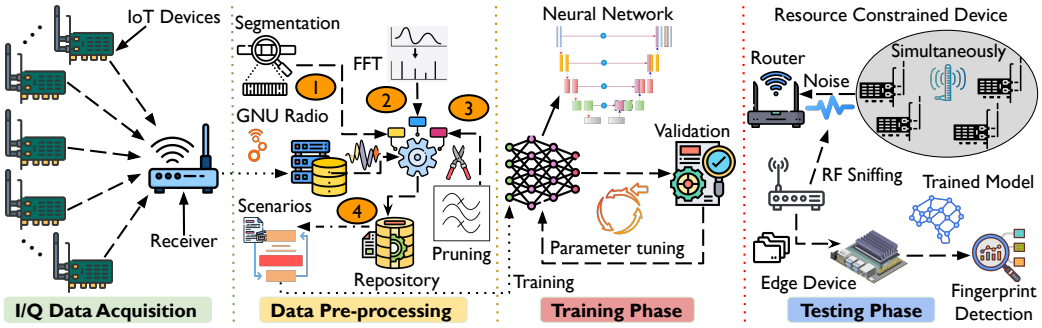


Figure 3: Overview of `M2RF`, from I/Q data acquisition through data pre-processing, training, and testing in real-world scenarios.

4.1. Data Pre-processing

Data collected in controlled environments often fails to generalize to complex real-world spectrum conditions. However, directly collecting spectrum data in open environments presents its own difficulties: unknown signals may be transmitting simultaneously, creating interference that compromises algorithm performance. Additionally, labeling data containing these unknown signals is inherently challenging, as their sources and characteristics cannot be readily identified.

To address **RQ1**, we introduce a novel data pre-processing pipeline that effectively simulates comprehensive spectrum conditions through controlled data collection. This pipeline comprises two components: controlled data collection to establish a comprehensive signal repository, and a data augmentation process that simulates real-world scenarios using signals from the controlled environment. Note that this pipeline is only applied during the training phase. During inference, the fingerprinting algorithm operates directly on real-world spectrum data.

4.1.1. Signal Repository

We first build a curated collection of individual, high-quality radio signals captured under controlled conditions. These signals are recorded sequentially,

ensuring that only one transmission occurs at a time, with known center frequencies f_c and bandwidths B. This controlled environment ensures that each signal is free from external interference or overlapping transmissions, capturing the characteristics necessary for subsequent processing and accurate radio fingerprinting. Once collected, each radio signal $s(t)$ undergoes the following procedure:

❶ **Segmentation**: The continuous time-domain signal $s(t)$ is segmented into smaller, fixed-length portions to capture individual signal instances. We divide the continuous recording into segments of duration T, which corresponds to a fixed number of I/Q samples, to isolate relevant transmissions. The segmentation process can be represented as:

$$s_{\text{seg}}[n] = s[n] \cdot w[n], \tag{1}$$

where $w[n]$ is a rectangular windowing function defined in discrete time as:

$$w[n] = \begin{cases} 1 & \text{for } 0 \leq n < N \\ 0 & \text{otherwise} \end{cases}, \tag{2}$$

where N is the fixed number of I/Q samples in each segment and $s_{\text{seg}}[n]$ represents the n-th sample of the segmented signal. The segmented signal $s_{\text{seg}}[n]$ contains a fixed duration of the transmission, ready for frequency domain processing.

❷ **Fast Fourier Transform (FFT)**: The segmented time-domain signal $s_{\text{seg}}[n]$ is then converted to the frequency domain using the Fast Fourier Transform (FFT). This transformation yields the frequency spectrum $S_{\text{fft}}(f)$, which represents the signal's frequency components:

$$S_{\text{fft}}(f) = \text{FFT}\{s_{\text{seg}}[n]\}. \tag{3}$$

❸ **Pruning of Unwanted Frequency Components**: To focus on the signal band of interest and eliminate out-of-band noise, we apply frequency pruning. This step retains only the frequency components within the bandwidth B around the center frequency f_c, effectively isolating the relevant spectral portion for radio fingerprinting. The pruned signal $S_{\text{pruned}}(f)$ is obtained by applying a binary mask in the frequency domain:

$$S_{\text{pruned}}(f) = S_{\text{fft}}(f) \cdot M(f), \tag{4}$$

where $M(f)$ is a frequency mask defined as:

$$M(f) = \begin{cases} 1 & \text{for } f_c - \frac{B}{2} \leq f \leq f_c + \frac{B}{2} \\ 0 & \text{otherwise} \end{cases}. \tag{5}$$

❹ **Storage in Signals Repository**: The pruned frequency-domain I/Q samples $S_{\text{pruned}}(f)$ are then stored in the signals repository. This repository serves as a comprehensive and clean dataset of individual radio fingerprints in the frequency domain. It is designed for subsequent use in analysis, scenario simulation, and training of the fingerprinting model, ensuring the necessary data quality for accurate device identification. The pre-processing pipeline, based on our implemented steps, can be summarized as the transformation:

$$s(t) \rightarrow s_{\text{seg}}[n] \rightarrow S_{\text{fft}}(f) \rightarrow S_{\text{pruned}}(f) \rightarrow sig_{\text{repo}}.$$

4.1.2. Simulation of Scenarios

Scenario generation is a crucial step that simulates a wide variety of real-world environments. By integrating multiple signals into a "stitched" wideband signal, this approach reduces the need for extensive real-world data collection.

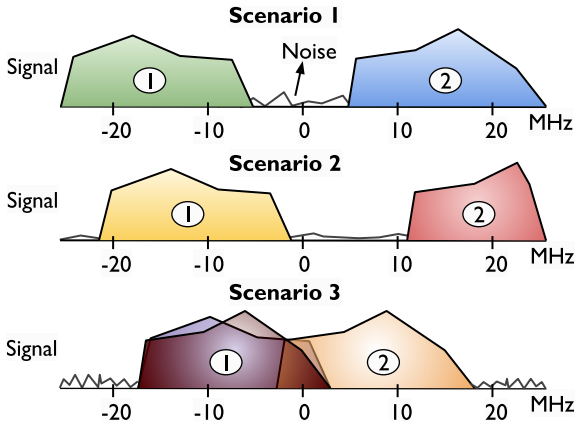


Figure 4: Visual representation of scenarios: (Top) Scenario 1 with non-overlapping signals, (Middle) Scenario 2 with randomly positioned signals, and (Bottom) Scenario 3 showcasing partial or full overlapping signals, all within a 50 MHz bandwidth.

By using such stitching procedure made with individual signal widths, our pipeline dynamically generates numerous training samples that can work

with different signal bandwidths. Figure 4 shows an example of scenario generation with 50 MHz observable bandwidth and 20 MHz wide signals. The scenarios are generated with a procedure described in Algorithm 1, which assembles signals into a complete training sample via 'spectrum stitching'.

Algorithm 1 Sample generation using spectrum stitching.

Require: sig_{repo}, buf, B, $signal_{bw}$, n_{iq}, $max_{signals}$, $prob_{empty}$, $prob_{centered}$
Decide if the bandwidth is empty based on $prob_{empty}$
if bandwidth is not empty **then**
    Select randomly $n_{signals}$ from 1 to $max_{signals}$
    Randomly choose transmitters from sig_{repo}
    for each card signal **do**
        Extract corresponding signal from sig_{repo}
        Determine center frequency f_{center}
        if Scenario 1 **then**
            Sequential placement in buffer
        else if Scenario 2 **then**
            Random placement without overlap
        else if Scenario 3 **then**
            Allow overlaps
        end if
        Update $label_{matrix}$ and buf
    end for
end if
Add background noise from sig_{repo}
Stitch signals and noise to finalize $input_{samp}$
return $input_{samp}$ and $label_{matrix}$

The algorithm first determines if the observable bandwidth will be empty, based on the probability $prob_{empty}$. If not empty, it selects a number $n_{signals}$ between 1 and $max_{signals}$ of signals from the repository sig_{repo}. Placement within the bandwidth is guided by a center frequency f_{center}, chosen randomly based on the parameter $prob_{centered}$, which controls whether signals are centered or distributed. The final algorithm step adds background noise sourced from the signals repository to the stitched signal to simulate realistic conditions. The resulting sample is then stored with its label matrix. The resolution of scenarios, defined by the frequency sub-band size into which the observable bandwidth is divided, is given by: $resolution\ (R) = \frac{B}{n_{iq}}$. This resolution sets the granularity for analyzing and classifying the signal spectrum.

The label matrix, structured as $C \times n_{iq}$ (where C is the number of classes), enables fine-grained classification across the bandwidth.

4.2. Spectrum Fingerprinting

As discussed in Section 2, conventional fingerprinting algorithms that classify one device at a time cannot scale to multi-device authentication. To address **RQ2**, we propose an approach based on semantic segmentation. This approach directly takes wideband RF data as input and labels each I/Q waveform in the frequency domain, enabling simultaneous localization and detection of multiple devices across the spectrum. The following subsections outline the multi-label semantic segmentation methodology, the structure of the DNN model, its adaptability to varying input sizes, generalization strategies, and the scalable processing techniques we implemented. Additionally, we provide a detailed discussion of the specific adversarial detection technique used in this approach. A detailed explanation of the various loss functions is provided in Appendix A.

4.2.1. Multi-Label Signal Segmentation

Our approach utilizes semantic segmentation, a technique originally developed for computer vision tasks. The key idea is to segment objects from the background by labeling each pixel belonging to those objects based on their semantic information within the frame. Similarly, we apply this idea to spectrum sharing tasks, identifying target signals within noisy spectrum environments.

Specifically, we transform the captured waveform into the frequency domain and divide it into multiple sub-channels. A DL-based semantic segmentation algorithm is then applied to detect signals across the bandwidth. Similar to image-based semantic segmentation, our signal segmentation approach labels each sub-channel based on waveform-level features. This enables the simultaneous detection and classification of multiple overlapping signals within the bandwidth.

One significant difference between the image segmentation and signal segmentation is the multi-label nature of the radio environment. In an image, the object in the behind will be blocked by the object in the front by its non-transparent nature. In contrast, different signals can coexist within the same frequency band without occluding each other in the radio environment. As a result, each frequency bin can be assigned to multiple classes simultaneously.

Therefore, we extend the semantic segmentation algorithm to output a binary segmentation map for each class, where each map indicates the presence or absence of the corresponding class within the given frequency bin. The final segmentation output is a matrix where each row corresponds to a class and each column corresponds to a frequency bin.

4.2.2. DNN Model Architecture

Our backbone is inspired from U-Net, which was initially proposed for biomedical image segmentation [30]. We adapted this architecture for radio fingerprinting by replacing the 2D convolutional layers with 1D convolutions to process I/Q samples effectively. As illustrated in Figure 5, the architecture comprises five encoding and five decoding blocks. The encoding path systematically downsamples the input data, capturing features at varying levels of abstraction through 1D convolutional layers, batch normalization, and ReLU activations. Max pooling layers are employed within each encoding block to reduce the spatial dimensions.

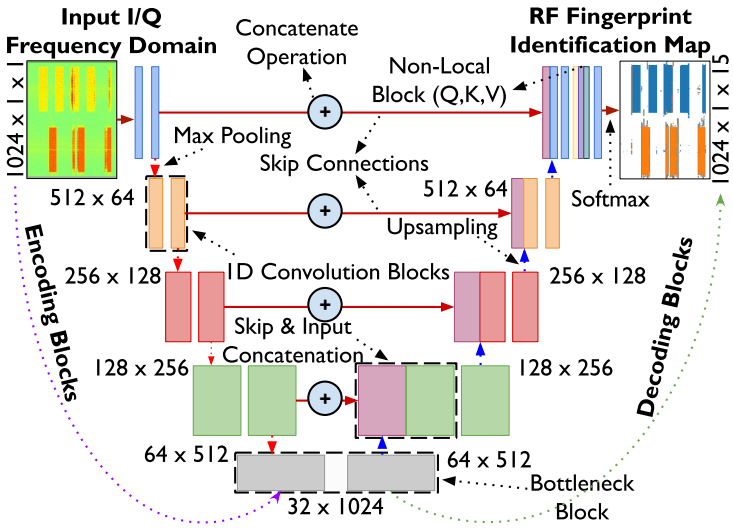


Figure 5: Adapted U-Net architecture for radio fingerprinting. The left side shows the encoding path for feature abstraction, the middle includes the bottleneck and non-local block for long-range dependencies, and the right side represents the decoding path.

The decoding path mirrors the encoding process, progressively reconstructing the data to its original size using upsampling layers. Skip connections between corresponding encoding and decoding blocks ensure that

spatial information, crucial for accurately identifying device-specific characteristics in radio signals, is preserved throughout the DNN. The final layer applies a 1x1 convolution to produce a multi-channel output, resulting in a $C \times n_{iq}$ matrix, where each channel corresponds to a different class in the multi-label segmentation task.

4.2.3. Integration of Non-local Block

One issue of U-Net is that its architecture is fully comprised of convolutional neural networks (CNNs). However, conventional CNNs often struggle with capturing long-range dependencies [31], especially in radio applications where signals are wide-spread to large bandwidth in the spectrum. As such, conventional U-Net may misclassify a portion of sub-channels as the output is based on a group of neighboring features in sub-channels without considering the dependencies across frequency. As such, a non-local block is incorporated into the last layer to enhance the performance. The non-local block addresses this by introducing a self-attention mechanism that allows the network to consider the global context of the feature maps. The self-attention is defined as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V, \tag{6}$$

where Q, K, and V represent the Queries, Keys, and Values, respectively, derived from the feature maps using 1x1 convolutions. Here, d is the embedding dimension. This operation computes a weighted sum of the entire feature map, effectively enabling the model to capture long-range dependencies. The integration of the non-local block ensures that the DNN can accurately classify RF signals, even in scenarios where signals overlap or interfere with each other.

During training, we estimate the noise floor across the training dataset by recording the minimum values of the smoothed signal power in the frequency domain. This noise estimate is then used to normalize the input signals during inference, ensuring that the model can generalize to different noise levels encountered in real-world deployments. The effectiveness of this approach is enhanced by the scenario generation process discussed in Section 4.1.2, which introduces a variety of signal placements, overlaps, and noise conditions into the training data. This diversity ensures that the model can handle a wide range of RF conditions without overfitting to specific scenarios.

4.3. Adaptive Signal Bandwidth and Channel Bandwidth Processing

As discussed in **RQ3**, spectrum sharing often requires the system to monitor a broader spectrum than their observable channel bandwidth B. Additionally, transmitted signals may dynamically adjust their signal bandwidth W based on available spectrum resources and throughput requirements. To address this challenge, we implement adaptive signal and channel bandwidth processing by scanning channels and aggregating the results.

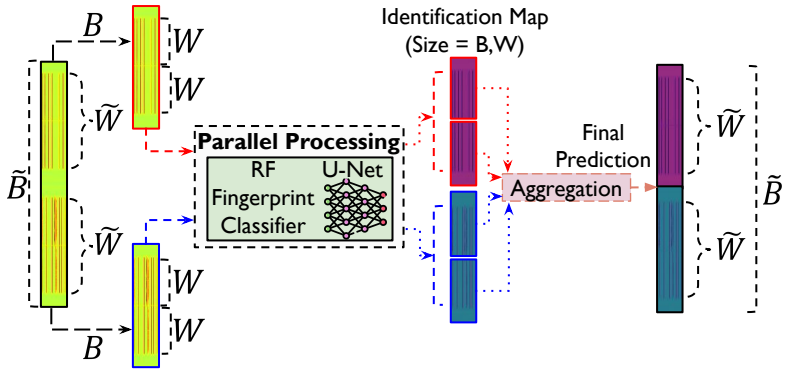


Figure 6: Adaptive wide-band and channel bandwidth processing pipeline: The input signal covering a larger bandwidth $\tilde{B}$ or wider channel width $\tilde{W}$ is divided into overlapping segments of size B or W, processed by the U-Net model. The outputs are then aggregated to produce the final prediction across the full signal or bandwidth.

Our key intuition is that hardware imperfections are intrinsic to the physical components of the device and are thus independent of the signal bandwidth. Thus, when a signal with a bandwidth larger than W is received, the `M2RF` divides it into smaller segments, each matching the W for which the model was trained. Similarly, when faced with a signal that spans a larger observable bandwidth $\tilde{B} > B$, `M2RF` divides the larger bandwidth into smaller, partially overlapping segments, each of size B. Each segment is processed individually by the DNN and the outputs are combined to form a final output that covers the entire bandwidth $\tilde{B}$. After processing, the predictions from these individual segments—whether divided by signal width or observable bandwidth—are aggregated to form a cohesive understanding of the entire wider signal $\tilde{W}$ or bandwidth $\tilde{B}$. This aggregation step ensures that even when the signal spans a larger width or bandwidth, the model's predictions

are consistent and accurate, effectively identifying the unique radio fingerprint embedded within the signal. This capability highlights the scalability and portability of our approach, making it highly versatile for deployment across various RF environments where both signal widths and observable bandwidths can vary significantly.

Figure 6 shows the adaptive pipeline that efficiently processes signals exceeding the standard training bandwidth. Specifically, for a input that has larger signal bandwidth W or observable bandwidth B, the pipeline first divide it into multiple overlapping chunks and each segment will have a individual score output by the signal segmentation model. After that, the aggregation block is used to average the overlapped output of multiple chunks while keeping the same score for non-overlapping output. This method ensures that our system can accurately process and analyze signals across a wide range of bandwidths and channel widths, maintaining high precision in radio fingerprinting.

4.4. Anomaly Detection

Beyond identifying legitimate devices, we are also interested in detecting interference and malicious traffic in the spectrum. To address **RQ4**, we introduce a novel anomaly detection approach by leveraging the uncertainty in the DNN output. During training, only legitimate signals are used, which results in confident predictions for authorized devices. Conversely, a malicious signal or interference not seen during training will be less confident, thus enabling `M2RF` to detect adversaries by evaluating the randomness of the spectrum map.

To quantify adversarial activity, we apply total variation, which evaluates the consistency of the DNN output across the frequency domain. Higher total variation indicates a higher likelihood of a malicious signal. For a 1D vector x, the total variation is defined as:

$$TV(x) = \sum_{i=0}^{N-1} |x_{i+1} - x_i|, \tag{7}$$

where x_i is the i-th element in vector x while N denotes the dimension of the input. While total variation is first introduced for denoising [32, 33], the element-wise distance $|x_{i+1} - x_i|$ evaluates the consistency in DL module output in our case, making it a good metric to detect the malicious user who constantly has a noisier output than legitimate user. For example, the total

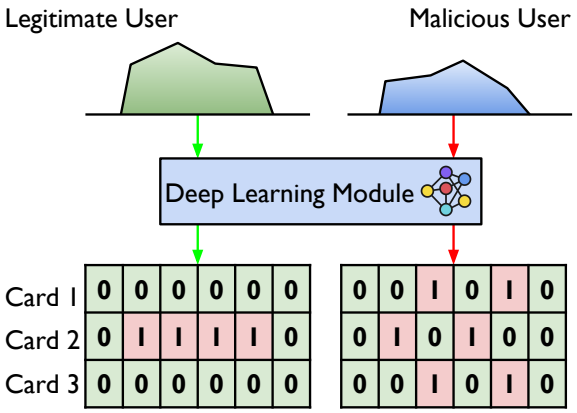


Figure 7: An example of DNN output for a legitimate user vs malicious user. The inference map presents high randomness while legitimate user output has more consistency.

variation of the legitimate user output in Figure 7 is 2 while the malicious user output has a total variation of 12. By increasing the resolution in frequency domain (e.g., in our experiment we use 4096 as the input and output size), the difference of total variation between legitimate and malicious users will increase significantly.

By comparing the TV values for legitimate and malicious signals, we set a detection threshold λ_m:

$$TV(x) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_m, \tag{8}$$

where H_0 denotes the hypothesis that signal x is not an adversary and H_1 denotes the alternative hypothesis that x is considered as adversary.

5. Experimental Setup

Our data collection setup captures radio fingerprints under two distinct scenarios—wireless and wired—using sophisticated hardware to ensure accuracy and reliability. As shown in Figure 8, the configuration includes a Multiple-Input Multiple-Output (MIMO) system for the simultaneous transmission of 15 PCI-E wireless LAN cards, all identical in model and version (802.11ac/ax). This choice of identical devices creates a challenging test scenario, generating highly correlated signals to rigorously test `M2RF`'s ability to

distinguish between identical transmitters. The ASUS RT-AX86U router is used as the primary receiver, and I/Q data is captured via an USRP X310 and USRP B200mini, each equipped with VERT2450 and L-com antennas.

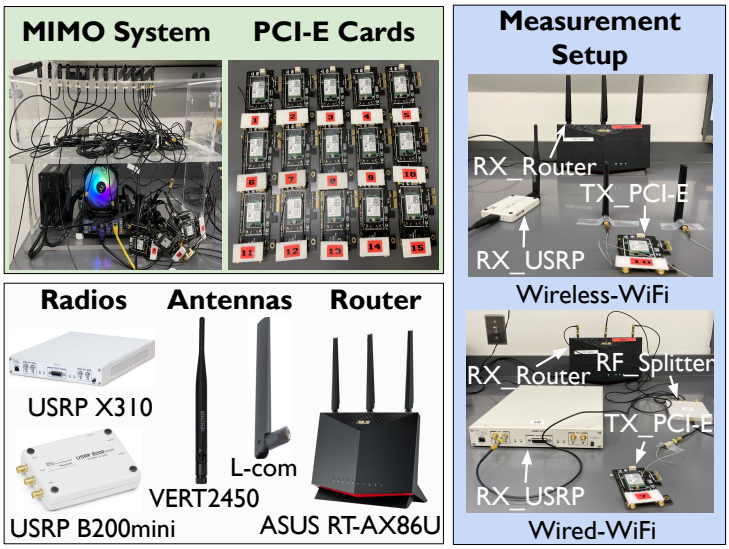


Figure 8: Overview of the data collection testbeds; (Top Left): MIMO system with PCI-E cards for simultaneous transmission; (Top Right): Individual PCI-E cards; (Bottom Left): Radios, antennas, and router setup; (Right): Measurement setup for wireless and wired data collection.

We created two testbeds to evaluate radio fingerprinting performance under both wireless and wired data collection methods. The wireless setup, where radio transmissions from the PCI-E cards are captured by the USRP radios through antennas, simulates a realistic, uncontrolled environment, typical of actual deployments. This setup provides insight into radio fingerprint behavior in dynamic conditions affected by interference, multi-path effects, and environmental variability. Conversely, the wired setup provides a controlled baseline with a higher signal-to-noise ratio (SNR) of around 20-25 dB, compared to 15-20 dB in the wireless setup. This comparison highlights the robustness of our radio fingerprinting approach across varying SNR levels and operating conditions.

Data were collected over three days within a laboratory setting to capture a wide range of signal conditions. This approach accounted for environmental factors like temperature fluctuations and electromagnetic interference. Data was collected across two specific frequency bands – 5.5 GHz (channel 100) and

5.6 GHz (channel 120) – within a 50 MHz observable bandwidth, with each Wi-Fi card transmission occupying a 20 MHz bandwidth. To prevent signal overlap and ensure distinct RF fingerprints, each PCI-E card's transmission was captured separately, achieving the precision necessary to differentiate between devices with nearly identical hardware profiles.

5.1. Hardware Characteristics

DNN training was conducted on a system featuring 4 NVIDIA A100 80GB PCIe GPUs, 512 GB RAM, and dual Intel Xeon Silver 4410Y processors, ensuring efficient processing for large-scale DNN tasks. After training, the DNNs were run on the Jetson Orin Nano module, powered by a 6-core ARM Cortex-A78AE 64-bit CPU. The system is equipped with 8 GB of 128-bit LPDDR5 memory and a 1024-core NVIDIA Ampere architecture GPU with 32 Tensor Cores, capable of delivering up to 40 TOPS.

5.2. Experimental Dataset and Training

Experiments were conducted in both wireless and wired modes, primarily focusing on signals within a 50 MHz bandwidth. The primary dataset includes signals from 15 authorized devices, comprised 1.25 million samples, with 80% allocated for training and 20% for testing across three distinct scenarios: non-overlapping, overlapping, and partially observed signals. For training, we used the Adam optimizer with a StepLR learning rate scheduler, starting at 0.001 and reducing by a factor of 0.1 every 30 epochs. The DNN was trained for 100 epochs with a batch size of 1024, with early stopping applied after 30 epochs of no improvement to prevent overfitting.

To rigorously test the `M2RF` resilience against potential attacks, we prepared additional testing datasets for specific attack scenarios:

• **Adversary:** For informed adversary, we collected Wi-Fi data from unauthorized devices identical in hardware to the authorized devices but excluded these samples from training. For uninformed adversary, we collected data using Wi-Fi protocol but from devices having different hardware characteristics than those used during training;

• **Interference**: We collected data with devices using Bluetooth Low Energy (BLE) with bandwidth 1 MHz, Long Term Evolution (LTE) with bandwidth 10 MHz, Zigbee with bandwidth 2 MHz as interference, as well as with additional Wi-Fi devices in the 2.4 GHz band.

6. Performance Evaluation

6.1. Performance Across Input Sizes and Scenarios

6.1.1. Wireless Mode

We started with wireless data, and compared F1-score for different input sizes (1024, 2048, 4096) between the three scenarios defined in Figure 4 to analyze how well `M2RF` performs. As depicted in Figure 9, the F1-scores for all scenarios significantly improve as we increase the input size. For scenario 1, the F1 score approaches as high as 86.65% with 1024 input size but increases nearly perfectly to a value of 94.99% to even an input size of 4096, showing that the trained model without overlap predicts very well between each signal class. In scenario 2, a similar trend is observed, where F1-scores increases from 82.11% at 1024 to reach the value of around 90.45% at 4096. The significant improvement is in scenario 3, where the F1-score improves from 69.71% at 1024 to 77.06% at 4096, demonstrating the `M2RF` robustness against overlapping signals. On the other hand, larger input sizes imply higher processing latency.

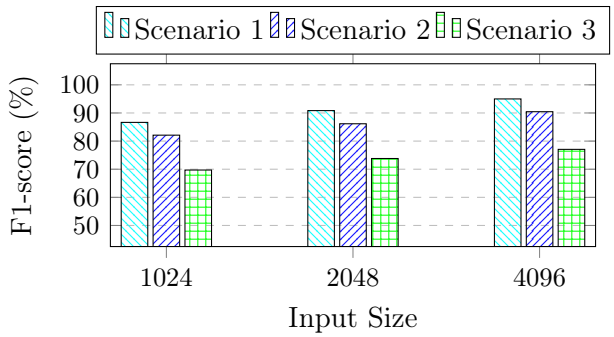


Figure 9: F1-score comparison for various input sizes across scenarios using wireless data.

Similarly, the IoU metric is a good indicator for segmentation accuracy. As expected, the IoU reaches 90.54% for an input size of 4096 in scenario 1, while scenario 2 slightly drops to 82.75%, due to the complexity introduced with random signal placements. In scenario 3 IoU decreases to 63.39% as signals now overlap and become harder to localize within the spectrum correctly. However, these results highlight the adaptability and robustness of the `M2RF` in varying RF environments. Detailed metrics for each input size across scenarios are provided in Appendix B.

6.1.2. Wired Mode

We further evaluate the robustness of M2RF in a wired setup, using the same input size of 4096 to ensure consistency with the wireless mode. Figure 10 shows that the wired setup achieved notably higher scores, which is expected given the absence of interference.

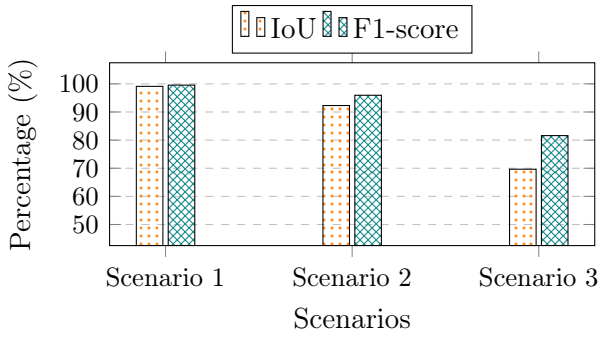


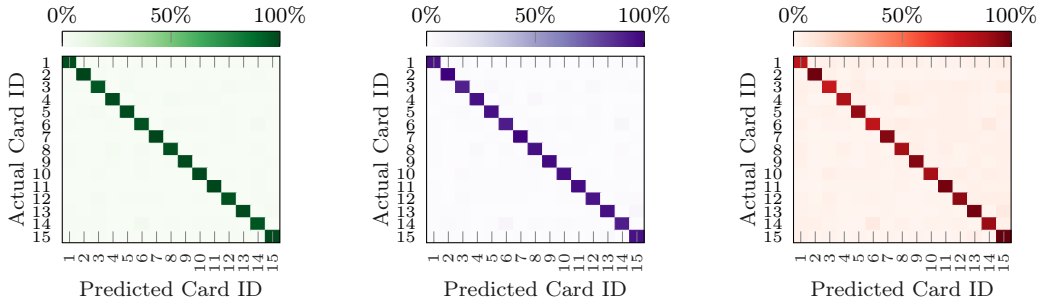
Figure 10: IoU and F1-score for different scenarios with an input size of 4096 using wired data.

In scenario 1 where the signals do not overlap, M2RF obtained a IoU of 99.12% and an F1-score of 99.56%, implying almost no misclassification. This shows the ability of M2RF to identify independent, non-overlapping signal sources. In a moderately challenging conditions (scenario 2) with partially observed signals, we obtain an IoU of 92.29% and F1-score of 95.95%, demonstrating the resilience of M2RF. The wired setup resulted in 69.66% IoU and 81.6% F1-score, even under scenario 3 – the worst-case conditions with both fully and partially overlapping signals.

6.1.3. Confusion Matrices Analysis

Deeper insights into the M2RF's classification accuracy were gained by analyzing the confusion matrices for each scenario, using wireless data with an input size of 4096. These matrices offer a detailed breakdown of true and false identifications across all 15 devices. Further analysis of confusion matrices for the wired mode is provided in Appendix C.

In scenario 1 (Figure 11a), the confusion matrix of non-overlapping signals provides that M2RF achieves a high classification accuracy, and most devices get an accuracy greater than 90%. For instance, card 1 achieves 94.1%, and card 2 reaches 97.2%, reflecting the system's effectiveness in distinguishing devices when signals are isolated and clearly separated.



(a) Scenario 1. (b) Scenario 2. (c) Scenario 3.

Figure 11: Confusion matrices for three scenarios in wireless mode.

Scenario 2, shown in Figure 11b, introduces an increased level of complexity with randomly placed signals within the bandwidth. This setup represents a more realistic real-world scenario, where signals can be partially or fully present into the observable bandwidth, generating signal glitches on duration and shifting. Although M2RF achieves high performance for multiple devices (e.g., card 2 at 94.8%) there is a small drop in accuracy for others, such as card 3 (85.2%). The variation in this performance illustrates the M2RF's robustness to non-ideal conditions where the signals are not perfectly isolated, thus making the classification problem harder and M2RF learning to generalize across unpredictable scenarios.

Scenario 3 is shown in Figure 11c which is full or partial overlapping of signals. At this point, accuracy rates further decrease, where card 1 gets only 68.1% and card 3 gets only 65.9%. Such a decrease is intuitive as if the signals overlap in the same frequency range, finding the characteristics of signal would be difficult. These obstacles notwithstanding, the M2RF still performs well in accurately classifying most devices and shows it could operate in congested RF environments where overlaps are frequently observed.

The distinguishing of devices in scenario 3 by the M2RF also suggests its ability to test against jamming attack and even detect it. In those cases, we may actually have intentional disruptions within the spectrum represented by overlapping signals. The M2RF's strong localization of these overlapping signals would enable the development of a target countermeasure where the M2RF can detect and suppress jamming in real-time with minimal impact to other neighboring communications. This capability adds a critical dimension to the M2RF 's utility in dynamic, interference-prone environments, where

prompt and accurate signal identification is essential for maintaining security and integrity.

6.2. Scalability and Generalization

In this section, we examine the scalability and generalizability of M2RF across different bandwidths and channel widths. The DNN was initially trained with a 50 MHz bandwidth and 20 MHz-wide signals and then tested under varying bandwidths to assess whether the hardware imperfections it leverages remain consistent. These evaluations provide insights into M2RF's adaptability to changing spectral conditions. Key cases are illustrated in Figures 12 and 13, using data collected on a different day with unseen signal conditions, which typically vary due to environmental factors.

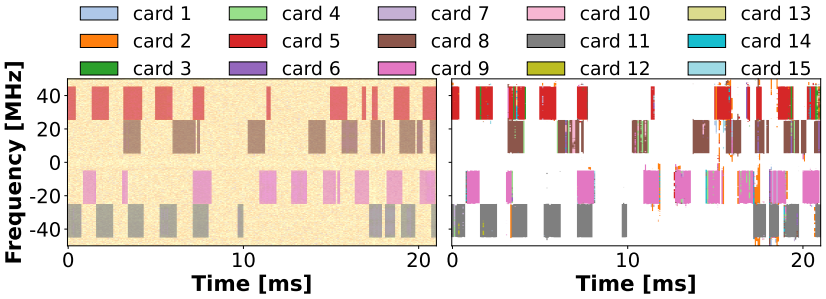


Figure 12: 100 MHz bandwidth, 4 signals 20 MHz wide. (Left) Ground Truth, (Right) Model Prediction.

In the first test, the M2RF system was evaluated with a 100 MHz bandwidth containing four devices, each occupying 20 MHz. As shown in Figure 12, M2RF successfully distinguishes between these signals, achieving an F1-score of 90.22% and an IoU of 82.32%. Although performance shows a slight reduction from the 50 MHz baseline, this decrease can be attributed to the increased spectral complexity and channel noise. Nonetheless, M2RF demonstrates high scalability, maintaining effective detection and localization of multiple signals within the broader bandwidth without requiring retraining.

In a second test, M2RF was evaluated with two signals that were 40 MHz wide in a 100 MHz bandwidth. M2RF achieved F1-score of 87% and an IoU of 77.29%. The broader signal, along with its associated noise, added complexity to this configuration. Although the accuracy drops a little, M2RF correctly identifies and localizes every signal, as shown in Figure 13. The above result

illustrates the robustness of M2RF to changes in channel width and signal structure. Most importantly, the system performed a correct classification without retraining, making our M2RF approach even more robust.

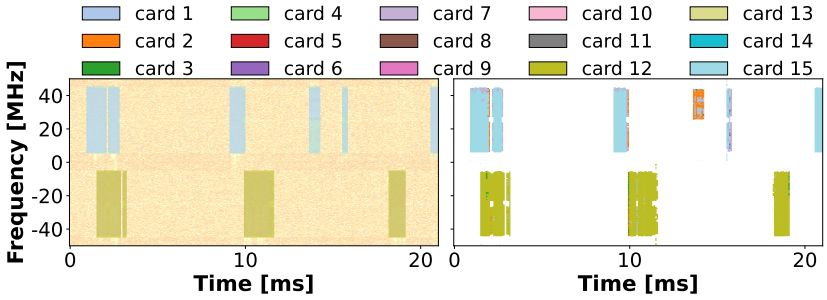


Figure 13: 100 MHz Bandwidth, 2 signals 40 MHz wide. (Left) Ground Truth, (Right) Model Prediction.

This establishes that, although we see some degradation in performance at higher bandwidth demand cases, the fundamental approach remains resilient. The signal-recognition capabilities of the M2RF, including its scalability and robustness to changing channel width and signal overlap, rely on hardware-induced imperfections that vary consistently across devices. These frequency-specific distortions as a result of hardware imperfection persist across different bandwidths allowing for reliable classification and consequently localization of the signal, yielding a robust fingerprinting approach that can work well with both diverse and dynamic RF environments. Such scalability and flexibility is essential for practical applications where spectral conditions will change, and the system has to function properly across a wide range of bandwidths and interference levels.

6.3. System Defense under Malicious Activity

We rigorously evaluated M2RF's defense capabilities against both *informed* and *uninformed adversaries* within Wi-Fi networks, as well as interference from other wireless technologies in congested spectrum environments.

In the informed adversary scenario, the attacker has full knowledge of the authentication method and ML model used, as well as access to identical hardware as the legitimate devices. On the other hand, the uninformed adversary represents a more realistic scenario in which the attacker uses potentially different hardware versions unknown to the system at training

time. To evaluate the robustness of M2RF in crowded spectrum environments, we also performed inference detection with respect to signals from other non-Wi-Fi technologies like BLE, LTE and Zigbee. This multi-technology challenge tests the capability of M2RF to detect unauthorized transmissions in various signal types, reflecting real-world conditions in densely populated RF environments.

Table 1: M2RF Performance under Attacks in Wi-Fi Networks.

Attack Type	User	P (%)	R (%)	F1 (%)	MR (%)	FAR (%)
Uninformed	**Authorized**	98.29	87.66	92.67	12.34	1.52
	Malicious	88.90	98.48	93.44	1.52	12.34
Informed	**Authorized**	95.71	87.51	91.43	12.49	3.91
	Malicious	88.51	96.09	92.14	3.91	12.49
					Overall Accuracy: 92.44%	

We used approximately 100 000 samples, evenly split between authorized and malicious transmissions, to evaluate system performance for each attack type. As shown in Table 1, M2RF consistently performed well in distinguishing authorized from malicious signals using a TV-based threshold of $TV = 8$. In uninformed attack scenario, the system reached an F1-score of 92.67% for authorized devices and 93.44% for malicious devices, with a mean Miss Rate (MR) of 2.72% across both attack types, demonstrating reliable detection of unauthorized signals even when attackers use different hardware. For informed attack, where the attacker's hardware matches that of the authorized devices, M2RF maintained a high F1-score of 91.43% for authorized devices and 92.14% for malicious devices, showing robustness against highly sophisticated adversaries, without requiring retraining on specific attack data.

Table 2: M2RF Performance under Interference from Other Technologies.

User	P (%)	R (%)	F1 (%)	MR (%)	FAR (%)
Authorized	78.23	87.47	82.59	12.53	24.45
Interference	85.72	75.55	80.31	24.45	12.53
				Overall Accuracy: 81.52%	

Table 2 highlights M2RF 's ability to differentiate authorized Wi-Fi signals from non-Wi-Fi signals in a multi-technology setting, achieving an over-

all accuracy of 81.52%. This multi-technology evaluation confirms M2RF's adaptability in congested spectrum environments, such as the 2.4 GHz band, effectively distinguishing authorized Wi-Fi devices from other signals without requiring retraining on these technologies. Using unique radio fingerprints, M2RF maintains high performance under various RF conditions, demonstrating its ability to secure spectrum management and defense against unauthorized access in high-traffic, multi-technology scenarios.

6.4. Energy-Latency Trade-off for Different Input Sizes on an Edge Device

In our experiments on energy-latency trade-offs, we measured the performance of different input sizes—1024, 2048, and 4096—on both GPU and CPU through Jetson Orin Nano device. The primary metrics were mean inference time (MIT) and mean energy consumption (MEC), which are crucial for evaluating real-time M2RF efficiency. Despite the CPU having lower mean power consumption (MPC) per millisecond (ms), the significantly longer MIT leads to much higher total energy consumption compared to the GPU. For example, at an input size of 4096, the inference time on the GPU was 20.87 ms compared to 391.62 ms on the CPU, demonstrating the GPU's substantial speed advantage for time-critical applications.

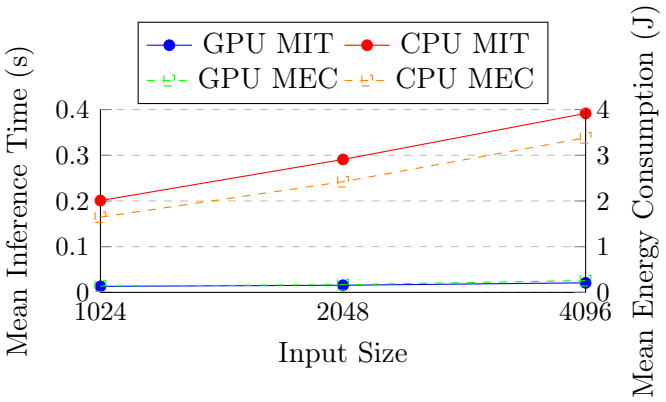


Figure 14: Energy-latency trade-off for different input sizes.

MEC was calculated using the formula:

$$\text{MEC (mJ)} = \text{MIT (ms)} \times \text{MPC (mW)}. \tag{9}$$

This formula highlights that, although the CPU consumes less power at 8.63 milliwatts (mW) for an input size of 4096, its much longer inference time

results in a significantly higher energy consumption of 3381.50 millijoules (mJ), compared to the GPU's energy consumption of 257.95 mJ for the same input size. Thus, the GPU's slightly higher power draw (12.36 mW) is more than offset by its superior processing speed, making it far more energy-efficient in total energy use. Figure 14 further illustrates this trade-off, showing how both MEC and MIT increase with input size, but at a much steeper rate for the CPU than for the GPU.

7. Related Work

Spectrum Sensing. Early spectrum sensing research focused on binary classification to detect whether the frequency band is in use or not [34, 35]. These approaches fall short of supporting sophisticated spectrum policies where devices may require different levels of priority in spectrum access. Recent work proposed to jointly classify multiple signals in the spectrum based on wireless technologies [24] and modulation types [27]. However, [24, 27] fail to provide finer-grained identification of devices.

Radio Fingerprinting. Initial fingerprinting approaches, based on hand-crafted features such as phase and amplitude noise to characterize device fingerprints, performed well in controlled environments but suffered from overfitting issues in more complex settings [36, 25, 26]. CNNs and RNNs solve this issue by performing feature extraction over raw I/Q data automatically [37, 38]. However, existing radio fingerprinting methods has a series of limitations making them unpractical in spectrum sharing scenarios. For example, the authors [21] had over 95% accuracy with 16 devices but classified only one signal at time while it was assumed the center frequencies were known. In contrast, our method based on U-Net dynamically classifies transmissions with different, even overlapping center frequencies and brings spectrum localization by segmenting in frequency bins. In addition, traditional models often struggle to generalize effectively between controlled and real-world environments, leading to frequent model failure and the need for retraining [39, 40]. In contrast, our work demonstrates scalability and generalization capability in different scenarios without retraining.

8. Conclusions

This paper demonstrates that radio fingerprinting in multi-device, multi-band environments requires effective management of overlapping signals, di-

verse bandwidths, and inherent hardware imperfections. We present a U-Net-based model for scalable and robust semantic segmentation of RF signals to tackle these challenges. We obtain the following results: (i) the combined loss function significantly enhances performance, achieving an IoU of 90.54% and an F1-score of up to 94.99% for non-overlapping signals, and an F1-score of up to 77.06% in overlapping scenario; (ii) the approach maintains reliable detection across varied bandwidths, achieving an F1-score of 90.22% in scenarios with a 100 MHz bandwidth; (iii) the system effectively detects malicious Wi-Fi activities with an overall accuracy of 92.44% and a mean MR of 2.72%, without prior exposure to attack data, and successfully differentiates authorized Wi-Fi devices from non-Wi-Fi technologies with an accuracy of 81.52%, even in congested spectrum environments; and (iv) our system achieves an mean inference time of 20.87 ms and a mean energy consumption of 257.95 mJ on edge device. These results confirm that the `M2RF` has the potential to provide an efficient, scalable solution for IoT security applications with real-time constraints.

References

[1] Federal Communications Commission (FCC), "Spectrum Crunch," https://www.fcc.gov/general/spectrum-crunch.

[2] L. Zhang, M. Xiao, G. Wu, M. Alam, Y.-C. Liang, and S. Li, "A Survey of Advanced Techniques for Spectrum Sharing in 5G Networks," *IEEE Wireless Communications*, vol. 24, no. 5, pp. 44–51, 2017.

[3] F. Hu, B. Chen, and K. Zhu, "Full Spectrum Sharing in Cognitive Radio Networks Toward 5G: A Survey," *IEEE Access*, vol. 6, pp. 15 754–15 776, 2018.

[4] H. Shokri-Ghadikolaei, F. Boccardi, C. Fischione, G. Fodor, and M. Zorzi, "Spectrum Sharing in mmWave Cellular Networks via Cell Association, Coordination, and Beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 11, pp. 2902–2917, 2016.

[5] L. Lv, J. Chen, Q. Ni, Z. Ding, and H. Jiang, "Cognitive Non-Orthogonal Multiple Access with Cooperative Relaying: A New Wireless Frontier for 5G Spectrum Sharing," *IEEE Communications Magazine*, vol. 56, no. 4, pp. 188–195, 2018.

[6] Linda Hardesty (Fierce Wireless), "What is a CBRS spectrum access system?" https://www.fiercewireless.com/private-wireless/what-a-cbrs-spectrum-access-system, 2020.

[7] L. Zhang, Y.-C. Liang, and M. Xiao, "Spectrum sharing for Internet of Things: A Survey," *IEEE Wireless Communications*, vol. 26, no. 3, pp. 132–139, 2018.

[8] S. Kwon and H.-K. Choi, "Evolution of Wi-Fi Protected Access: Security Challenges," *IEEE Consumer Electronics Magazine*, vol. 10, no. 1, pp. 74–81, 2021.

[9] A. Koutsos, "The 5G-AKA Authentication Protocol Privacy," in *Proceedings of IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2019, pp. 464–479.

[10] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A Survey on Wireless Security: Technical Challenges, Recent Advances, and Future Trends," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[11] L. Baldesi, F. Restuccia, and T. Melodia, "ChARM: NextG Spectrum Sharing Through Data-Driven Real-Time O-RAN Dynamic Control," in *Proceedings of IEEE International Conference on Computer Communications (INFOCOM)*. IEEE, 2022, pp. 240–249.

[12] T. D. Vo-Huu, T. D. Vo-Huu, and G. Noubir, "Fingerprinting Wi-Fi Devices Using Software Defined Radios," in *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. ACM, 2016, pp. 3–14.

[13] Y. Xu, M. Liu, L. Peng, J. Zhang, and Y. Zheng, "Colluding rf fingerprint impersonation attack based on generative adversarial network," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 3220–3225.

[14] M. Alhaidary and S. M. M. Rahman, "Security vulnerability analysis and corresponding mitigation for password-based authentication using an offline personal authentication device," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE, 2016, pp. 842–849.

[15] Z. Yao, Y. Peng, Y. Wang, C. Xu, J. Wang, Y. Lin, and G. Gui, "A novel radio frequency fingerprint concealment method based on iq imbalance compensation and digital pre-distortion," *IEEE Transactions on Information Forensics and Security*, 2024.

[16] M. R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, "Calculation of the performance of communication systems from measured oscillator phase noise," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 5, pp. 1553–1565, 2014.

[17] D. Zanetti, S. Capkun, and B. Danev, "Types and origins of fingerprints," *Digital Fingerprinting*, pp. 5–29, 2016.

[18] I. Alla, S. Yahia, V. Loscri, and H. Eldeeb, "Robust device authentication in multi-node networks: Ml-assisted hybrid pla exploiting hardware impairments," in *2024 Annual Computer Security Applications Conference (ACSAC)*, 2024, pp. 1172–1185.

[19] J. Zhang, G. Shen, W. Saad, and K. Chowdhury, "Radio frequency fingerprint identification for device authentication in the internet of things," *IEEE Communications Magazine*, 2023.

[20] T. Jian, B. C. Rendon, A. Gritsenko, J. Dy, K. Chowdhury, and S. Ioannidis, "Mac id spoofing-resistant radio fingerprinting," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2019, pp. 1–5.

[21] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: Optimized Radio Classification through Convolutional Neural Networks," in *Proc. of IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2019, pp. 370–378.

[22] A. Al-Shawabka, F. Restuccia, S. D'Oro, T. Jian, B. C. Rendon, N. Soltani, J. Dy, K. Chowdhury, S. Ioannidis, and T. Melodia, "Exposing the Fingerprint: Dissecting the Impact of the Wireless Channel on Radio Fingerprinting," *Proc. of IEEE Conference on Computer Communications (INFOCOM)*, 2020.

[23] P. Angueira, I. Val, J. Montalban, Ó. Seijo, E. Iradier, P. S. Fontaneda, L. Fanari, and A. Arriola, "A survey of physical layer techniques for se-

cure wireless communications in industry," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 2, pp. 810–838, 2022.

[24] D. Uvaydov, M. Zhang, C. P. Robinson, S. D'Oro, T. Melodia, and F. Restuccia, "Stitching the Spectrum: Semantic Spectrum Segmentation with Wideband Signal Stitching," *Proceedings of IEEE Conference on Computer Communications (INFOCOM)*, 2024.

[25] J. Zhang, R. Woods, M. Sandell, M. Valkama, A. Marshall, and J. Cavallaro, "Radio frequency fingerprint identification for narrowband systems, modelling and classification," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3974–3987, 2021.

[26] N. Soltanieh, Y. Norouzi, Y. Yang, and N. C. Karmakar, "A review of radio frequency fingerprinting techniques," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 3, pp. 222–233, 2020.

[27] A. Mittal, M. Zhang, T. Gourousis, Z. Zhang, Y. Fei, M. Onabajo, F. Restuccia, and A. Shrivastava, "Sub-6 ghz energy detection-based fast on-chip analog spectrum sensing with learning-driven signal classification," *IEEE Internet of Things Journal*, 2024.

[28] A. Al-Shawabka, P. Pietraski, S. B. Pattar, F. Restuccia, and T. Melodia, "Deeplora: Fingerprinting lora devices at scale through deep learning and data augmentation," in *Proceedings of the Twenty-second International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing*, 2021, pp. 251–260.

[29] Y. Xiao, Y. He, X. Zhang, Q. Wang, R. Xie, K. Sun, K. Xu, and Q. Li, "From hardware fingerprint to access token: Enhancing the authentication on iot devices," *arXiv preprint arXiv:2403.15271*, 2024.

[30] R. Azad, E. K. Aghdam, A. Rauland, Y. Jia, A. H. Avval, A. Bozorgpour, S. Karimijafarbigloo, J. P. Cohen, E. Adeli, and D. Merhof, "Medical image segmentation review: The success of u-net," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

[31] D. W. Romero, D. M. Knigge, A. Gu, E. J. Bekkers, E. Gavves, J. M. Tomczak, and M. Hoogendoorn, "Towards a general purpose cnn for long range dependencies in n d," *arXiv preprint arXiv:2206.03398*, 2022.

[32] Z. Wang, F. Ma, P. Ji, and C. Fu, "Image denoising based on an improved wavelet threshold and total variation model," in *International Conference on Intelligent Computing*. Springer, 2024, pp. 142–154.

[33] C. Donnat, O. Klopp, and N. Verzelen, "One-bit total variation denoising over networks with applications to partially observed epidemics," *arXiv preprint arXiv:2405.00619*, 2024.

[34] D. Uvaydov, S. D'Oro, F. Restuccia, and T. Melodia, "DeepSense: Fast wideband spectrum sensing through real-time in-the-loop deep learning," in *Proc. of IEEE Intl. Conf. on Computer Communications (INFOCOM)*, Vancouver, BC, Canada, May 2021.

[35] C. Liu, J. Wang, X. Liu, and Y.-C. Liang, "Deep CM-CNN for Spectrum Sensing in Cognitive Radio," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 10, pp. 2306–2321, 2019.

[36] X. Wang, Y. Zhang, H. Zhang, X. Wei, and G. Wang, "Identification and authentication for wireless transmission security based on rf-dna fingerprint," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 230, 2019.

[37] K. Sankhe, M. Belgiovine, F. Zhou, L. Angioloni, F. Restuccia, S. D'Oro, T. Melodia, S. Ioannidis, and K. Chowdhury, "No Radio Left Behind: Radio Fingerprinting Through Deep Learning of Physical-Layer Hardware Impairments," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 1, pp. 165–178, 2019.

[38] T. Yang, S. Hu, W. Wu, L. Niu, D. Lin, and J. Song, "Conventional neural network-based radio frequency fingerprint identification using raw i/q data," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, p. 8681599, 2022.

[39] A. Elmaghbub and B. Hamdaoui, "No blind spots: On the resiliency of device fingerprints to hardware warm-up through sequential transfer learning," in *Proceedings of the 17th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2024, pp. 134–144.

[40] A. Saeif, S. Savio, and O. Gabriele, "The day-after-tomorrow: On the performance of radio fingerprinting over time," in *Proceedings of the 39th Annual Computer Security Applications Conference*, 2023, pp. 439–450.

[41] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*. Springer, 2017, pp. 240–248.

[42] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *International symposium on visual computing*. Springer, 2016, pp. 234–244.

[43] G. Zhao, W. Yang, X. Ren, L. Li, Y. Wu, and X. Sun, "Well-classified examples are underestimated in classification with deep neural networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 8, 2022, pp. 9180–9189.

[44] Y. Huang, J. Qi, X. Wang, and Z. Lin, "Asymmetric polynomial loss for multi-label classification," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[45] T.-Y. Ross and G. Dollár, "Focal loss for dense object detection," in *proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2980–2988.

Appendix A. Loss Functions and Optimization

We have investigated a number of loss functions. Here, the loss functions are part of *local-level* and *region-level* metrics. Local-level loss measures accuracy at each frequency bin within the raw I/Q data, capturing fine-grained variations essential for distinguishing signal characteristics. Region-level loss, on the other hand, considers larger segments within the data, promoting consistency across continuous sections and enhancing detection of broader patterns, such as distinct signal regions or transmission boundaries. We summarize them as follows:

Dice Loss (DiL): DiL [41] is a region-based metric used to assess the similarity between predicted labels and ground truth, which is derived from the

Dice coefficient, a widely used measure of similarity. It is defined as:

$$\text{DiL} = 1 - \frac{2 \times \sum_{i=1}^{n} y_i \times \hat{y}_i}{\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \hat{y}_i + \epsilon}, \tag{A.1}$$

where y_i and $\hat{y}_i$ are the ground truth and predicted values, respectively, and ϵ is a small constant to avoid division by zero. DiL focuses on maximizing the overlap between the predicted and ground truth masks, making it suitable for tasks where precise segmentation is critical.

Intersection over Union Loss (IoUL): The IoUL [42] is another region-based loss function that measures the overlap between the predicted and ground truth. It is defined as:

$$\text{IoUL} = 1 - \frac{\sum_{i=1}^{n} y_i \times \hat{y}_i}{\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \hat{y}_i - \sum_{i=1}^{n} y_i \times \hat{y}_i + \epsilon}. \tag{A.2}$$

This loss is particularly useful in cases where there is significant class imbalance, as it penalizes both false positives and false negatives.

Cross-Entropy Loss (CEL): The CEL [43] is a loss function for classification tasks and is defined as

$$\text{CEL} = -\sum_{i=1}^{n} y_i \log(\hat{y}_i). \tag{A.3}$$

This loss provides a local-level accuracy of semantic segmentation which is used as a baseline in our experiments and is combined with other loss functions to improve the model performance.

Binary Cross-Entropy Loss (BCEL): BCE [44] is widely used for binary classification tasks and is similar to CEL but adapted for binary output. It is defined as:

$$\text{BCE} = -\sum_{i=1}^{n} y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i). \tag{A.4}$$

BCE is effective for multi-label segmentation tasks where each frequency bin can belong to more than one class, making it particularly relevant for radio fingerprinting scenarios with overlapping signals.

Focal Loss (FL): FL [45] is designed to address the issue of class imbalance by down-weighting the contribution of easy examples during training and focusing on hard-to-classify examples. It is defined as:

$$\text{FL} = -\alpha(1 - \hat{y}_i)^{\gamma} \log(\hat{y}_i), \tag{A.5}$$

where α is a balancing factor and γ is the focusing parameter. FL is particularly effective in improving the performance of the DNN on underrepresented classes.

Combined Loss (CL): To leverage the strengths of different loss functions, we have implemented a CL function that integrates both local-level and region-level losses, defined as:

$$\mathrm{CL} = \beta \times \mathrm{CEL} + (1 - \beta) \times \mathrm{IoUL}, \tag{A.6}$$

where β is a weighting factor that balances the contribution of each component. As such, we can optimize both fine-grained accuracy and overall region consistency.

Evaluation of Loss Functions. We discuss a comparative analysis of various loss functions applied to the radio fingerprinting task of non-overlapping signals. Table A.3 shows that the CL function achieves the highest performance across all metrics, with an Intersection over Union (IoU) of 77.37%. This highlights the effectiveness of combining local-level and region-level loss functions.

Table A.3: Performance Metrics for Different Loss Functions in Scenario 1 with an Input Size of 1024.

Loss Function	**IoU (%)**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**
BCEL	77.08	86.48	86.50	86.49
CEL	77.04	86.44	86.43	86.43
CL	**77.37**	**86.66**	**86.65**	**86.65**
DiL	60.34	83.13	69.75	75.85
FL	76.07	85.79	85.74	85.76
IoUL	59.84	83.11	68.81	75.29

DiL and IoUL show significantly lower IoU values (around 60%). While FL performed better than DiL and IoUL, it does not perform as BCEL, CEL and especially CL. As such, we chose CL for our next experiments.

Appendix B. Detailed Metrics for Each Modality

In Table B.4, we analyze the performance of `M2RF` in detail over various input sizes (1024, 2048, and 4096) with respect to the wireless mode for three different scenarios. We quantify the aspect of device-specific characterization

from each input size, especially in a wireless scenario, when signals are likely to be received and disturbed due to environmental interference. The constant high performance of the 4096 input size shows its appropriateness as a benchmark. Table B.5 presents the corresponding results in wired mode, where the improved signal quality and reduced interference further enhance performance, establishing an "ideal" scenario for comparison.

Table B.4: Metrics for Different Input Sizes and Scenarios in Wireless Mode.

Input Size	**Scenario**	**Precision (%)**	**Recall (%)**	**F1-Score (%)**	**IoU (%)**
1024	1	86.66	86.65	86.65	77.37
	2	82.21	82.06	82.11	70.68
	3	76.28	64.35	69.71	54.99
2048	1	90.86	90.83	90.84	83.50
	2	86.24	86.10	86.16	76.12
	3	80.87	68.09	73.81	59.45
4096	1	95.02	94.97	94.99	90.54
	2	90.52	90.39	90.45	82.75
	3	84.56	70.97	77.06	63.39

Input Size 1024: When an input has a size of 1024, the metrics show that while the `M2RF` captures features at the device level, it does not perform as well with the lower IoU values, as in scenario 3, where it only achieves an IoU of 54.99%. The small input size limits the signal information that can be used for model processing. This hinders the handling of more complicated signal scenarios with heavy interference or overlaps. Precision, recall, and F1-scores also decrease across scenarios. For example, the F1-score is 86.65% in scenario 1 compared to a much lower F1-score of 69.71% in scenario 3. The aforementioned limitations, however, imply that an input size of 1024 is poor for robust radio fingerprinting within dynamic wireless environments.

Input Size 2048: Increasing the input size to 2048 gives significant improvements on all metrics. In scenario 1, the F1-score is raised to 90.84%, and IoU increases to 83.50%, which demonstrates that it can tell apart device characteristics more easily with a bigger data sample through `M2RF`. On the other hand, in scenario 3, where signals frequently overlap, we observe that the IoU and F1-score are still low at 59.45% and 73.81%, respectively. Though we set the `M2RF` to this input size of 2048 to capture finer intricacies within the signal, it is clear from our results that larger values will be needed in order to obtain strong accuracy under complicated wireless conditions.

Input Size 4096: The M2RF obtains its best performance in terms of all the metrics when having an input size of 4096. Scenario 1 achieves a precision of 95.02% (F1-score of 94.99%, IoU of 90.54%), which suggests that the M2RF is capable of leveraging the broader input size to identify device-specific hardware imperfections in the RF signals. In scenario 3, where it is logical to believe the overlapping signals will challenge the model, we still achieve an F1-score of 77.06% and an IoU of 63.39%, which is a far better performance than using any smaller input size. This means that an input size of 4096 is better for accounting for full radio fingerprinting details, even under wireless mode scenarios in which signal interference occurs and where the transmissions can be more accurately localized within the spectrum.

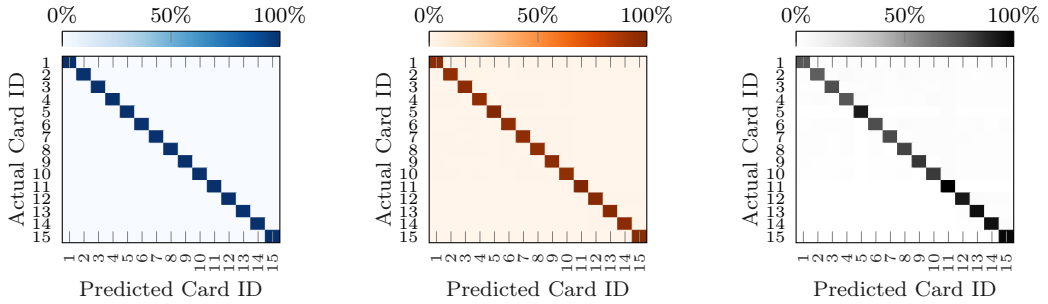
Based on these results, we set 4096 as our input size baseline which strikes a balance between accuracy and generalizability for radio fingerprinting problems in wireless and wired mode.

Table B.5: Metrics for 4096 Input Size and Scenarios in Wired Mode.

Scenario	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)
1	99.56	99.55	99.56	99.12
2	96.01	95.90	95.95	92.29
3	89.41	75.24	81.60	69.66

Benchmark Results in Wired Mode: As we can see in Table B.5, these metrics also indicate that a wired mode with an input size of 4096 performs better. In scenario 1, the precision, recall, and F1-score were all greater than 99%, with an IoU value of 99.12%, suggesting almost perfect identification accuracy. In scenario 2, the F1-score and IoU are both high (95.95% and 92.29%, respectively), which shows that even in optimal conditions, the M2RF accommodates a small portion of signal belonging to a structure within its part of the observable spectrum. When things get difficult in scenario 3 with overlapping signals, the M2RF achieves an F1-score of 81.60% and an IoU of 69.66%, also higher than the its wireless counterpart. The effectiveness of the radio fingerprinting under better channel conditions with sufficient SNR margin and less interference is reflected in this result, suggesting that 4096 input size is preferable for high-accuracy device authentication.

In general, the comparison of wireless and wired modes justifies the choice of an input size of 4096 for radio fingerprinting purposes. Although performance in the wireless mode depends on signal complexity, the results for the



(a) Scenario 1. (b) Scenario 2. (c) Scenario 3.

Figure C.15: Confusion matrices for three scenarios in wired mode.

wired mode provide a consistent baseline, validating that this input size is capable of capturing intricate device-specific imperfections even under difficult conditions.

Appendix C. Additional Results for Wired Mode

This part provides an overview of the performance for three scenarios of wired mode under an input size of 4096. This wired connection means that the signals are much cleaner and less susceptible to interference, enabling `M2RF` to attain higher accuracy. The confusion matrix for each of the scenarios demonstrates how well the system performs under various signal conditions. **Scenario 1: Non-overlapping Signals.** In the first wired scenario, we have excellent performance, also seen in the confusion matrix in Figure C.15a, as all card IDs provide accuracy values above 99%. Specifically, card 1 gives us an accuracy of 99.5%, card 3 provides 99.6%, and card 5 reaches 99.8%. Since the signals are non-overlapping in this use case, it is easy for the model to identify each device separately. This scenario underscores the effectiveness of radio fingerprinting under ideal conditions where each device's unique signal characteristics are isolated, thereby minimizing potential misclassification. The high performance in this case exemplifies how hardware imperfections can be effectively leveraged for device identification in a controlled setting. **Scenario 2: Partial Observation with Random Frequency Centers.** Scenario 2, illustrated in Figure C.15b, does not degrade accuracy greatly but will degrade it bit by bit more than was the case in scenario 1 due to more complex signals. Note that, for example, cards 1 and 3, we

can see the accuracies are 95.3% and 95.7%, but card 5 has a higher accuracy of 98.6%. Partial signals and random frequency centers make it more challenging to separate device-specific features from the data. Nevertheless, the M2RF performs reliably well, being able to contend with more complicated signal patterns but still uniquely separating devices due to hardware imperfections inherent in each. These results suggest the robustness of M2RF to small changes in signal characteristics, a necessary trait for dynamic RF environments where signal properties experience moderate fluctuations.

Scenario 3: Partial/Full Signals Overlapping. Figure C.15c shows the confusion matrix generated for scenario 3, which again highlights how challenging it is when signals are fully or partially overlapping. The accuracy is significantly lower in comparison to the previous scenarios. For instance, cards 1 and 3 get lower accuracies of 71.1% and 74% respectively, while good performances are attained on cards 9 through 15 with accuracies above 80%. Because several signals occupy the same or adjacent frequency bins, it becomes challenging for the M2RF to discriminate between them and they are more often misclassified. However, improved signal quality in the wired mode helps mitigate some interference, enabling M2RF to maintain reasonable performance despite the challenging conditions. In this scenario, the heavy overlap among signals reveals the limitations of radio fingerprinting, while also demonstrating that M2RF can still extract discriminative features even in challenging RF environments.