DEEP LEARNING FOR RF FINGERPRINTING

From Statistical Enhancement to Scalable Architectures for Wireless Security

A Thesis Submitted to the Division of Graduate Studies
of the Royal Military College of Canada
by

Nordine Quadar, P.Eng

In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

March, 2026

To my late father, and to my family and mentors, for believing in me.

Acknowledgements

The completion of this doctoral thesis would not have been possible without the guidance, support, and encouragement of many individuals and organizations.

First and foremost, I would like to express my deepest gratitude to my supervisor, Professor Abdellah Chehri at the Royal Military College of Canada, for his invaluable guidance, patience, and unwavering support throughout this research journey. His insightful feedback and encouragement have been instrumental in shaping this work and my development as a researcher.

I am profoundly grateful to Benoit Debaque and the entire team at Thales Defense & Security for their collaboration and for providing the industrial perspective that grounded this research in practical applications. This research was conducted through a three-year Mitacs Accelerate fellowship (2022-2025) in partnership with Thales, whose financial support and technical expertise were essential to the success of this work.

I would like to thank the members of my thesis examination committee, Profs. Mariam El Mezouar (RMC), François Chan (RMC) and Ashraf Matrawy (Carleton University) for their time, expertise, and constructive feedback that improved the quality of this dissertation.

Finally, and most importantly, I owe my deepest gratitude to my family. To my wife, for her endless love, patience, full support, and for caring for our children during countless late nights and weekends of research. Your sacrifices made this possible. To my children, for reminding me daily what truly matters and for insisting I take breaks to play, those moments kept me grounded and sane. To my mother, for her constant prayers and encouragement. And to my entire family, for their love and belief in me.

Abstract

RF fingerprinting offers a way to authenticate wireless devices based on their unique hardware characteristics at the physical layer, without relying on cryptographic credentials. However, after almost two decades of research, practical deployment remains limited. Our analysis of over 40 recent studies reveals a consistent pattern: systems that achieve above 95% accuracy in the laboratory suffer 20–70% accuracy drops under realistic conditions, scale poorly beyond 20–30 devices, and cannot detect unknown devices. In this thesis, we address these limitations through four complementary approaches.

We first develop a statistical feature enhancement framework that combines magnitude, phase, power, entropy, and spectral features to reach 99.6% same-day accuracy while improving cross-day performance from 37.5% to 52.1%. We then introduce cyclostationary feature engineering that captures periodic statistical properties in communication signals, reaching 85.9% cross-transmission accuracy with 38% memory reduction through progressive learning. To handle mobility, we design a CNN-LSTM-Attention architecture that maintains 85.8% accuracy under 100Hz Doppler shift conditions typical of UAV operations. Finally, our MDAE-Transformer architecture shifts from classification to embedding-based identification, reaching 92.9% cross-day accuracy and 75.2% at the 80-device scale with near-constant inference latency, while providing architectural support for open-set recognition through reconstruction-based anomaly detection.

We evaluate all approaches across six datasets spanning WiFi, LoRa, UAV controllers, and Bluetooth protocols. Our comparative analysis shows that no single approach dominates all scenarios; method selection depends on operational requirements. Statistical enhancement suits controlled environments, cyclostationary features handle transmission parameter variations, temporal modeling addresses mobility, and the MDAE-Transformer supports large-scale deployment with better temporal adaptation.

This work moves RF fingerprinting closer to operational deployment and provides foundations for physical-layer security in next-generation wireless

networks, particularly for UAV communications and IoT security where cryptographic approaches alone prove insufficient.

Keywords: RF fingerprinting, physical-layer security, deep learning, UAV security, IoT authentication, transformer architectures, cyclostationary analysis, temporal modeling

Résumé

L'identification par empreinte radiofréquence (RF) permet d'authentifier les dispositifs sans fil à partir de leurs caractéristiques matérielles uniques au niveau de la couche physique, sans dépendre des mécanismes cryptographiques. Cependant, après près de deux décennies de recherche, le déploiement pratique reste limité. Notre analyse de plus de 40 études récentes révèle un constat récurrent : les systèmes qui atteignent plus de 95% de précision en laboratoire subissent des chutes de 20 à 70% en conditions réalistes, passent mal à l'échelle au-delà de 20–30 dispositifs, et ne peuvent pas détecter les dispositifs inconnus. Dans cette thèse, nous abordons ces limitations à travers quatre approches complémentaires.

Nous développons d'abord un cadre d'amélioration statistique qui combine des caractéristiques de magnitude, phase, puissance, entropie et spectre pour atteindre 99,6% de précision le même jour tout en améliorant les performances inter-journalières de 37,5% à 52,1%. Nous introduisons ensuite l'ingénierie de caractéristiques cyclostationnaires qui capture les propriétés statistiques périodiques des signaux de communication, atteignant 85,9% de précision inter-transmission avec une réduction de mémoire de 38% grâce à l'apprentissage progressif. Pour répondre aux défis de mobilité, nous concevons une architecture CNN-LSTM-Attention qui maintient 85,8% de précision dans des conditions de décalage Doppler de 100Hz typiques des opérations de drones. Enfin, notre architecture MDAE-Transformer passe de la classification à l'identification par représentations vectorielles, atteignant 92,9% de précision inter-journalière et 75,2% à l'échelle de 80 dispositifs avec une latence d'inférence quasi constante, tout en offrant un support architectural pour la reconnaissance en ensemble ouvert par détection d'anomalies basée sur la reconstruction.

Nous évaluons toutes les approches sur six ensembles de données couvrant les protocoles WiFi, LoRa, contrôleurs de drones et Bluetooth. Notre analyse comparative montre qu'aucune approche unique ne domine tous les scénarios; le choix de la méthode dépend des exigences opérationnelles. L'amélioration

statistique convient aux environnements contrôlés, les caractéristiques cyclostationnaires gèrent les variations de paramètres de transmission, la modélisation temporelle traite la mobilité, et le MDAE-Transformer supporte le déploiement à grande échelle avec une meilleure adaptation temporelle.

Ce travail rapproche l'empreinte RF du déploiement opérationnel et pose les bases de la sécurité au niveau physique dans les réseaux sans fil de nouvelle génération, particulièrement pour les communications de drones et la sécurité IoT où les approches cryptographiques seules s'avèrent insuffisantes.

Mots-clés : empreinte radiofréquence, sécurité couche physique, apprentissage profond, sécurité des drones, authentification IoT, architectures transformer, analyse cyclostationnaire, modélisation temporelle

Contents

List of Tables

Note: Tables appearing within the published papers included in Chapters 4–7 are not listed here. They can be found in the respective papers as presented in each chapter.

List of Figures

Note: Figures appearing within the published papers included in Chapters 4–7 are not listed here. They can be found in the respective papers as presented in each chapter.

List of Acronyms

ADC	Analog-to-Digital Converter
AWGN	Additive White Gaussian Noise
CFO	Carrier Frequency Offset
CI	Confidence Interval
CNN	Convolutional Neural Network
CNN-LSTM	Convolutional Neural Network - Long Short-Term Memory
CPU	Central Processing Unit
CSS	Chirp Spread Spectrum
DFT	Discrete Fourier Transform
DNN	Deep Neural Network
DoS	Denial of Service
DRNN	Deep Residual Neural Network
DSSS	Direct Sequence Spread Spectrum
DUT	Device Under Test
FFN	Feedforward Network
FFT	Fast Fourier Transform
FHSS	Frequency Hopping Spread Spectrum
FPGA	Field-Programmable Gate Array
GAN	Generative Adversarial Network
GELU	Gaussian Error Linear Unit
GoF	Goodness of Fit
GPS	Global Positioning System
GPU	Graphics Processing Unit
IQ	In-phase/Quadrature
IoT	Internet of Things
IoU	Intersection over Union
ISAC	Integrated Sensing and Communication
ISM	Industrial, Scientific, and Medical
kNN	k-Nearest Neighbors
LEO	Low Earth Orbit

LNA	Low Noise Amplifier
LoRa	Long Range
LOS	Line-of-Sight
LSTM	Long Short-Term Memory
MAC	Media Access Control
MDAE	Masked Denoising Autoencoder
MitM	Man-in-the-Middle
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NLOS	Non-Line-of-Sight
NTN	Non-Terrestrial Network
OSI	Open Systems Interconnection
OvA	One Versus All
PCA	Principal Component Analysis
PDF	Probability Density Function
PKI	Public Key Infrastructure
PSNR	Peak Signal-to-Noise Ratio
PSD	Power Spectral Density
ReLU	Rectified Linear Unit
RF	Radio Frequency
RFF	Radio Frequency Fingerprint
RFFI	Radio Frequency Fingerprint Identification
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SDR	Software Defined Radio
SF	Spreading Factor
SIEM	Security Information and Event Management
SigMF	Signal Metadata Format
SNR	Signal-to-Noise Ratio
SoA	State of the Art
THz	Terahertz
TinyML	Tiny Machine Learning
tSNE	t-distributed Stochastic Neighbor Embedding
TTDD	Train Today, Test Day-after-Tomorrow
TTDS	Train Today, Test Day-after-Same
TTSS	Train Today, Test Same-day
UAV	Uncrewed Aerial Vehicle
UMAP	Uniform Manifold Approximation and Projection
USRP	Universal Software Radio Peripheral
VOC	Visual Object Classification
WiFi	Wireless Fidelity
WiSIG	Wireless Specific Emitter Identification - Signal Dataset
YOLO	You Only Look Once
5G	Fifth Generation
6G	Sixth Generation

Chapter 1

1 Introduction

1.1 Motivation and Context

Modern wireless systems increasingly operate in contested, mobile, and resource-constrained environments where the consequences of authentication failure extend well beyond data confidentiality. Uncrewed Aerial Vehicle (UAV) fleets conduct surveillance and reconnaissance in dynamic airspace, Internet of Things (IoT) sensors support critical infrastructure monitoring, and autonomous systems coordinate across distributed networks with limited human oversight. In these settings, a compromised device is not simply a privacy breach. It can enable unauthorized access to operational networks, disrupt command and control, or allow adversarial devices to blend into legitimate traffic. Recent conflicts have made the stakes visible: commercial and military UAVs are routinely jammed, spoofed, or intercepted [1], while Canadian defense initiatives for Arctic domain awareness explicitly identify the detection of unauthorized aerial platforms as an operational priority. Securing these systems requires authentication that remains robust when devices move, when channels vary, when adversaries probe the network, and when computational budgets are constrained.

Internet of Things (IoT) and Uncrewed Aerial Vehicles (UAV) systems face serious security vulnerabilities where conventional cryptographic authentication is inadequate to secure against sophisticated attacks. Numerous credential-based attack vectors have been documented by recent security research where analyses show that GPS spoofing attacks allow full drone hijacking and wireless protocol flaws allow unauthorized device infiltration, and authentication bypasses in UAV swarm communications [1, 2, 3]. In October 2022, cybercriminals demonstrated a real-world exploitation of these vulnerabilities when they successfully compromised a US financial services company by using drones equipped with modified Wi-Fi devices on the building roof to obtain network credentials [4]. These incidents highlight a basic weakness in contemporary wireless security: even though cryptographic authentication

is mathematically sound, it is still susceptible to vectors like implementation errors and social engineering. The limitations of credential-based authentication become more crucial as more than 20 billion IoT devices are expected to be deployed worldwide by 2025 and more than 40 billion by 2034 [5, 6].

1.1.1 The Wireless Security Challenge

The new security challenges that our modern wireless networks face are driven by different convergent factors. One of the most challenging is the proliferation of resources-constrained IoT devices in uncontrolled and dynamic environments, this creates attack surfaces that traditional security techniques cannot adequately protect. These devices often lack computational resources to embed complex cryptographic operations or updating mechanisms used for patching vulnerabilities [7, 8]. Moreover, the emerging threat actors employ more sophisticated attacks that include advanced channel analysis, firmware extraction, and trained Machine Learning (ML) models that target implementation weaknesses rather than mathematical primitives [9]. Another fact that makes the actual countermeasures more complex is the increasing autonomy of wireless systems, from UAV swarms to industrial IoT networks. This demands new security mechanisms that can work without continuous human oversight or centralized credential management.

Traditional cryptographic approaches operate exclusively in the digital domain that consider the physical wireless channel as a transparent conduit for bit transmission. This abstraction, while enabling protocol standardization and interoperability, it ignores the rich information embedded in the physical signal itself that can be used for security purpose [10]. The electronic component tolerances and hardware imperfections that result from manufacturing variations, create unique "fingerprints" in transmitted RF signals that are intrinsic to the hardware and difficult to replicate [11]. These physical-layer features result from microscopic differences in oscillator crystals, power amplifier linearity, filter responses, and I/Q modulator balance, which are properties that even the original manufacturer cannot precisely control or duplicate, in contrast to cryptographic keys that are stored in memory and susceptible to extraction [12, 13].

1.1.2 RF Fingerprinting: Physical-Layer Device Authentication

RF fingerprinting (RFF) core idea is to exploit hardware-specific imperfections in wireless transmitters; this will create unique device signatures similar

to human biometric identification. When a wireless device transmits a signal, some small deviations from the ideal waveform create characteristic patterns that stay stable across transmissions and remain consistent regardless of the transmitted data or protocol [14]. Even devices from the same production line can have some measurable differences in oscillator frequency offset, power amplifier third-order intercept point (IP3), phase noise spectral density, and I/Q modulator gain/phase imbalance [15]. These small variations create device-specific signal distortions that machine learning algorithms can detect and classify.

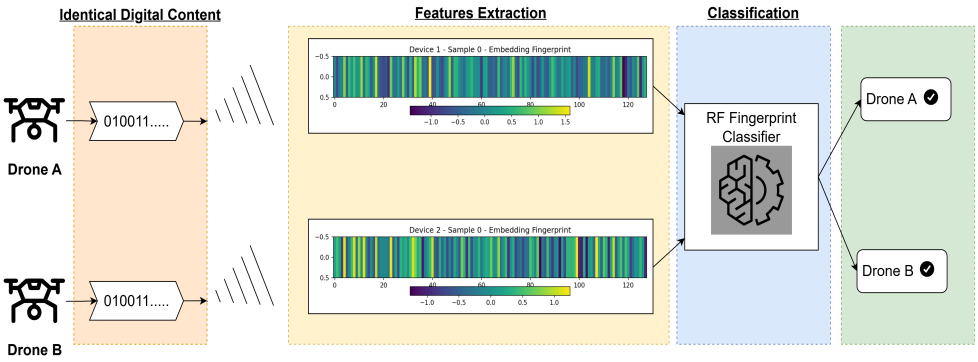


Figure 1.1: RF fingerprinting concept: Hardware imperfections in wireless transmitters create unique signal characteristics that enable device identification independent of transmitted content. Despite transmitting identical data, Device A and Device B produce distinguishable RF signatures due to manufacturing variations in oscillators, amplifiers, and other analog components.

Figure 1.1 illustrates this concept, two identical devices that are transmitting the same digital message will produce RF signals that contain distinct hardware-dependent characteristics. These fingerprints operate transparently at the physical layer, which provides authentication that adversaries cannot change by protocol manipulation. Apart from being resistant to credential compromise, physical-layer authentication has other benefits. One of these advantages is that RFF can operate independently of the communication protocol, this enables cross-protocol device identification that can be applied to WiFi, cellular, Bluetooth, LoRa, and proprietary wireless systems like for UAVs, without protocol-specific integration [16]. Also, the authentication process can be done without any hardware/software upgrade or cooperation from the target device, and just passive monitoring of transmitted signal can be enough for identification [17]. Furthermore, the RFF is a supplement to cryptographic security rather than its replacement; it offers a defense-in-depth

in which system access requires both logical and physical authentication to be successful [14]. However, translating these theoretical advantages into practical deployments is facing substantial challenges that we are addressing in this thesis.

1.1.3 The Deployment Challenge: Laboratory to Field Performance Gap

RFF research is not new; it's almost two decades of work trying to demonstrate its potential. However, the operational deployment remains limited. In the literature review we detailed in Chapter 2, analysis of more than 40 recent studies (2020–2024) showed that while RF fingerprinting systems achieve excellent accuracy under controlled laboratory conditions (typically $> 95\%$), performance degrades severely when deployed under realistic operational conditions [14, 16].

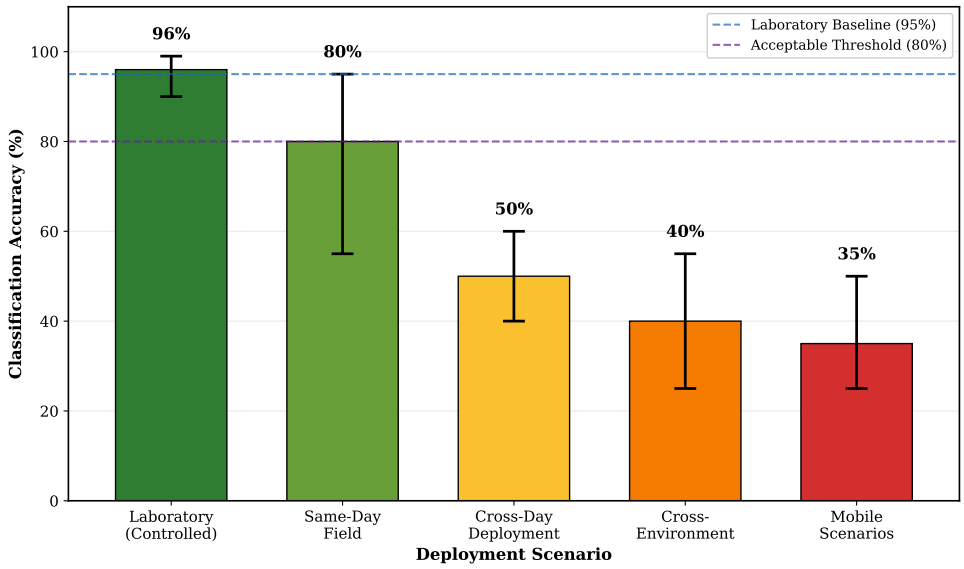


Figure 1.2: RF fingerprinting performance degradation from laboratory to field deployment. Meta-analysis of 40+ studies (2020–2024).

This deployment challenge is quantified across various operational dimensions in Figure 1.2. We have identified five key challenges that are preventing practical adoption, as follows:

Temporal Degradation: Studies report accuracy drops of 20–60% when ML models are trained using data from one day and tested on data from next

days, with performance keep dropping over weeks and months [18, 19]. This temporal sensitivity is due to many parameters such as the environmental changes (temperature, humidity, interferences, etc.) or device aging (oscillator drift, component parameter shifts, etc.).

Transmission Parameter Variations: RF fingerprinting systems usually overfit to specific transmission configuration like the modulation schemes or channel frequency. When devices operate under varying parameters, like frequency-hopping systems, the accuracy drops 15–45% as learned features fail to generalize when transmission parameter have changed [20].

Mobility and Channel Dynamics: This is one of the main challenges for systems like UAVs where the mobility introduces Doppler frequency shifts and time-varying multipath fading effects that can mask the hardware fingerprints. These systems experiment accuracy drops of 15–40% when they operate under Doppler shift more than 50Hz [21]. This issue becomes even worse in non-line-of-sight scenarios where multipath creates frequency-selective fading that distorts signal characteristics used to extract the device fingerprints.

Scalability Limitations: With the big number of connected devices like IoT in wireless networks, scalability becomes very important, and it should be taken into consideration. Classification-based RFF approaches suffer from computational complexity that keeps growing linearly with device population [22]. For example, to add new device to a trained model, a complete retrain is required with data from all devices, unless advanced techniques are used to preserve learned parameters. This issue becomes a barrier for large-scale deployment of IoT networks where devices continuously join and leave.

Unknown Device Detection: The closed-set classification methods cannot identify devices absent from the training data. They tend to force classification to known devices regardless of the confidence. This is an issue when dealing with spoofing attacks, this is particularly problematic in security-critical applications where adversarial devices represent the main threat [23]. The open-set recognition is still an unsolved challenge that need particular attention [24].

1.1.4 Research Opportunity and Impact

The convergence of advanced deep learning techniques with RF engineering knowledge creates great opportunities to address the deployment challenges we identified in the previous section. Recent developments in the integration of attention mechanisms, continual learning, self-supervised training, and embedding-based architectures [25, 26, 27, 28] provide architectural founda-

tions that can enable robust RF fingerprinting when combined with domain-specific characteristics such as signal processing insights.

In our thesis, we address each deployment barrier through four complementary approaches that integrate domain knowledge with modern deep learning: statistical feature enhancement for temporal adaptation, temporal modeling for mobility robustness, cyclostationary feature engineering for transmission parameter invariance, and transformer-based embeddings for scalability with open-set recognition. Each of our proposed approach targets specific limitations while building on insights from previous methods in order to propose practical deployment across diverse operational scenarios. The goal is to investigate the physical-layer authentication potential to benefit different fields such as critical infrastructure protection, military communications, industrial IoT, and autonomous vehicle coordination where traditional approaches have some limitations.

1.2 Threat Model and Security Scope

RF fingerprinting addresses a specific class of wireless security threats distinct from those targeted by cryptographic authentication. Understanding the threat model precisely is essential for evaluating the contributions of this thesis and positioning them within the broader security landscape.

1.2.1 Threats Addressed by This Thesis

- **Device impersonation.** An adversary transmits RF signals while claiming to be a legitimate authorized device. Because the adversary's hardware exhibits different manufacturing imperfections than the genuine device, RF fingerprinting detects the mismatch between the claimed identity and the observed physical-layer signature. This threat is addressed across all four contributions.
- **Unauthorized device access.** An unregistered or unauthorized device attempts to communicate within a secured network without possessing credentials. Chapter 7 addresses this through reconstruction-based anomaly detection, where devices absent from the enrollment registry produce elevated reconstruction error, enabling rejection at the physical layer.
- **Transmission parameter spoofing.** An adversary attempts to evade detection by varying transmission parameters such as frequency, power level, or modulation scheme to disrupt the fingerprinting system. Chapter 6 addresses this by operating in higher-order cyclostationary domains

that capture hardware-specific signatures invariant to transmission parameter changes.

- **Temporal drift exploitation.** An adversary exploits the temporal degradation of fingerprinting models by operating at times or in conditions that differ from model training, causing the system to fail or misclassify. Chapter 4 and Chapter 5 address this through multi-domain feature aggregation and temporal modeling strategies that maintain discriminative capability across varying operational conditions.

1.2.2 Threats Outside This Thesis Scope

- **Hardware-level device cloning.** An adversary with physical access to a target device precisely replicates its microscopic manufacturing variations at the component level. This attack is theoretically possible but practically infeasible given current manufacturing precision and requires specialized laboratory equipment beyond typical adversary capabilities.
- **Adversarial signal crafting.** An adversary deliberately crafts transmitted signals using optimization techniques to produce RF fingerprints matching a target device's enrollment embedding. Evaluating robustness to such attacks requires dedicated adversarial robustness testing not performed in this thesis, as stated in Section 1.6.
- **Relay and replay attacks.** An adversary captures and retransmits legitimate device signals to impersonate that device. While RF fingerprinting provides some resistance through timing and channel-dependent features, systematic evaluation of replay attack robustness falls outside the thesis scope.

1.2.3 Positioning Relative to Cryptographic Authentication

This thesis is explicitly framed as developing **complementary physical-layer security** that supplements rather than replaces cryptographic authentication. The two layers address orthogonal threat vectors: cryptography verifies identity claims at the protocol level, while RF fingerprinting verifies that the transmitting hardware matches the claimed identity at the physical level. Together, they create a defense-in-depth architecture where an adversary must simultaneously defeat both the cryptographic and physical-layer mechanisms to successfully impersonate a legitimate device.

1.3 Problem Statement

The literature review we present in Chapter 2 demonstrated that while RF fingerprinting shows strong theoretical foundations and laboratory performance, existing challenges still prevent reliable operation under realistic deployment conditions. The analysis of existing approaches reveals that the performance degradation is not only due to insufficient model capacity or training data quantity, but mainly from fundamental architectural and methodological limitations that fail to take into consideration the unique characteristics of RF signal variation.

1.3.1 Core Research Problem

This thesis addresses the fundamental research problem of developing RF fingerprinting approaches that can maintain robust device identification performance across diverse and dynamic operational conditions similar to those in real-world wireless environments, while meeting scalability and efficiency requirements that are necessary for practical deployment in UAV and IoT systems.

Building on the threat model established in Section 1.2, the problem decomposes into five technical challenges. Each challenge represents a distinct axis of variation that current approaches inadequately address, and each corresponds to one or more of the attack vectors defined in the threat model.

1.3.2 Research Challenges

Challenge 1: Temporal Generalization. As we mentioned earlier, device characteristics evolve continuously. Current approaches treat device signatures as static, learning features that capture momentary characteristics rather than stable hardware properties. This temporal sensitivity manifests as rapid performance degradation, with accuracy dropping from more than 90% to less than 50% when tested on non-seen data [14]. The challenge requires distinguishing gradual characteristic changes from adversarial manipulation such as spoofing attempts.

Challenge 2: Transmission Parameter Robustness. The variety of wireless devices requires in some cases operation across varying frequency channels, power levels, modulation schemes, and data rates. Existing RF methods usually overfit to specific parameters rather than learn hardware-invariant characteristics. This causes them to fail when transmission condi-

tions differ from training scenarios [20], producing degradation when devices switch channels or modulation schemes.

Challenge 3: Mobility and Dynamic Channels. Device mobility is a frequent case where communication introduces Doppler frequency shifts and multipath propagation. This creates time-varying channel responses that distort transmitted signals. For example, UAV applications experience Doppler shifts exceeding 100 Hz, which results in accuracy reductions of 30–50% [21].

Challenge 4: Scalability and Open-Set Recognition. Classification-based approaches show growing computational complexity with device count, and this becomes more critical when we try to classify 20–30 devices [22]. Another challenge is closed-set classification: when we try to identify unknown devices, these approaches tend to force classification to known classes regardless of confidence. This requires new methods that can provide reliable open-set recognition to distinguish known devices under varying conditions from unknown devices trying to join the network.

Challenge 5: Computational Efficiency. A further challenge is to fit advanced deep learning approaches that achieve good performance on devices with limited resources. IoT edge devices and embedded processors require solutions operating within strict constraints such as memory footprint under 1 GB and power consumption compatible with battery operation [29].

1.4 Research Questions and Thesis Structure

We translate the identified challenges into five research questions that guide the research and technical investigations, each one is addressed using different approaches presented in Chapter 4 to 7.

RQ1: How can RF fingerprinting achieve robustness to transmission parameter variations while maintaining hardware-specific discriminability?

In this question we investigate the feature extraction methods that capture hardware specific characteristics that can be robust when we face different transmission settings such as frequency or modulation scheme. We employ cyclostationary signal analysis to exploit the periodic statistical properties of RF signal that remain stable across transmission.

Addressed in Chapter 6: Advanced Feature Engineering with Continual Learning

RQ2: How can systems maintain device identification accuracy under temporal variations of device characteristics caused by aging and environmental drift?

Hardware characteristics are tiny variations that can be hidden due to environment conditions or hardware aging, this causes trained models to become outdated. In this question, we explore statistical enhancement techniques and apply them to multi-domain feature aggregation. The main goal is investigate all possible characteristics of RF signal and how they can contribute to temporal stability.

Addressed in Chapter 4: Statistical Feature Enhancement for RF Fingerprinting

RQ3: How can RF fingerprinting handle mobile scenarios with Doppler effects and fading conditions?

In this question we investigate the effects of mobility scenarios and how they can affect the identification accuracy. We explore temporal modeling architectures that explicitly capture both short-term dynamics (Doppler, fading) and long-term evolution (aging). We developed attention-based mechanisms that focus on discriminative temporal segments less affected by mobility.

Addressed in Chapter 5: Temporal Modeling for Mobile Scenarios

RQ4: How can systems scale to large device populations and still provide open-set recognition of unknown devices?

We change the direction in this part of our research to move from classification-based approaches, that are known for their limitation regarding exponential complexity growth and unknown devices detection, to embedded-based approaches. These techniques provide constant inference complexity regardless of population size. When combined with reconstruction-based anomaly detection, they can enable open-set recognition as we demonstrate in Chapter 7.

Addressed in Chapter 7: Anomaly Detection and Scalable Architecture

RQ5: How can sophisticated RF fingerprinting operate within computational constraints of IoT and edge devices?

Advanced signal processing and deep learning architectures often exceed deployment platform capabilities. We address this challenge by proposing memory-efficient continual learning approaches and lightweight attention mechanisms suitable for embedded deployment.

Addressed across Chapters 4–7 with explicit efficiency analysis in each

We emphasize that these research questions are not independent but represent complementary aspects of the broader deployment challenge. To this end, the thesis addresses these challenges through a systematic progression, whereby we first establish robust feature engineering foundations via statistical enhancement, extend these to dynamic conditions using temporal modeling, and subsequently achieve parameter invariance through cyclostationary analysis. Finally, we address large-scale deployment by utilizing embedding-based

architectures, each proposed approach contributes unique capabilities while revealing limitations that motivate subsequent methods.

1.5 Research Contributions

In this thesis, we make four primary technical contributions that specifically address the identified deployment challenges, each of which has been validated through peer-reviewed publications. We note here that our research was conducted through industry collaboration with Thales Defense & Security under a three-year Mitacs fellowship (2022–2025).

1.5.1 Contribution 1: Statistical Feature Enhancement for Temporal Robustness

Technical Contribution: We have developed a multi-domain feature aggregation framework that integrates different signal characteristics such as magnitude, phase, power spectral, entropy, and statistical features, which collectively exhibit complementary temporal stability characteristics. Our approach investigates the extent to find out which specific feature domains maintain consistency under temporal drift relative to environmental variations.

Performance Achievement: We note that the proposed statistical enhancement approach achieves a robust 99.6% same-day accuracy on the WiSIG dataset. Furthermore, it improves cross-day performance to 52.1% where traditional ResNet-50 baseline achieves only 37.5% under the same conditions [30]. The multi-domain aggregation demonstrates a 37.2 percentage point improvement compared to single-domain approaches, which we claim validates the necessity of feature diversity in providing robustness against temporal variations.

Novelty and Significance: We argue that prior works typically rely on single-domain features or end-to-end learning models that overlook the analysis of temporal stability properties. Our contribution establishes that different feature domains exhibit distinct aging characteristics, while spectral features remain stable over several weeks, phase-based features require daily recalibration. We emphasize that this insight enables hybrid approaches that effectively balance temporal robustness with discriminative power, consequently providing the foundation for our subsequent work on continual adaptation. We have published our results in peer-reviewed conference proceedings, demonstrating their practical applicability to UAV and IoT security scenarios [30].

Detailed in Chapter 4

1.5.2 Contribution 2: Temporal Modeling Architecture for Mobility Robustness

Technical Contribution: We present the development of a CNN-LSTM-Attention architecture where we explicitly model temporal dependencies within RF signals. This design enables a critical distinction between mobility-induced variations and the device-specific hardware characteristics. The use of a bidirectional LSTM component captures both forward and backward temporal relationships, while our attention mechanisms dynamically focus on the discriminative signal segments that are least affected by mobility.

Performance Achievement: Our temporal modeling approach successfully maintains a high 85.8% accuracy, even under demanding 100 Hz Doppler shift conditions that equivalent to high-speed UAV scenarios, compared to approaches lacking explicit temporal modeling which achieved only 73.2% [31]. Consequently, we utilized data augmentation strategies that synthesize diverse mobility conditions during training to improve generalization accuracy to Doppler shifts.

Novelty and Significance: We discuss that while LSTM architectures exist for general time-series analysis, their specific application to RF fingerprinting under conditions of extreme mobility can enhance performance. We emphasize that the integration of attention mechanisms with Doppler-aware data augmentation enables robust performance under operational conditions that would otherwise cause the complete failure of static approaches. We have published these results in international conference proceedings, with experimental validation conducted on commercial UAV controller datasets [31].

Detailed in Chapter 5

1.5.3 Contribution 3: Cyclostationary Feature Engineering with Progressive Learning

Technical Contribution: We present a novel neural network layer that implements second, fourth, and sixth-order cyclostationary feature extraction, that we combined with a progressive learning framework based on continual learning principles. This design effectively enables incremental device addition without catastrophic forgetting. We note here that this approach exploits the periodic statistical properties inherent in communication signals, lesson that we have learned from Chapter 2.

Performance Achievement: We emphasize that the cyclostationary approach achieves a notable 85.9% cross-transmission accuracy, consistently representing device identification across varying conditions. This constitutes

a 12.5 percentage point improvement compared to the 73.4% we achieved from our CNN-LSTM baseline [32]. Simultaneously, the progressive learning component successfully reduces memory requirements by 38% when compared to standard joint training and still maintain competitive classification performance.

Novelty and Significance: While cyclostationary analysis exists as a classical tool in signal processing literature, we claim that its integration with deep learning through custom differentiable layers, coupled with a progressive continual learning strategy, represents a novel contribution to the field. This work demonstrates that domain-specific signal processing knowledge, when properly integrated with modern network architectures, substantially outperforms purely data-driven approaches. Furthermore, the progressive learning framework provides the first demonstration of continual learning for RF fingerprinting, directly addressing practical deployment scenarios where devices must join networks incrementally. These significant results are currently under review in IEEE Internet of Things Journal [32].

Detailed in Chapter 6

1.5.4 Contribution 4: Transformer-Based Embedding Architecture for Scalable Deployment

Technical Contribution: We detail the development of the Masked Denoising AutoEncoder with Transformer (MADE) architecture, where we use embedding-based representation learning. We designed a dual training paradigm that combines masked reconstruction objectives with contrastive learning, consequently enabling both the generation of high-quality device representations and robust anomaly detection through reconstruction error analysis.

Performance Achievement: We confirm that the MADE achieves 99.9% accuracy under optimal signal conditions, which demonstrates its theoretical capability to approach perfect identification. Furthermore, evaluation on UAV controllers yielded 92.9% accuracy, a 7 percentage point improvement over the CNN-LSTM baseline proposed in Chapter 6, this validates that the architectural design provides consistent benefits across heterogeneous hardware platforms and communication protocols. We note here that the architecture additionally achieves 75.2% accuracy at the 80-device scale, exhibiting only a 5.0 percentage point degradation from the 5-device baseline, compared to classification approaches, which are known to exhibit exponential degradation.

Novelty and Significance: To the best of our knowledge, this work represents the first demonstration of transformer-based embeddings for RF fingerprinting that achieves realistic scalability while simultaneously provid-

ing open-set recognition capabilities through reconstruction-based anomaly detection. We claim that this architecture fundamentally advances beyond the conventional closed-set classification paradigm that has dominated prior work, thereby enabling practical deployment in large-scale IoT networks comprising hundreds of devices. Moreover, the dual training methodology establishes a crucial template for developing foundation models in signal processing domains, with pre-training objectives furnishing both discriminative and generative capabilities. These significant findings are currently being drafted for submission to a top-tier journal, supported by additional validation derived from our industry collaboration with Thales Defense & Security.

Detailed in Chapter 7

1.5.5 Impact and Validation

Publications and Peer Review: Our research has resulted in multiple peer-reviewed publications in competitive international venues:

Journal Articles

- N. Quadar, A. Chehri, and B. Debaque, "Advanced Security Frameworks for UAV and IoT: A Deep Learning Approach," *Internet of Things (Elsevier)*, vol. 32, pp. 101594, 2025. [30] → Chapter 4
- N. Quadar, A. Chehri, and B. Debaque, "Scalable Deep Learning for RF Fingerprinting: The MADE Architecture for Robust Physical-Layer Device Identification," *IEEE Open Journal of the Communications Society*, vol. 7, pp. 1973-1993, 2026. [33] → Chapter 7
- N. Quadar, A. Chehri, and B. Debaque, "Integrating Sensing, Communication, and Computing for Robust RF Fingerprinting in Consumer Electronics," IEEE Transactions On Consumer Electronics, 2026. (Under review) [32] → Chapter 6
- N. Quadar, A. Chehri, B. Debaque, H. Yanikomeroglu, and G.K. Kurt, "Unseen Signals, Real Threats: A Comprehensive Survey of AI Generalization in RF Fingerprinting for Wireless Security," *IEEE Open Journal of the Communications Society*, 2026. (Under review) [34] → Chapter 2

Conference Papers

- N. Quadar, A. Chehri, and B. Debaque, "Feature Learning and Continual Adaptation for Robust RF Fingerprinting in UAV Networks," in *Proc. IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, UK, 24–28 May 2026, to appear. [35] → Chapter 6

- N. Quadar, A. Chehri, and B. Debaque, "Leveraging Cyclostationary Features and Continual Learning for Robust RF Fingerprinting in IoT and UAV Networks," in *Proc. IEEE 103rd Vehicular Technology Conference (VTC2026-Spring)*, Nice, France, 9–12 June 2026, to appear. [36] → Chapter 6
- N. Quadar, A. Chehri, and B. Debaque, "Robust RF Fingerprinting for LoRa IoT Devices in Mobile Scenarios Using CNN-LSTM-Attention," in *Proc. IEEE 101st Vehicular Technology Conference (VTC2025-Spring)*, 2025, pp. 1–5. [31] → Chapter 5
- N. Quadar, A. Chehri, and B. Debaque, "Wireless Security and IoT Device Identification using RF Fingerprinting and Deep Learning," in *Proc. IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, 2024, pp. 1–5. [37]
- N. Quadar, A. Chehri, and B. Debaque, "Audio Recognition-based Method for RF Transmitters Classification using CNN-LSTM model," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Vancouver, BC, Canada, 2025, pp. 464–468. [38]

Magazine Articles

- N. Quadar, A. Chehri, and B. Debaque, "TinyML Datasets as Enablers of 6G Edge Intelligence: Key Insights and Research Gaps," *IEEE Wireless Communications*, 2026. [39]
- N. Quadar, A. Chehri, and B. Debaque, "TinyML Dataset Challenges in Enabling Scalable Intelligence for 6G Consumer Electronics," *IEEE Consumer Electronics Magazine*, 2026. [40]
- N. Quadar, M. Rahouti, M. Ayyash, S.K. Jagatheesaperumal, and A. Chehri, "IoT-AI/Machine Learning Experimental Testbeds: The Missing Piece," *IEEE Internet of Things Magazine*, vol. 7, no. 1, pp. 136–143, 2024. [41]

Industry Collaboration: Three-year Mitacs fellowship (2022–2025) with Thales Defense & Security where we worked with different teams to validate the practical relevance and deployment potential of our findings. The collaboration focuses on UAV authentication and critical infrastructure protection scenarios where physical-layer security can provide an extra security layer.

The collective impact of this work demonstrates RF fingerprinting's transition from laboratory investigation to deployment ready. Figure 1.3 illustrates our technical progression, showing how each contribution builds upon previous work while addressing limitations and future work.

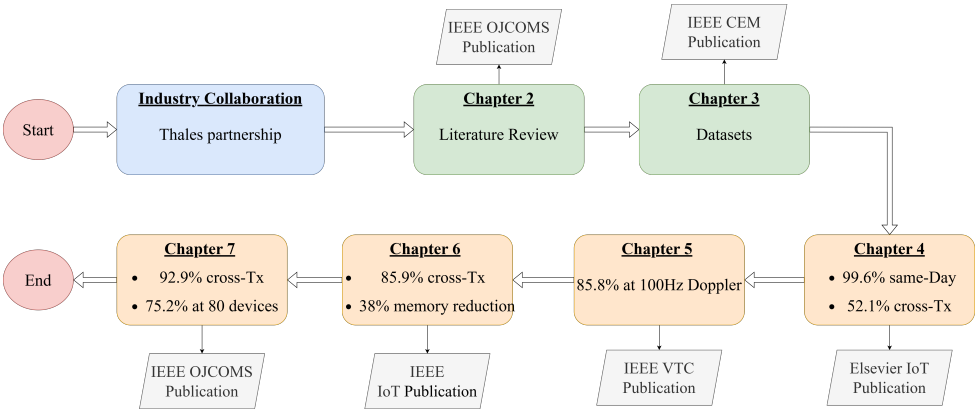


Figure 1.3: Thesis contributions timeline showing progression addressing complementary deployment challenges. Each contribution validated through peer-reviewed publication.

Table 1.1 summarizes the four primary contributions of this thesis, mapping each to its research question, key performance achievement, evaluation dataset, and publication venue.

1.6 Thesis Scope

We should state that our research focuses on developing deep learning approaches for RF fingerprinting in UAV and IoT deployment scenarios with priority to address generalization robustness across diverse operational conditions. Our investigation specifically addresses real-world challenges, such as natural environmental variations and mobility dynamics, but also maintaining computational efficiency for edge deployment. We note that adversarial attack resistance and deliberate hardware circuit modifications are beyond the purview of this scope; to this end, our primary emphasis remains on developing practical software-based solutions deployable with standard wireless transceivers under realistic operational constraints. The experimental validation, therefore, actively prioritizes assessment under practical deployment conditions over idealized laboratory scenarios, with all performance metrics unequivocally emphasizing robustness to real-world variations rather than simply achieving peak accuracy under tightly controlled settings.

Table 1.1: Summary of thesis contributions, research questions addressed, key results, and publication status.

Contribution	RQ	Key Result	Dataset	Publication
Statistical Feature Enhancement (Ch. 4)	RQ2	52.1% cross-day (+14.6 pp vs baseline)	GLOBECOM22, ACMWiSec21	*Internet of Things*, Elsevier, 2025
CNN-LSTM Temporal Modeling (Ch. 5)	RQ3	85.8% at 100 Hz Doppler (+12.6 pp)	LoRa-60	IEEE VTC-Spring, 2025
Cyclostationary Features + Progressive Learning (Ch. 6)	RQ1	85.9% cross-TX; 38% memory reduction	UAV 8/17-device	IEEE OJ-COMS; IEEE ICC 2026; IEEE VTC 2026
MADE Architecture (Ch. 7)	RQ4	75.2% at 80 devices; 5.0% degradation over 16× scale	WiSIG (174 devices), UAV	IEEE OJ-COMS, 2026

1.7 Thesis Organization

This thesis follows a **paper-based format**, also known as a manuscript-style thesis, as approved by the Royal Military College of Canada. In this format, Chapters 4 through 7 each incorporate a peer-reviewed or submitted journal or conference publication as their core technical content. Each technical chapter is structured in two parts: the published or submitted manuscript, presented in its original peer-reviewed form, followed by a critical analysis section that contextualizes the paper's contributions within the broader thesis narrative, examines how it integrates with the experimental framework established in Chapter 3, and identifies limitations that motivate the subsequent chapter. This structure means that notation, formatting, and terminology may vary slightly between chapters to reflect the conventions of each publication venue. Readers are encouraged to treat each technical chapter's manuscript section as a self-contained contribution and the critical analysis section as the thesis wrapper that connects it to the overall research arc.

The remainder of this thesis is structured to provide a clear progression, transitioning from theoretical foundations through specific technical contributions to final deployment guidance. Figure 1.4 illustrates this organizational

flow, showing how each chapter builds upon the previous one.

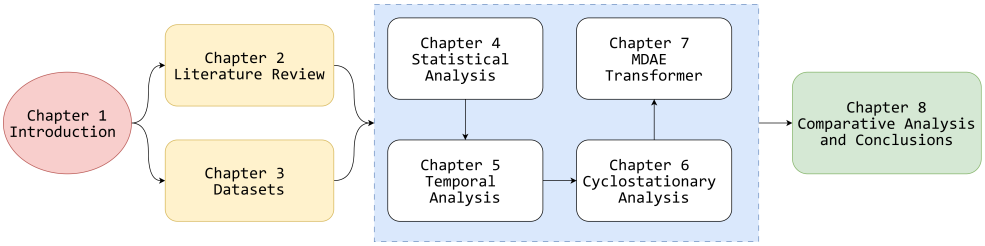


Figure 1.4: Thesis organization.

Chapter 2: Literature Review and Physical-Layer Security Foundations: establishes the theoretical background for RF fingerprinting. This chapter provides a critical analysis of over 40 recent studies to quantify the main lab-to-field performance gap and establish the problem formulation that motivates this research.

Chapter 3: Datasets and Experimental Framework: details the evaluation methodology employed throughout our work. We describe six distinct datasets, including WiFi, LoRa, and UAV controllers, to ensure that our protocols enable fair comparison across approaches while maintaining reproducibility.

Chapters 4–7: Technical Contributions constitute the core technical contributions of this thesis, each addressing a unique aspect of the deployment challenge:

- **Chapter 4: Statistical Feature Enhancement:** focuses on multi-domain feature engineering. By utilizing feature aggregation and adaptation strategies, this chapter establishes the foundations required to achieve a 99.6% same-day accuracy and a significantly improved 52.1% cross-day performance.
- **Chapter 5: Temporal Modeling for Mobile Scenarios:** presents the CNN-LSTM-Attention architecture. This design maintains 85.8% accuracy under 100Hz Doppler shifts.
- **Chapter 6: Cyclostationary Feature Engineering with Progressive Learning:** we utilize higher-order statistical analysis to achieve 85.9% cross-transmission generalization. This chapter demonstrates that domain-specific signal processing knowledge, when paired with a progressive learning framework, can reduce memory requirements by 38% while substantially outperforming purely data-driven methods.
- **Chapter 7: Masked Denoising Autoencoder with Transformer Architecture:** we shift the paradigm toward scalable deployment. We

introduce reconstruction-based anomaly detection as a foundation for open-set recognition and demonstrate how we achieved 75.2% accuracy at the 80-device scale.

- **Chapter 8: Comparative Analysis and Conclusions:** in this chapter, we synthesizes the findings across all four approaches. This concluding chapter provides the performance assessment and identify remaining research limitations and future directions based on the gaps observed during our investigation.

Chapter 2

2 Literature Review and Physical-Layer Security

2.1 Introduction

In this chapter, we present our examination of RF fingerprinting technology where we analyze its historical evolution and the fundamental limitations that define its position within the contemporary physical-layer security landscape. Our primary goal is to establish a solid foundation and identify gaps that need new contributions. We have organized this chapter to provide both a broad coverage of existing literature and a clear problem identification that justifies our technical work. Specifically, Section 2.2 presents the literature review, prepared as a comprehensive survey, that presents the evolution of RF fingerprinting from traditional, manual feature engineering until state-of-the-art deep learning architectures. This review identifies the persistent generalization challenges that, as we argue, continue to prevent practical deployment despite the impressive accuracy achieved in controlled settings.

Ultimately, this chapter serves as the foundation for understanding the possible potential and current constraints of RF fingerprinting technology. By providing this essential context, we demonstrate the clear necessity for new solutions that can bridge the gaps between laboratory performance and the deployment requirements of modern UAV and IoT networks.

2.2 Comprehensive Literature Analysis

In this section, we present our survey of RF fingerprinting technology, submitted to IEEE Open Journal of the Communications Society. We include the full analysis below to provide detailed technical background and literature analysis.

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version 11 January, 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.011100

Radio Frequency Fingerprinting: A Survey of AI Generalization Challenges and Solutions for Wireless Security

Nordine QUADAR[1] *(Memeber, IEEE)*, Abdellah CHEHRI[1] *(Senior Memeber, IEEE)*, Benoit DEBAQUE.[2], Halim YANIKOMEROGLU[3] *(Fellow, IEEE)* AND Gunes Karabulut KURT[4] *(Senior Memeber, IEEE)*

[1]Royal Military College of Canada, Kingston, ON, Canada
[2]Thales Research and Technology, Quebec, QC, Canada
[3]Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada
[4]Polytechnique Montréal, Montréal, QC, Canada

CORRESPONDING AUTHOR: Nordine Quadar (e-mail: quadar@rmc.ca).

This work was supported by Mitacs Accelerate Fellowship program.

ABSTRACT Radio frequency (RF) fingerprinting has emerged as a promising physical-layer security technique for device authentication in wireless networks, offering protocol-independent identification by exploiting hardware-induced signal characteristics and imperfections. This paper explores its challenges and potential within the heterogeneous wireless networks, including UAV communications and IoT deployments which are characterized by dynamic operating conditions, heterogeneous device profiles, and elevated security demands. Although deep learning-based methods have substantially improved controlled recognition accuracy from 70–90% to as high as 95–99%, the generalization gap, identified as the severe performance degradation when deployment conditions differ from training environments, remains the fundamental barrier to reliable deployment in practical environments. Systems frequently demonstrate performance degradation, ranging from 30–70%, when confronted with unseen datasets and real-world operating conditions that differ from those encountered during training. This survey systematically traces the evolution of RF fingerprinting, from traditional manual feature engineering to cutting-edge learning architectures, with a particular emphasis on generalization challenges. Key research gaps are identified in multi-factor generalization modeling, real-world validation under diverse conditions, scalable architecture design for broad device deployments, adaptive learning mechanisms for dynamic contexts, and computational efficiency for resource-constrained platforms. The analysis provides foundational insight for advancing next-generation RF fingerprinting techniques capable of sustaining robust and consistent performance across highly variable and complex wireless environments. By clarifying current limitations and future directions, this work contributes to the practical realization of secure RF-based authentication for heterogeneous wireless networks.

INDEX TERMS RF fingerprinting, physical-layer security, generalization challenges, deep learning, wireless authentication, IoT security, UAV communications.

I. Introduction

The convergence of heterogeneous wireless technologies, including the Internet of Things (IoT), Uncrewed Aerial Vehicles (UAVs), and emerging 6G networks, is reshaping distributed cyber-physical systems across smart cities, industrial automation, defense, and emergency response. These systems operate across a fragmented spectrum of wireless protocols (WiFi, Bluetooth, LoRa, cellular, and specialized UAV control links), creating heterogeneous environments where billions of devices must be securely identified and authenticated [1]–[4]. As outlined in the 3GPP roadmap, Release 20 initiates 6G foundational studies while Release 21 addresses normative standardization [5], with Non-Terrestrial Networks (NTNs) extending infrastructure

TABLE 1. Comprehensive Acronym Glossary

3GPP	3rd Generation Partnership Project
5G	Fifth Generation (Mobile Network)
6G	Sixth Generation (Mobile Network)
A2G	Air-to-Ground
ADC	Analog-to-Digital Converter
AI	Artificial Intelligence
CNN	Convolutional Neural Network
CNN-LSTM-Attention	CNN–LSTM–Attention (DL Architecture)
CFR	Channel Frequency Response
CIR	Channel Impulse Response
CMOS	Complementary Metal-Oxide-Semiconductor
CRM	Challenge-Response Mechanism
CSI	Channel State Information
DARPA	Defense Advanced Research Projects Agency
DAConv	Dilated Attention Convolutional Network
DC	Direct Current
DeepCRF	Deep Contrastive RF Fingerprinting
DFT	Discrete Fourier Transform
DNN	Deep Neural Network
DoS	Denial-of-Service
DSSS	Direct-Sequence Spread Spectrum
dB	Decibel
dBm	Decibel-milliwatts
FMCW	Frequency-Modulated Continuous-Wave
FFT	Fast Fourier Transform
FPGA	Field-Programmable Gate Array
FSK	Frequency Shift Keying
GAN	Generative Adversarial Network
GDM	Generative Diffusion Model
GLOBECOM	IEEE Global Communications Conf.
GPS	Global Positioning System
GPU	Graphics Processing Unit
G2G	Ground-to-Ground
HAP	High-Altitude Platform
ICC	Int'l Conf. on Communications
INFOCOM	IEEE Conf. on Computer Communications
I/Q	In-phase / Quadrature
IIoT	Industrial Internet of Things
IoT	Internet of Things
k-NN	k-Nearest Neighbors
LoRa	Long Range
LoRaWAN	Long Range Wide Area Network
LSSVM	Least Squares Support Vector Machine
LSTM	Long Short-Term Memory
MAC	Media Access Control
MCC	Matthews Correlation Coefficient
MDA/ML	Multi. Discriminant Analysis w/ ML
MitM	Man-in-the-Middle
ML	Machine Learning
MPE	Multidimension Permutation Entropy
MSCNN	Multi-Sampling CNN
N/A	Not Applicable
NTN	Non-Terrestrial Network
OSI	Open Systems Interconnection
PCA	Principal Component Analysis
PKI	Public Key Infrastructure
PLA	Physical-Layer Authentication
PSD	Power Spectral Density
PUF	Physical Unclonable Function
QAM	Quadrature Amplitude Modulation
RBF	Radial Basis Function
RF	Radio Frequency
RF-Diffusion	RF-tailored Diffusion Architecture
RFMLS	RF Machine Learning Systems
RFE	Recursive Feature Elimination
RiffNet	Custom DL Model (Name)
RNN	Recurrent Neural Network
RSSI	Received Signal Strength Indicator
ResNet	Residual Network
SI	International System of Units
SRW	Soldier Radio Waveform
SNR	Signal-to-Noise Ratio
SSIM	Structural Similarity Index Measure
SVM	Support Vector Machine
UAV	Uncrewed Aerial Vehicle
URLLC	Ultra-Reliable Low-Latency Communication
USRP	Universal Software Radio Peripheral
VGG	Visual Geometry Group (CNN Style)
WiFi / Wi-Fi	Wireless Fidelity
X310	USRP Model by Ettus Research

via satellites, high-altitude platforms, and UAVs to support global coverage and resilience, introducing new authentication challenges [6]. The scale and heterogeneity of these deployments render conventional cryptographic approaches increasingly impractical, key management across thousands of diverse devices operating under different protocols presents scalability challenges that traditional Public Key Infrastructure (PKI) frameworks struggle to address [7].

In this context, Radio Frequency (RF) fingerprinting emerges as a compelling physical-layer security technique. By exploiting unique hardware-induced signal characteristics, manufacturing tolerances, component aging, and circuit imperfections that produce device-specific transmission signatures, RF fingerprinting enables lightweight, protocol-independent device identification without cryptographic overhead [8]–[10]. Recent survey data reveals the urgency of this need: approximately 97% of organizations report difficulty securing IoT devices, and 98% have experienced digital certificate outages with average losses exceeding $2.25 million per incident [11]. Despite achieving 95–99% accuracy in controlled laboratory conditions [12], [13], however, RF fingerprinting systems consistently degrade to 30–70% accuracy [14]–[16] in operational deployments, a generalization gap that constitutes the central barrier to practical deployment and the organizing theme of this survey.

To illustrate the operational diversity and protocol heterogeneity within wireless network deployments, Table 2 presents a comparative summary of communication standards and functional roles employed across deployment scenarios.

As these technologies increasingly converge, they underpin mission-critical applications that are highly susceptible to sophisticated and evolving threat vectors. In smart city deployments, UAVs integrated with IoT sensors perform real-time monitoring of traffic flow, environmental conditions, and integrity of the urban infrastructure. Unauthorized access or system compromise in these contexts can lead to widespread service disruption, data leakage, or violations of citizen privacy [4]. Within industrial IoT (IIoT) environments, UAV-assisted surveillance of manufacturing processes and logistical operations introduces novel attack surfaces for cyber-physical sabotage and industrial espionage. Adversaries may exploit RF vulnerabilities to infiltrate telemetry systems or manipulate process control [3].

Additionally, in precision agriculture systems, UAVs operating in tandem with terrestrial IoT sensors facilitate autonomous crop monitoring and irrigation management. Security breaches in such setups not only threaten sensitive agronomic data but may also interrupt operational continuity, potentially leading to resource waste, yield losses, or ecosystem imbalance. Collectively, these use cases exemplify the high stakes of securing heterogeneous wireless deployments, especially in light of persistent generalization challenges in RF fingerprinting-based authentication frameworks. Addressing these multi-layered risks requires robust, adaptable,

TABLE 2. Functionalities and Communication Protocols of UAV and IoT Systems

System	Functionality	Communication Protocols
UAV	Autonomous navigation and surveillance, capturing aerial data for reconnaissance or target acquisition.	5G and beyond, 6G, NTN, Cellular, Link 16, WiFi
UAV	Communication relay, extending secure network coverage for tactical operations.	5G and beyond, 6G, NTN, Cellular, Link 16, SRW
IoT	Sensor-based data collection, monitoring environmental or operational conditions (e.g., temperature, humidity).	Bluetooth, LoRa, WiFi, Sigfox, 6LoWPAN, LoRaWAN
IoT	Device-to-device communication, enabling local data exchange or control within IoT networks.	Bluetooth, LoRa, WiFi, Sigfox, 6LoWPAN, LoRaWAN, 5G

and scalable security mechanisms designed to withstand dynamic conditions and diverse adversarial strategies across wireless communication layers.

A. Limitations of Protocol-Level Security in Dynamic Environments

Traditional encryption and authentication mechanisms, while effective in static environments, encounter fundamental challenges in dynamic wireless networks. Four interconnected limitations constrain their applicability in UAV and IoT deployments.

Dynamic topology creates constantly changing network configurations as UAVs enter and leave communication range unpredictably, while IoT device heterogeneity compounds this challenge across devices with varying computational capabilities, power constraints, and communication protocols [7]. This intermittent connectivity disrupts traditional session-based security protocols that assume continuous communication channels.

Resource constraints further limit conventional cryptographic approaches. UAVs face strict power consumption limitations affecting computationally intensive encryption operations, while many IoT devices operate under even more severe computational and energy constraints [17]. With 69% of organizations reporting increased cyber-attacks on IoT devices [11], the inadequacy of currently deployed lightweight security measures is well-documented [2].

Key management complexity becomes particularly acute at scale. Managing cryptographic keys across thousands of heterogeneous devices operating under different protocols presents scalability challenges that traditional PKI frameworks struggle to address [7], [18], a problem intensified when UAVs using cellular or 5G links must securely communicate with IoT devices operating on LoRa, WiFi, or Bluetooth.

Protocol vulnerabilities expose additional security gaps. Many IoT devices rely on legacy standards with known weaknesses [19], while UAV communication protocols were designed with operational efficiency rather than security as the primary concern [1]. Documented vulnerabilities include unencrypted GPS signals and susceptibility to jamming and spoofing attacks [17]. These limitations collectively motivate physical-layer security approaches that operate independently of protocol-level assumptions. Figure 1 contrasts protocol-level and physical-layer security across resource and scalability dimensions [9], [10].

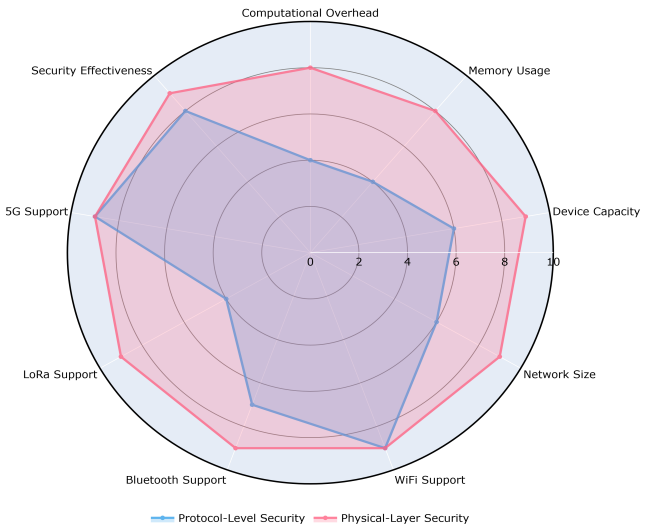


FIGURE 1. Protocol-Level vs. Physical-Layer Security: Resource and Scalability Trade-offs

B. The Generalization Challenge

Despite the promising advantages of RF fingerprinting for physical-layer security, a critical barrier prevents its widespread practical deployment: the generalization challenge. RF fingerprinting models trained under controlled laboratory conditions consistently fail to maintain performance when deployed in real-world environments, a gap that represents the central research problem addressed by this survey and the primary reason RF fingerprinting has not yet achieved widespread operational adoption despite years of impressive laboratory results.

The root cause lies in how deep learning models learn: they optimize for statistical patterns in training data, memorizing environment-specific conditions rather than isolating the hardware-invariant characteristics that make RF finger-

printing theoretically robust. When deployment conditions differ from training conditions, as they inevitably do in operational wireless networks, these learned associations break down regardless of architectural sophistication.

Generalization failures manifest across five dimensions that Section III analyzes in depth: cross-transmission parameter variations, temporal degradation due to hardware aging and environmental drift, spatial and environmental sensitivity, mobility-induced Doppler effects, and scalability limitations as device populations grow. Each dimension presents distinct technical challenges, and their compound interaction in real deployments creates degradation that exceeds the sum of individual factor impacts, a nonlinear compounding effect that current solutions, which typically address dimensions in isolation, are not designed to handle.

The practical implications extend beyond technical performance. Systems requiring frequent retraining incur substantial operational costs, performance uncertainty limits reliability guarantees for security-critical applications, and the mismatch between laboratory validation and field deployment creates fundamental barriers to commercial and military adoption. This survey organizes the entire literature around understanding, quantifying, and addressing these limitations, positioning generalization as a strategic priority rather than a peripheral metric.

C. Survey Contributions

The field of RF fingerprinting has been surveyed from various perspectives over the past decade. To establish the unique contribution of this work, Table 3 provides a comparison with existing surveys, highlighting differences in focus, scope, and treatment of generalization challenges.

As shown in Table 3, this survey distinguishes itself through generalization as the primary organizing principle, structuring the entire discussion around this central barrier to practical deployment rather than organizing chronologically or by methodology. We provide quantitative evidence from various studies documenting performance degradation across all five generalization dimensions. Beyond cataloging existing approaches, we analyze why current solutions address only specific scenarios while fundamental limitations persist, with emphasis on resource constraints and deployment practicality often overlooked in laboratory-focused surveys. The key contributions are as follows:

- **Foundational Context:** Establishes the theoretical underpinnings of RF fingerprinting, emphasizing hardware-induced imperfections and their role in device-level authentication.
- **Technique Evolution and Performance Analysis:** Surveys the evolution from manual feature engineering and statistical modeling to advanced deep learning architectures (e.g., CNN, LSTM), highlighting performance characteristics in static or controlled environments.
- **Critical Generalization Challenge Assessment:** Offers an in-depth analysis of limitations in generalization across transmission types, temporal shifts, mobility, and scalability, pinpointing factors that hinder real-world deployment.
- **Existing Solutions and Gaps:** Reviews current mitigation strategies, including domain adaptation, data augmentation, and transfer learning, while discussing their capabilities and persistent limitations.
- **Survey and Research Agenda:** Tracks the methodological progression from handcrafted features to modern learning frameworks, and identifies five critical challenges for future research:
  - Modeling multi-factor generalization
  - Real-world validation under diverse and unpredictable conditions
  - Scalable design architectures for mass device deployment
  - Adaptive learning mechanisms suited to dynamic contexts
  - Computational efficiency on resource-constrained platforms

D. Survey Organization and Methodology

This survey is organized to progressively build understanding of RF fingerprinting's potential and limitations, with generalization challenges as the unifying thread connecting all sections. Figure 2 illustrates the logical flow and connections between sections.

Section II establishes the theoretical foundation of RF fingerprinting, hardware imperfection principles, signal processing methods, and protocol characteristics across WiFi, Bluetooth, LoRa, ZigBee, and UAV control links, providing the technical context necessary to understand why generalization challenges emerge. Section III, the core analytical contribution, characterizes generalization failures across cross-transmission, temporal, environmental, mobility, and scalability dimensions with quantitative evidence, identifying root causes rather than merely documenting performance gaps. Section IV critically reviews existing deep learning solutions, domain adaptation, data augmentation, and continual learning, analyzing achievements versus persistent limitations with attention to edge deployment feasibility. Section V identifies concrete research directions including public dataset gaps, foundation model fine-tuning pathways, and prioritized research agenda grounded in the gap analysis of Sections III and IV. Section VI synthesizes key findings and reinforces the imperative to solve generalization challenges for practical deployment in heterogeneous wireless networks.

Methodology: This literature review examines RF fingerprinting techniques and their generalization challenges through an analysis of recent publications (2020–2025) and seminal works in the field. The review prioritizes publications from high-impact venues including IEEE Transactions on Wireless Communications, IEEE Transactions on Infor-

TABLE 3. Comparison of RF Fingerprinting Surveys: Focus and Generalization Treatment

Survey	Year	Primary Focus	Gen. Analysis	Unique Contribution
[20]	2017	Public safety applications	Brief	Early RF fingerprinting in emergency scenarios
[21]	2022	ML for IoT wireless	Dedicated	Machine learning integration for IoT
[22]	2024	Traditional vs DL methods	Central	Traditional-DL comparison
[23]	2024	LoRa-specific	Dedicated	Technology-specific (LoRa) deep dive
[24]	2024	IoT physical layer auth	Partial	Security-centric IoT authentication
[25]	2025	Statistical vs DL features	Partial	Dual-axis feature taxonomy organization
[26]	2024	Deployment challenges	Central	Tutorial on lab-to-deployment gap
This Survey	**2026**	**Generalization-centric**	**Organizing theme**	**Generalization as primary organizing principle**

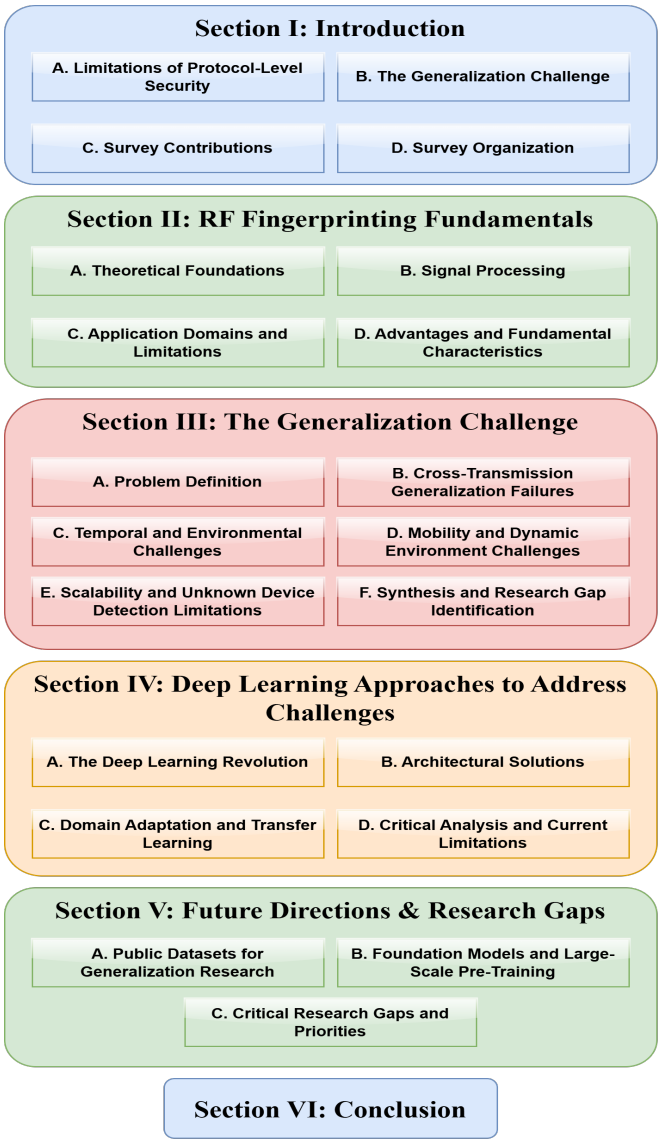


FIGURE 2. Survey Structure

mation Forensics and Security, and IEEE Internet of Things Journal, supplemented by relevant conference proceedings from INFOCOM, ICC, and GLOBECOM.

II. RF Fingerprinting: Principles and Applications

This section establishes the theoretical and practical foundations necessary to understand RF fingerprinting and its generalization challenges. We begin with the hardware imperfection principles that create unique device signatures, examine signal processing methods for fingerprint extraction, explore protocol-specific applications, and review traditional approaches that preceded deep learning methods. Understanding these fundamentals is essential for comprehending why generalization challenges emerge and persist in RF fingerprinting systems.

A. Theoretical Foundations and Hardware Imperfections

RF fingerprinting is fundamentally based on the exploitation of unintentional hardware variations that occur during the manufacturing process of wireless devices. These variations produce unique, persistent signatures in transmitted RF signals that can be leveraged for device identification independently of digital content or communication protocol [27], [28]. The theoretical foundation rests on the principle that manufacturing tolerances in analog front-end components introduce distinctive distortions into transmitted waveforms, distortions that are intrinsic to the physical structure of each device and fundamentally difficult to replicate or forge [29].

Several distinct categories of hardware imperfections contribute to these unique RF signatures, each manifesting in different aspects of the transmitted signal.

Oscillator Instabilities and Phase Noise: Local oscillators in wireless devices exhibit manufacturing variations that produce distinctive phase noise profiles and frequency stability characteristics. Modeling the transmitted carrier as:

$$s(t) = A\cos\big(2\pi f_c t + \phi(t)\big) \tag{1}$$

where A is the signal amplitude, f_c is the nominal carrier frequency, and $\phi(t)$ is a device-specific stochastic phase

noise process. The power spectral density of $\phi(t)$ follows a $1/f^2$ profile at offsets close to the carrier, with the exact profile determined by oscillator component tolerances [30]. These variations manifest as unique spectral signatures and temporal phase variations that persist across transmissions.

Power Amplifier Nonlinearities: Manufacturing tolerances in power amplifier components create device-specific nonlinear distortion patterns. Using a polynomial model for the PA output $y(t)$ as a function of input $x(t)$:

$$y(t) = \sum_{k=1,3,5,\ldots}^{K} a_k \cdot x^k(t) \tag{2}$$

where the odd-order coefficients $\{a_1, a_3, a_5, \ldots\}$ are device-specific constants determined by component fabrication [28]. The third-order term a_3 dominates intermodulation distortion, while higher-order terms introduce harmonic content that forms distinctive spectral signatures unique to each device.

I/Q Imbalance and DC Offset: Manufacturing variations in mixers, filters, and analog-to-digital converters create amplitude and phase mismatches between the in-phase (I) and quadrature (Q) signal components. For amplitude imbalance ε_A and phase imbalance ε_ϕ, the distorted complex baseband signal $\tilde{s}(t)$ relates to the ideal signal $s(t)$ as:

$$\tilde{s}(t) = \alpha_1 \cdot s(t) + \alpha_2 \cdot s^*(t) \tag{3}$$

where $\alpha_1 = \frac{1}{2}(1 + \varepsilon_A e^{j\varepsilon_\phi})$ and $\alpha_2 = \frac{1}{2}(1 - \varepsilon_A e^{-j\varepsilon_\phi})$ are device-specific imbalance coefficients [27]. The image component introduced by $\alpha_2 \cdot s^*(t)$ creates characteristic constellation asymmetries and spectral distortions that serve as reliable identification features.

Clock Skew and Timing Variations: Differences in crystal oscillator characteristics and clock distribution networks affect sampling rates, symbol timing, and frame synchronization, creating device-specific temporal patterns observable in preamble and guard interval structures [31].

Filter Response Variations: Component tolerances in analog filters introduce variations in center frequency, bandwidth, and group delay characteristics, creating unique spectral shaping effects evident in frequency-domain analysis [32].

Figure 3 illustrates how these hardware imperfections manifest throughout the RF transmitter chain, with each component contributing distinct device-specific characteristics.

The persistence and uniqueness of RF fingerprints derive directly from these physical imperfections. Because the coefficients in Equations (1)–(2) are determined by analog component fabrication, they remain stable across transmission sessions, environmental conditions, and power cycling [29], [33]. The combination of multiple independent imperfection sources, each with its own manufacturing tolerance distribution, creates a high-dimensional characteristic space where the probability of two devices sharing identical parameters approaches zero, providing strong statistical uniqueness guarantees even under practical measurement constraints [29].

B. Signal Characteristics and Processing Methods

Extracting discriminative RF fingerprints from wireless transmissions requires sophisticated signal processing techniques that isolate hardware-induced variations while suppressing transmission-dependent effects. The signal processing pipeline begins with I/Q signal acquisition, where the complex baseband representation of the received signal preserves both amplitude and phase information necessary for fingerprint extraction.

I/Q signal acquisition must meet specific requirements to ensure reliable fingerprint extraction. Adequate sampling rates are essential to capture the bandwidth of interest while avoiding aliasing effects that could mask or distort hardware signatures. Signal-to-noise ratio (SNR) requirements vary depending on the specific fingerprinting technique employed, but generally require SNR levels above 10 dB for reliable operation [28]. Preprocessing steps typically include carrier frequency offset estimation and compensation, timing synchronization, and amplitude normalization to remove transmission-dependent variations while preserving hardware-specific characteristics [29], [34]. Feature extraction operates across multiple signal domains, each providing complementary information about hardware characteristics.

Time domain features capture temporal variations in signal amplitude and phase. These include transient signal characteristics during power-up and power-down sequences, steady-state amplitude variations, and phase evolution patterns. Time-domain analysis is particularly effective for capturing characteristics of power amplifier behavior and oscillator stability [31].

Frequency domain characteristics reveal spectral signatures created by hardware imperfections. Fast Fourier Transform (FFT) analysis exposes frequency-dependent variations in signal power, spectral asymmetries caused by I/Q imbalance, and harmonic distortion patterns from nonlinear components. Frequency-domain features provide complementary information to time-domain characteristics [32].

Statistical features characterize the probability distributions of signal parameters. Moments of signal amplitude and phase distributions (mean, variance, skewness, kurtosis) capture device-specific statistical signatures. Advanced entropy-based measures, such as multidimension permutation entropy, can reveal subtle nonlinear dynamics in hardware behavior that are invisible to conventional statistical analysis [28].

Transform domain features employ sophisticated mathematical transforms to reveal hidden signal structures. Wavelet transforms provide time-frequency localization that can isolate transient events and capture multi-scale signal variations. Discrete Fourier transforms and their variants expose periodic patterns and spectral correlations that characterize filter responses and oscillator behavior [32].

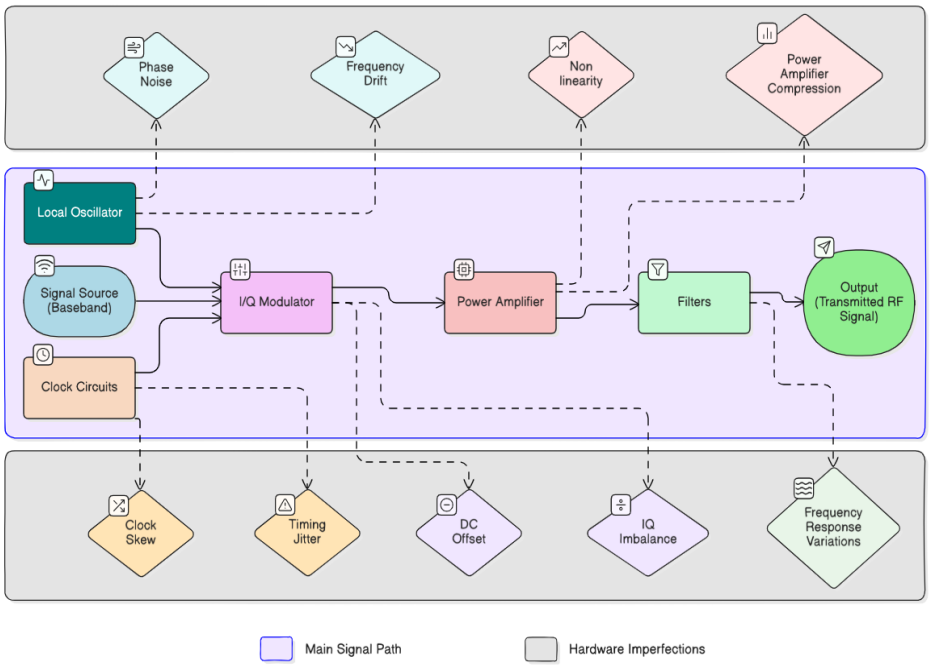


FIGURE 3. Hardware Imperfection Sources in RF Transmitter Chain: each analog component introduces characteristic distortions, phase noise at the oscillator, I/Q imbalance at the modulator, nonlinear distortion at the power amplifier, and spectral shaping at the filter, that collectively produce a device-specific RF fingerprint.

The signal processing pipeline from raw RF to classification involves several critical stages, as illustrated in Figure 4. Initial signal conditioning removes known transmission-dependent variations while preserving hardware signatures. Feature extraction algorithms then compute discriminative characteristics across multiple signal domains. Feature selection techniques, such as recursive feature elimination or principal component analysis, identify the most informative subset of extracted features. Finally, normalization and scaling ensure that features from different domains can be combined effectively for classification [27].

C. Application Domains and Traditional Approach Limitations

RF fingerprinting has been deployed across diverse wireless protocols and deployment scenarios, providing hardware-based device authentication that complements traditional cryptographic approaches. The cybersecurity context motivating these deployments spans the full protocol stack, from physical-layer RF attacks to application-layer vulnerabilities, as summarized in Table 4, which classifies the predominant threat vectors impacting UAV and IoT systems across OSI layers [3], [4]. Real-world incidents including the unauthorized acquisition of a U.S. military drone by Iran in 2011 and controlled WiFi-based UAV hijacking demonstrations [17] illustrate the operational stakes of these threats across heterogeneous wireless deployments.

Against this threat landscape, RF fingerprinting provides physical-layer authentication across diverse wireless protocols. Table 5 summarizes performance characteristics across protocols, demonstrating both broad applicability and significant variation in effectiveness across deployment scenarios.

WiFi Device Fingerprinting: WiFi applications have achieved remarkable success, with deep learning-based approaches demonstrating accuracy exceeding 99% for populations of 100+ devices [12], [29]. These systems can distinguish between devices from the same manufacturer and model, providing fine-grained identification capabilities that support network access control and device inventory management. Recent advances include CSI-based approaches achieving 99.53% accuracy using only four measurements, demonstrating efficiency improvements for real-world deployment [12].

Bluetooth Device Authentication: Bluetooth RF fingerprinting leverages hardware signatures to enhance pairing security and detect unauthorized device access attempts. Studies have shown that Bluetooth RF fingerprints remain stable across different environmental conditions and can achieve high identification accuracy even in the presence of interference from other wireless devices [30]. Advanced approaches using Hilbert-Huang Transform techniques have demonstrated 99.8% accuracy for populations of 10 devices [35].

LoRa Network Security: LoRa protocols benefit significantly from RF fingerprinting due to their emphasis on long-range, low-power communications. RF fingerprinting provides an additional security layer for LoRaWAN networks, enabling detection of rogue devices and preventing

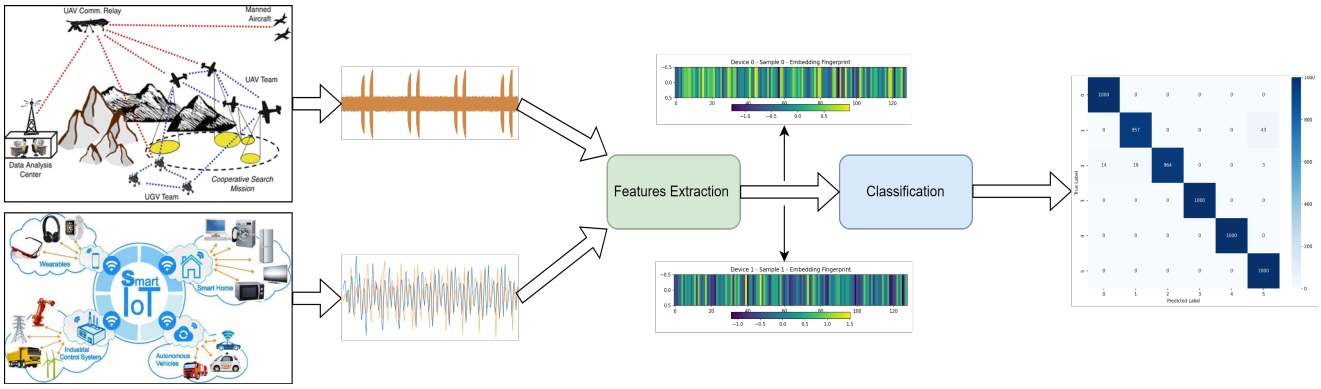


FIGURE 4. **RF Fingerprinting Methodology: General pipeline showing signal processing and classification approach with potential applications in UAV and IoT device authentication**

TABLE 4. **Threat Taxonomy for UAV and IoT Systems by OSI Layer**

OSI Layer	UAV Threats	IoT Threats
Physical	RF jamming, GPS spoofing, device capture, power drain attacks, hardware sabotage	Sensor tampering, electromagnetic interference, physical theft, side-channel attacks, battery exhaustion
Data Link	MAC spoofing, frame injection, collision attacks	MAC address spoofing, replay at link layer, eavesdropping, packet sniffing
Network	Routing attacks, IP spoofing, MitM, denial-of-service, black-hole/sinkhole attacks	Protocol exploits, DDoS, address spoofing, packet injection, misrouting
Transport	TCP/UDP flooding, port scanning, session hijacking	Connection hijacking, SYN flooding, buffer overflows, port exploitation
Session	Session key theft, replay of encrypted commands, disruption of authentication sequences	Broken authentication, credential hijacking, protocol state manipulation
Presentation	Payload obfuscation, format-specific attacks, encoding exploitation	Injection via malformed data formats, data serialization abuse, encryption vulnerabilities
Application	Command injection, malware payloads, firmware exploitation, unauthorized command and control	Unauthorized API access, data exfiltration, firmware backdoors, insecure update mechanisms

unauthorized network access. Performance evaluations have demonstrated accuracy levels approaching 95–99% for LoRa device identification, with some approaches maintaining stable performance across days when properly designed to account for environmental variations [36], [37].

ZigBee Device Identification: ZigBee fingerprinting has shown excellent results, with some studies reporting accuracy exceeding 95% under realistic SNR conditions. ZigBee's spread spectrum modulation (DSSS with O-QPSK) creates distinct hardware-dependent signal characteristics that enable reliable fingerprinting [31], [38]. Recent work has demonstrated robust performance even at low SNR levels (-5 dB), with open-set detection capabilities for identifying unknown devices [38].

UAV Control Link Security: UAV security applications leverage RF fingerprinting for controller authentication, rogue UAV detection, and command link validation. Controller authentication ensures that UAV command and control communications originate from authorized ground stations, providing hardware-level verification resistant to MAC address spoofing. Studies have demonstrated UAV controller identification accuracy exceeding 99% under controlled conditions, with robust approaches maintaining over 85% accuracy in challenging mobile scenarios with Doppler effects [39]. Advanced systems can distinguish between flight modes (hovering, flying) and detect interference, achieving 99.20% classification accuracy in realistic ISM-band environments [40].

Table 5 summarizes performance characteristics across wireless protocols, demonstrating both the broad applicability of RF fingerprinting and significant variation in effectiveness across deployment scenarios.

Traditional RF Fingerprinting and Generalization Failures: Prior to deep learning advances, RF fingerprinting relied on manual feature engineering combined with classical machine learning classifiers (SVM, k-NN, ensemble methods). Features were extracted from time, frequency, and transform domains, with approaches achieving 90–99%

TABLE 5. RF Fingerprinting Performance Characteristics by Protocol

Protocol/Method	Accuracy	Devices	SNR	Context & Key Characteristics
WiFi				
ORACLE [29]	99%	100+	10 dB	Controlled lab, I/Q samples
DeepCRF [12]	99.53%	19	–	CSI-based, 4 measurements only
Taşcıoğlu [41]	∼92%	20	15 dB	Near-Nyquist sampling
Bluetooth				
Ali et al. [30]	95%	50+	15 dB	Indoor, static conditions
Helluy-Lafont [35]	>99%	10	High	Packet averaging, real devices
Yuan et al. [42]	>75%	–	Low	Channel variations, BLE
LoRa				
Shen et al. [36]	95–96%	25	–	7-month CFO stability, indoor
Al-Shawabka [37]	91–99%	100+	–	Cross-day with augmentation
ZigBee				
Yu et al. [33]	100%	10+	20 dB	High SNR, controlled setup
Wang et al. [38]	75%	18	-5 dB	Low-SNR robust, open-set
Qing et al. [43]	91%	–	15 dB	Realistic SNR conditions
UAV Control				
Quadar et al. [39]	99%/85%	8+	12 dB	Static/Mobile with Doppler
Nemer et al. [40]	99.20%	Multiple	–	Flight modes, ISM interference
Ezuma et al. [44]	98.13%	15–17	25 dB	WiFi/Bluetooth interference

accuracy in controlled laboratory settings [27], [31], [32]. However, these traditional approaches encountered fundamental limitations when deployed in real-world conditions, ultimately motivating the transition to deep learning methods.

- **Manual Engineering Constraints:** The labor-intensive feature design process required substantial human expertise and time investment for each application domain. Domain experts needed deep understanding of both RF circuit design and signal processing principles to develop effective features [31]. Protocol-specific feature engineering prevented unified solutions, as features effective for one standard (e.g., ZigBee) often proved ineffective for others (WiFi, Bluetooth), requiring separate development efforts for each protocol and multiplying costs.
- **Scalability Limitations:** Performance degraded with large device populations as classification complexity grew exponentially. The curse of dimensionality affected high-dimensional feature spaces, where manually engineered features became less discriminative as device populations increased and feature overlap became more common. Computational complexity and memory requirements for storing high-dimensional feature vectors often exceeded capabilities of resource-constrained platforms [32].
- **Environmental Sensitivity:** Classification accuracy degraded significantly when SNR fell below controlled laboratory levels. Studies consistently demonstrated substantial performance drops below 10–15 dB SNR, with some approaches becoming unreliable under realistic noise conditions [28]. Channel variation effects from multipath propagation, fading, and interference altered signal characteristics that traditional features attempted to capture. Temperature and component aging caused hardware characteristics to drift over time, making trained classifiers less accurate without complete system retraining.
- **Real-World Deployment Gaps:** The controlled conditions used for traditional system development (stable SNR, minimal interference, static environments) rarely existed in operational scenarios. Performance differences between laboratory validation and field deployment often exceeded acceptable limits. Dynamic environments including mobile scenarios, varying interference patterns, and evolving device populations created conditions that exceeded the capabilities of static feature engineering approaches, leading to system failure rather than graceful degradation.

These fundamental limitations, particularly the scalability barriers, environmental sensitivity, and labor-intensive requirements, created compelling motivation for exploring deep learning approaches. The ability of deep neural networks to automatically learn discriminative features from raw data, while potentially providing better generalization capabilities and scalability, positioned deep learning as a promising solution. However, as Section III demonstrates, generalization challenges persist even with modern deep

learning methods, requiring continued research to bridge the gap between laboratory performance and practical deployment.

D. Advantages and Fundamental Characteristics

Despite the challenges discussed above, RF fingerprinting offers several fundamental advantages that distinguish it from traditional security approaches and motivate ongoing research efforts. These advantages stem from the hardware-based nature of RF fingerprints and their unique characteristics in wireless communication systems.

Cross-protocol applicability represents one of the most significant advantages. RF fingerprints manifest across different wireless protocols, though protocol-specific adaptations are typically required for optimal performance [31]. This enables security frameworks that can operate across diverse wireless technologies within integrated systems, whether dealing with WiFi, Bluetooth, LoRa, ZigBee, or proprietary control protocols.

Spoofing resistance emerges from the fundamental difficulty of replicating hardware characteristics. While digital identifiers such as MAC addresses can be easily modified or counterfeited, RF fingerprints are intrinsically tied to the physical properties of transmitter hardware [28]. The complex interaction of multiple hardware imperfections creates high-dimensional fingerprint signatures that would require precise replication of numerous analog components to forge successfully. Even sophisticated adversaries with access to detailed hardware specifications would find it extremely challenging to manufacture devices with identical RF characteristics to legitimate equipment.

Lightweight operation makes RF fingerprinting particularly suitable for resource-constrained environments. The computational requirements for fingerprint extraction and classification can be optimized for deployment on devices with limited processing power and energy budgets. Studies have demonstrated FPGA implementations achieving classification latencies as low as 22 microseconds while requiring only minimal memory resources [27]. This efficiency enables real-time authentication without imposing significant overhead on battery-powered devices.

Real-time capability supports online authentication scenarios where immediate device verification is required. Unlike cryptographic approaches that may require complex key exchange procedures or certificate validation processes, RF fingerprinting can provide authentication decisions within milliseconds of signal reception. This rapid response time is essential for applications such as UAV flight control where security decisions must be made quickly to maintain operational safety.

Complementary security allows RF fingerprinting to enhance rather than replace existing security mechanisms. RF fingerprints can serve as an additional authentication factor in multi-layered security architectures, providing hardware-level verification that complements cryptographic credentials and protocol-level security measures [29], [45], [46]. This layered approach creates redundant security barriers that significantly increase the difficulty of successful attacks while maintaining compatibility with existing security infrastructure and does not require hardware upgrades [47].

The fundamental characteristics of RF fingerprinting also include inherent robustness properties that make it suitable for challenging deployment environments. Studies have demonstrated that properly designed fingerprinting systems can maintain accuracy levels above 90% at SNR levels above 10 dB, though performance degrades under more severe conditions [28]. However, it is important to note that performance can degrade significantly in extreme conditions, highlighting the importance of robust system design and appropriate deployment considerations.

The scalability characteristics of RF fingerprinting vary depending on the specific implementation approach. Traditional machine learning methods can handle moderate device populations efficiently, while deep learning approaches have demonstrated the ability to scale to populations of thousands of devices [13], [48]. However, computational requirements and training data needs generally increase with the size of the device population, requiring careful system design for large-scale deployments.

While RF fingerprinting demonstrates significant potential across diverse applications with impressive performance in controlled environments, practical deployment faces substantial challenges that limit real-world effectiveness. The performance characteristics presented above typically assume static, controlled conditions with stable environmental parameters and consistent transmission settings. As Section III explores in detail, maintaining these performance levels becomes significantly more challenging when devices operate across varying transmission parameters, temporal conditions, and dynamic environments that characterize real-world deployments. Understanding these generalization challenges is essential for developing RF fingerprinting systems capable of reliable operation beyond laboratory settings.

III. The Generalization Challenge: Core Problem

This section examines the central challenge that motivates our survey: the persistent generalization failure of RF fingerprinting systems when deployed beyond controlled laboratory conditions. While Section II established that RF fingerprinting can achieve impressive accuracy exceeding 95–99% in controlled settings [12], [13], this section demonstrates why these systems consistently fail in real-world environments, with performance degrading to 30–70% accuracy [14]–[16] across five critical dimensions: cross-transmission variations, temporal drift, environmental sensitivity, mobility effects, and scalability limitations. Understanding why generalization fails, not merely that it fails, is essential for developing effective solutions. We organize our analysis around root causes and physical mechanisms,

comparing contradictory results and identifying fundamental versus solvable limitations.

A. Defining the Generalization Problem

The generalization challenge constitutes the fundamental barrier preventing practical deployment of RF fingerprinting technology. Despite achieving 95–99% accuracy in controlled laboratory conditions, systems consistently degrade to 30–70% accuracy when deployment conditions differ from training environments. Generalization refers to maintaining reliable device identification accuracy across variations in transmission parameters, environmental conditions, temporal factors, device mobility, and population scaling. Unlike traditional machine learning where generalization involves new samples from the same distribution, RF fingerprinting must maintain performance across fundamentally different signal propagation conditions and hardware operating states [49], [24].

The controlled versus real-world performance gap represents the most visible manifestation of generalization failures. Laboratory studies consistently report accuracy levels approaching theoretical limits, with systems achieving 99.53% accuracy for WiFi device identification [12] and 95.5% accuracy for populations exceeding 10,000 devices [13]. However, when these same systems are deployed in operational environments, performance typically degrades to 30–70% accuracy, often falling below the threshold required for practical security applications. This dramatic performance collapse indicates fundamental limitations in current approaches rather than minor implementation issues.

Root Cause of Performance Collapse: This degradation stems from a fundamental mismatch between how RF fingerprinting systems learn and how wireless devices operate. Deep learning models optimize for statistical patterns in training data, learning to recognize specific environmental conditions and transmission parameters rather than isolating hardware-invariant characteristics [50], [51]. When deployment conditions differ, learned statistical associations break down. Models memorize training-specific correlations (e.g., "Device A transmits at frequency F with channel response C") rather than extracting hardware signatures ("Device A's power amplifier exhibits nonlinearity N"). This explains why sophisticated architectures fail to generalize: they solve the wrong optimization problem.

Performance degradation patterns exhibit remarkable consistency across different studies, architectures, and applications, suggesting fundamental rather than implementation-specific limitations:

- **Cross-day testing**: Typically reveals accuracy drops of 20–60% compared to same-day evaluation.
- **Cross-location deployment**: Often results in performance degradation exceeding 70%.
- **Dynamic environmental conditions**: Can cause complete system failure in extreme cases.
- **Multi-factor interactions**: Compound effects when multiple variations occur simultaneously.

These consistent patterns indicate that generalization challenges represent systematic limitations that require novel technical approaches rather than incremental improvements to existing methods [52], [53].

Table 6 provides quantitative evidence of the severity and consistency of these generalization failures across different challenge dimensions, demonstrating performance degradation ranging from 12.4% to over 85% depending on the specific variation encountered.

TABLE 6. Quantitative Evidence of RF Fingerprinting Generalization Failures

Challenge Type	Controlled	Real-World	Degradation
Cross-transmission [54]	92.78%	32.05%	60.7%
Environmental (NLOS static) [33]	97%	84.6%	12.4%
Temporal (cross-day) [49]	98%	80%	18%
Spatial (distance) [15]	97%	16%	81%
Mobility (Doppler) [39]	99.6%	85.8%	13.8%
Environmental (A2G) [14]	95.9%	55.5%	40.4%
Cross-channel [55]	>90%	>5%	>85%

B. Cross-Transmission Generalization Failures

Having established the severity of generalization failures, we now examine their manifestation across transmission parameter variations, one of the most severe and well-documented failure categories. These failures manifest across multiple transmission parameter dimensions, creating a complex challenge that has proven resistant to traditional machine learning approaches.

Cross-transmission generalization failures manifest across multiple transmission parameter dimensions:

- **Frequency Channel Variations:** Single-channel models achieve MCC <0.9 on training channels but fail completely (MCC >0.05, random-guess) on distant channels [55]. Hardware components (PLLs, amplifiers, filters) exhibit frequency-dependent characteristics.
- **Power Level Changes:** Amplifier compression varies with transmission power; oscillator stability exhibits power-dependent thermal effects; dynamic power control adds complexity.
- **Modulation Scheme Differences:** 70% accuracy drop when QAM-trained models test on FSK signals (average 32.05% vs. over 90% within training modulation) [54]. Different schemes emphasize different hardware imperfections.

- **Bandwidth Modifications:** RF hardware exhibits bandwidth-dependent transfer functions (filters, amplifiers, DSP), causing feature extraction failures.

The root cause lies in frequency-dependent and power-dependent hardware characteristics. When a neural network trains on signals at frequency F1, it learns "how this device's amplifier behaves at F1" rather than learning a general representation of the device's power amplifier. At F2, the same hardware produces different distortion patterns due to transfer function variations. Multi-channel training partially addresses this by exposing the model to hardware behavior across multiple operating points, enabling interpolation but not true invariance [55], [54]. Quantitative evidence reveals severe failures: a standard raw-IQ classifier for 16 bit-similar USRP X310 radios achieved 98.60% accuracy in static conditions but collapsed to 35.96% when tested at a different location without controlled impairment injection [29], with multi-channel studies showing random-guess performance on distant frequency channels.

Cross-protocol limitations extend beyond single-protocol parameter variations to encompass fundamental differences between communication standards. Systems optimized for WiFi device identification typically fail when applied to Bluetooth, LoRa, or ZigBee devices, even when the underlying hardware may share similar components [23]. This protocol specificity stems from differences in signal structure, timing characteristics, and spectral occupancy that affect how hardware imperfections manifest in transmitted signals.

C. Temporal and Environmental Variation Challenges

Beyond transmission parameter variations, temporal and environmental factors introduce pervasive degradation that affects systems across extended deployment periods and diverse operating conditions. These challenges manifest as performance degradation over time scales ranging from hours to months, with environmental factors creating additional complexity that can cause immediate system failure when conditions change unexpectedly.

Temporal degradation patterns exhibit consistent characteristics across deployment scenarios, revealing fundamental limitations in the stability of learned fingerprint representations. The most commonly documented challenge is the same-day versus cross-day performance gap, with systems achieving 98% accuracy on same-day data but degrading to 80% when tested across days [49]. This degradation accelerates over extended periods: long-term stability studies have documented progressive accuracy decline over 5+ months due to systematic hardware aging and environmental drift [24], making continuous adaptation or periodic retraining a practical necessity for any real-world deployment.

Two distinct physical processes drive temporal failures: (1) irreversible hardware aging (thermal cycling, electromigration, mechanical stress) causing permanent parameter shifts, and (2) reversible environmental variations (temperature, humidity, power fluctuations) causing temporary changes [56], [57]. Current deep learning models cannot distinguish these processes, observing feature drift without determining whether it reflects permanent aging (requiring model updates) or temporary variation (requiring environmental compensation). This explains why periodic retraining only partially addresses drift: it accommodates aging but not environmental robustness. Approaches modeling these separately show improved stability [24], [49].

Beyond temporal factors, environmental sensitivity introduces immediate and severe performance degradation whenever operating conditions diverge from training environments. Temperature fluctuations, humidity, and weather variations can render systems ineffective, while location-dependent effects create spatial generalization challenges that interact with these factors to produce compound degradation: cross-location testing has documented accuracy drops from 97% to 16% when devices move from 1-meter to 2-meter distances, with performance collapsing to random-guess levels at 3 meters [15]. These compounding spatial and environmental sensitivities can alter the electromagnetic environment suddenly and without warning, making proactive adaptation particularly difficult and highlighting the fragility of fingerprint representations learned under static conditions.

D. Mobility and Dynamic Environment Challenges

Mobility introduces the most complex category of generalization failures, where Doppler effects, dynamic multipath, and time-varying channel conditions compound to create severe performance degradation. These challenges emerge when devices or receivers move during operation, when channel conditions change dynamically, or when systems must operate in environments with varying interference patterns and propagation characteristics.

Static versus mobile performance gaps reveal the immediate impact of movement on fingerprinting reliability. Even without full mobility, environmental degradation from line-of-sight to non-line-of-sight static conditions alone reduces accuracy from 97% to 84.6% [33]. Under genuine mobility with 100 Hz Doppler shift, systems maintaining 99.6% accuracy in stationary conditions degrade to 85.8% [39], with severe cases dropping under unpredictable movement patterns. Doppler shifts introduce systematic frequency offsets that mask subtle spectral characteristics, while high-speed scenarios such as UAVs traveling at 74.2–92 m/s create feature distribution shifts that further compound classification errors [14].

Dynamic channel conditions add a second layer of complexity beyond Doppler sensitivity. Multipath propagation patterns change continuously as objects move within signal paths, shadow fading introduces random signal strength variations that destabilize extractable features [58], [59], and fast fading creates rapid amplitude and phase fluctuations that static training data cannot anticipate. These effects interact with device orientation changes, distance variations, and power control responses to produce compound frequency-

time distortions that no single compensation strategy can fully address.

Mobility failures stem from the time-varying nature of the wireless channel, which violates the static assumptions underlying most fingerprinting approaches. Doppler shifts introduce frequency-domain distortions that are indistinguishable from hardware-specific spectral characteristics when analyzed with static models. A 100 Hz Doppler shift at 2.4 GHz creates frequency offsets that can mask or mimic the subtle phase noise signatures used for device discrimination. Furthermore, dynamic multipath creates time-varying amplitude and phase distortions that traditional feature extraction cannot separate from hardware effects. Successful approaches require explicit temporal modeling, using LSTM or attention mechanisms to track how channel conditions evolve while maintaining representations of time-invariant hardware characteristics [39], [60]. The 99.6% $\rightarrow$ 85.8% degradation under mobility is substantially better than static models, demonstrating that temporal modeling partially addresses the challenge but does not fully solve it.

E. Scalability and Unknown Device Detection Limitations

RF fingerprinting systems face fundamental scalability limitations as device populations grow and encounter unknown devices not present during training. These limitations manifest across three dimensions: population scaling degradation, open-set detection failures, and computational barriers that constrain real-world deployment.

Empirical analyses of scaling behavior reveal a consistent degradation trend with increasing device population. Identification accuracy typically exceeds 90% for small-scale deployments of 10–50 devices, declining to 70–80% for populations exceeding 100 devices, with further attenuation approaching 1,000 devices. Even systems demonstrating 95.5% accuracy at 10,000 devices [13] operate under controlled laboratory conditions that may not reflect deployment diversity.

The scalability limitations arise from the finite discriminative capacity of learned feature representations in high-dimensional classification problems. As device populations grow beyond hundreds of devices, distinguishing between increasingly similar hardware characteristics becomes fundamentally more difficult, leading to inevitable feature space overlap [51]. This creates inevitable overlap as devices with similar hardware characteristics become harder to distinguish. The problem is intensified by manufacturing consistency improvements: modern devices from the same production batch exhibit more similar hardware characteristics than older devices, reducing natural separation in feature space. Successful large-scale systems employ hierarchical classification strategies that partition the device population into coarse groups before fine-grained discrimination, effectively creating multiple smaller feature spaces [13].

Open-set recognition exposes another fundamental assumption failure in current approaches. Closed-set systems perform well when test devices are drawn from the training population, but real deployments inevitably encounter unknown devices that should be rejected rather than misidentified [38]. Current systems struggle to distinguish between known devices operating under varying conditions and genuinely unknown devices, a challenge compounded by the fact that temporal drift and device novelty produce similar feature distribution shifts, making principled rejection difficult without explicit open-set mechanisms.

Computational scaling barriers create practical constraints that limit deployment at operational scales. Memory requirements for feature vectors and model parameters grow linearly or super-linearly with device population, quickly exceeding embedded or edge platform capabilities [61], [62]. Processing requirements for real-time classification become prohibitive beyond thousands of devices, and storage requirements for training data across multiple operating conditions create logistical burdens that scale poorly with deployment size. These constraints collectively make it difficult to justify RF fingerprinting infrastructure costs relative to established alternatives such as cryptographic authentication, which offer more predictable scaling behavior.

F. Synthesis and Research Gap Identification

Apparent contradictions in reported generalization performance across studies can be reconciled by examining experimental conditions carefully. Cross-day temporal degradation ranges from approximately 18% under controlled collection conditions [49] to over 40% for statistical features under varying environmental conditions [16], with this variation explained by differences in collection time gaps (hours versus weeks), environmental stability (climate-controlled versus field conditions), and feature extraction methodology. Similarly, mobility-induced degradation varies from 6–7% under line-of-sight mobile conditions to 12–14% under high-Doppler scenarios with dynamic multipath [39], demonstrating that experimental setup profoundly affects measured generalization performance. These dependencies are essential context for interpreting performance claims across the literature and for predicting behavior in specific deployment scenarios.

Across all challenge dimensions examined in this section, current approaches share three fundamental limitations regardless of architectural choices. First, deep learning models optimize for statistical correlations in training data rather than causal hardware characteristics, making them sensitive to any condition shift. Second, no current approach can simultaneously accommodate cross-transmission variation, temporal drift, environmental sensitivity, mobility effects, and population scaling, each solution addresses its target challenge while remaining vulnerable to the others. Third, these challenges interact: a system tolerating moderate temporal drift may fail completely when drift occurs simultaneously with mobility and environmental variation, producing compound degradation that exceeds the sum of individual

factor impacts. Figure 5 illustrates these compound interaction patterns across the five generalization dimensions.

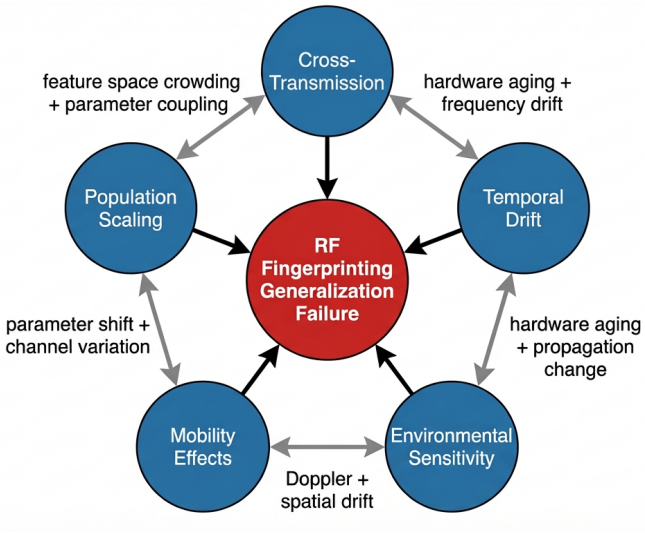


FIGURE 5. Compound interaction patterns among the five RF fingerprinting generalization challenge dimensions. Bidirectional edges indicate documented co-occurrence effects that produce degradation exceeding individual factor impacts. The central node represents the aggregate generalization failure outcome that motivates the solution frameworks surveyed in subsequent sections.

Table 7 provides a structured synthesis of all five challenge dimensions, their root causes, current solution approaches, and research priorities. Three challenges are rated critical, cross-transmission, environmental sensitivity, and multi-factor interactions, reflecting both the severity of their performance impact and the absence of robust general solutions in current literature.

The economic dimension of generalization failures creates deployment barriers that extend beyond technical performance metrics. Systems requiring frequent retraining or manual adaptation incur substantial ongoing operational costs, while performance uncertainty under real-world conditions limits the security guarantees that can be provided to end users. These factors make RF fingerprinting difficult to justify economically relative to established alternatives such as cryptographic authentication, which offer more predictable scaling behavior, unless generalization limitations are resolved sufficiently to provide reliable performance across anticipated operating conditions.

Addressing these limitations requires approaches that move beyond incremental architectural improvements. Physics-informed machine learning, causal inference frameworks, and adaptive system architectures represent the most promising directions [63], [64], as they target the root cause, statistical correlation learning, rather than its symptoms. The challenge interaction patterns identified here suggest that effective solutions must be holistic: a method resolving temporal drift in isolation will remain vulnerable to the compound failures that dominate real-world deployments. The following sections survey the state of these solution directions and assess their progress against the generalization requirements established here.

IV. Deep Learning Approaches to Address Generalization Challenges

Having established the generalization failure dimensions in Section III, this section examines how deep learning approaches have attempted to address them. We trace the transformation from manual feature engineering to end-to-end learning, analyze architectural innovations targeting specific generalization dimensions, explore domain adaptation and augmentation strategies, and critically assess what has been achieved versus what remains unsolved.

A. The Deep Learning Revolution: From Feature Engineering to End-to-End Learning

Deep learning transformed RF fingerprinting by replacing manual feature engineering with automatic feature discovery [50], [51]. Emerging in the mid-2010s, this shift addressed the scalability, adaptability, and performance constraints of traditional approaches. The convergence of large-scale datasets, the DARPA RFMLS dataset provided more than 400 GB from 10,000+ devices [51], [65], and GPU computational advances made intensive neural network training practically feasible, democratizing RF fingerprinting research and accelerating innovation across device types and protocols.

Early breakthroughs established the foundational capabilities driving widespread adoption. Riyaz et al. demonstrated that a 4-layer CNN achieves 98% accuracy for 5-device identification from raw I/Q samples, dramatically outperforming SVM at 33% on the same task [66], proving that networks could learn hardware-specific imperfections without manual feature engineering. ORACLE extended this to population scale, achieving 98.69% accuracy for 140 WiFi devices and 99.70% for 16 bit-similar USRP X310 radios through controlled impairment introduction [29], demonstrating that deep learning could distinguish between nominally identical devices where traditional methods failed entirely.

End-to-end learning capabilities enabled direct processing of raw I/Q samples, preserving all signal information and allowing networks to discover subtle hardware-specific characteristics, I/Q imbalance, carrier frequency offset, power amplifier nonlinearities, phase noise, without explicit programming [13]. Performance improvements were consistent and dramatic across studies: RiffNet demonstrated 95.5% accuracy at 10,000-device scale previously unattainable with traditional methods [13], and DeepCRF achieved 99.53% accuracy using only 4 CSI measurements [12]. Table 8 summarizes these milestones.

Despite these achievements, controlled-environment performance consistently failed to transfer to real-world conditions, with 20–40% accuracy drops under dynamic channel conditions, cross-day testing, and device mobility. These persistent generalization failures motivated the specialized

TABLE 7. RF Fingerprinting Generalization Challenges: Root Causes, Impact, and Research Priorities

Challenge	Performance Drop	Root Cause	Current Solutions	Research Requirements	Priority
Cross-transmission	65–85%	Parameter-hardware coupling	Multi-channel training, data augmentation	Physics-informed feature extraction, causal inference mechanisms	Critical
Temporal drift	20–60%	Hardware aging, environmental drift	Periodic retraining, adaptation algorithms	Continual learning frameworks, aging-aware models	High
Environmental sensitivity	40–80%	Propagation condition changes	Environmental modeling, robust features	Adaptive feature extraction, environment-invariant representations	Critical
Mobility effects	15–40%	Doppler shifts, multipath variations	Real-time processing, motion compensation	Dynamic adaptation mechanisms, temporal modeling	High
Population scaling	Variable	Feature space crowding, computational limits	Progressive learning, hierarchical approaches	Efficient architectures, scalable algorithms	Medium
Unknown devices	N/A	Open-set recognition limitations	Threshold-based rejection, confidence scores	Advanced anomaly detection, open-set learning	High
Multi-factor interactions	Compound	Simultaneous variation sources	Limited holistic approaches	Integrated adaptation frameworks, causal decomposition	Critical

TABLE 8. Deep Learning Performance Milestones in RF Fingerprinting

System	Architecture	Devices	Accuracy	Year
Riyaz et al. [66]	4-layer CNN	5	98%	2018
ORACLE [29]	CNN	140	98.69%	2019
MSCNN [33]	Multi-scale CNN	54	97%	2019
RiffNet [13]	Dilated Conv	10,000	95.5%	2021
DAConv [67]	CNN+Attention	10	95.6%	2023
DeepCRF [12]	CNN+Contrastive	19	99.53%	2025

architectural and learning-based solutions examined in the following subsections.

B. Architectural Solutions for Generalization Challenges

Building on end-to-end learning capabilities, researchers developed specialized architectures targeting specific generalization dimensions: CNNs for cross-transmission robustness, LSTMs for temporal stability, attention mechanisms for dynamic scenarios, and transformers with autoencoders for scalability.

Convolutional Neural Networks for Cross-Transmission Robustness: CNNs emerged as the dominant architecture for RF fingerprinting due to their natural alignment with signal processing principles and ability to learn spatially-invariant features across different signal conditions. Unlike manual feature engineering requiring separate designs for each transmission parameter, CNNs learn hierarchical representations capturing hardware-specific characteristics while maintaining partial invariance to frequency, power, and modulation variations, enabling multi-channel training where a single network handles multiple frequency channels or modulation schemes simultaneously. Multi-scale approaches such as MSCNN employed parallel branches with different downsampling factors to capture both short-term transients and long-term statistical patterns [33], while ResNet adaptations addressed the vanishing gradient problem for deeper networks [68], [69]. ORACLE achieved 98.69% accuracy for 140 WiFi devices [29] and DeepCRF demonstrated 99.53% accuracy using only 4 CSI measurements [12]. However, cross-transmission generalization remains limited when frequency gaps exceed approximately 100 MHz or modulation schemes differ fundamentally (QAM vs. FSK), as CNNs still learn transmission-specific patterns that fail to transfer across large parameter variations.

LSTM Networks for Temporal Dependencies: While CNNs excel at local feature extraction, LSTM networks provide the memory mechanisms necessary to capture long-term dependencies and temporal evolution patterns in transmitted signals [70]–[72]. Gating mechanisms allow selective retention of stable hardware-induced patterns while filtering transient channel effects, particularly valuable for cross-day stability where device characteristics drift gradually over hours or days. Hybrid CNN-LSTM architectures combine spatial feature extraction with temporal modeling for sce-

narios involving both transmission variations and temporal dynamics. Quadar et al. developed a CNN-LSTM-Attention architecture achieving 99.6% accuracy in stationary conditions and 85.8% under mobility conditions with 100 Hz Doppler shift [39], demonstrating that explicit temporal modeling substantially mitigates mobility-induced failures. Long-term temporal drift exceeding 6 months remains problematic, however, as finite LSTM memory cannot maintain arbitrarily long dependencies without performance degradation.

Attention Mechanisms for Dynamic Scenarios: Attention mechanisms address mobility-induced generalization failures by enabling networks to dynamically focus on temporal segments where hardware characteristics remain stable despite channel variations. In mobile scenarios, Doppler effects and dynamic multipath create time-varying distortions that mask hardware fingerprints during certain periods; multi-head attention down-weights these artifact-dominated periods while emphasizing segments where device-specific features are prominent [21], [67]. The Dual Attention Convolution (DAConv) module demonstrated this principle by integrating channel and spatial attention into convolutional layers, achieving 95.6% accuracy while providing interpretable attention weights revealing which signal features drive classification decisions [67]. Performance degrades substantially at SNR below 5 dB where signal quality is insufficient for reliable attention weight computation.

Transformers and Autoencoders for Scalability: Transformer architectures address scalability limitations through parallel processing and explicit modeling of device-to-device relationships via self-attention, enabling hierarchical device groupings that reduce effective classification complexity as population size grows. The MDAE-Transformer architecture demonstrated this scalability potential, achieving 99.9% accuracy under optimal conditions, 92.9% cross-day accuracy, representing only 7% degradation and a 78% improvement over statistical feature baselines, and 75.2% accuracy at 80-device scale [73]. Autoencoder approaches complement transformers by addressing open-set recognition through reconstruction-based feature learning: devices seen during training reconstruct accurately while unknown devices produce elevated reconstruction errors, enabling threshold-based rejection and addressing the closed-set assumption that limits most deep learning approaches [49], [74], [75]. Transformer architectures require substantial memory with quadratic scaling in sequence length, making edge deployment difficult without model compression, while autoencoders maintain moderate computational requirements offering better edge feasibility.

C. Domain Adaptation, Transfer Learning, and Data Augmentation

Learning-based approaches address generalization through domain adaptation, synthetic data augmentation, and progressive adaptation strategies that enable systems to leverage prior knowledge while adapting to new conditions.

Clarifying Few-Shot, Transfer, and Domain Adaptation: These distinct methodologies are frequently conflated in the literature. *Few-shot learning* employs meta-learning frameworks that train models to adapt quickly to new device classes using minimal samples per device, learning a general adaptation strategy rather than device-specific features [52], [76]. *Transfer learning* pre-trains a model on a source domain (e.g., WiFi devices) then fine-tunes on a target domain (e.g., Bluetooth devices), directly transferring learned weights [77], [78]. *Domain adaptation* minimizes statistical divergence between source and target feature distributions without requiring target labels, typically using maximum mean discrepancy or adversarial training. Understanding these distinctions is critical for selecting appropriate approaches for specific deployment scenarios.

Domain Adaptation for Cross-Transmission Generalization: Domain adaptation bridges the performance gap between controlled training and diverse deployment environments by learning transferable representations effective across signal domains. These approaches are most effective when source and target domains share underlying hardware characteristics but differ in transmission parameters or environmental conditions [14]. However, domain adaptation struggles when hardware architectures differ fundamentally across chipset families, or when the domain shift involves multiple confounding factors simultaneously, limiting practical applicability in heterogeneous deployment scenarios [15].

Data Augmentation and Generative Models: Data augmentation improves generalization by artificially expanding training datasets with synthetic variations simulating diverse operating conditions. Physics-based augmentation strategies that apply realistic frequency shifts based on channel models, such as the Doppler provide effective mobility robustness without the computational cost of generative model training [39]. Generative approaches offer higher-fidelity synthesis, Wen et al. developed a diffusion model framework for WiFi CSI augmentation preserving essential temporal and spectral characteristics [79], while RF-Diffusion adapted diffusion models through Time-Frequency Diffusion theory capturing time, frequency, and complex-valued signal domains [80]. However, effective pre-training requires datasets of large signal samples, and iterative denoising inference introduces latency that limits use to offline augmentation with edge deployment infeasible [80], [81]. Critically, studies show that synthetic data ratios beyond approximately one-third of the training set can cause performance degradation: synthetic signals cannot perfectly capture real-world variations, and over-reliance introduces generation artifacts [80]. Augmentation is most effective when it accurately models deployment variations, physics-based augmentation succeeds because it simulates real physical phenomena, but cannot introduce device types or environmental conditions absent from the base dataset.

D. Critical Analysis and Current Limitations

Table 9 synthesizes the deep learning approaches examined in Subsections B and C, evaluating each across performance, key strengths, limitations, computational requirements, and edge deployment feasibility.

What Deep Learning Has Achieved: Deep learning has delivered genuine progress across several dimensions. Controlled environment accuracy consistently reaches 95–99% [12], [13], eliminating the manual feature engineering bottleneck and enabling scalability to device populations of 100–10,000 units that were unattainable with traditional methods. Partial solutions have emerged for the most critical generalization dimensions: cross-transmission approaches now achieve 60–86% accuracy under varying transmission parameters [54], [55], temporal stability has improved to 80–93% cross-day performance [49], and attention-based temporal modeling achieves 85–94% accuracy under mobility conditions where traditional methods failed completely [39].

What Remains Fundamentally Unsolved: Despite this progress, critical challenges prevent robust real-world deployment [50], [82], [83]. Multi-factor compound effects remain the most severe unsolved problem: when temporal drift, environmental variations, and mobility occur simultaneously, as they routinely do in operational environments, performance degrades catastrophically beyond the sum of individual factor impacts, and no current approach effectively handles multiple simultaneous generalization dimensions. Long-term temporal drift exceeding 6 months remains poorly handled, with statistical feature approaches degrading to 52.1% cross-day accuracy [16], indicating that hardware aging over extended periods exceeds current adaptation capabilities. True cross-domain generalization, models trained on WiFi failing completely on Bluetooth or LoRa without retraining, remains unsolved despite shared underlying hardware characteristics. Finally, real-world validation is persistently absent: most reported results come from controlled laboratory datasets, and performance under unpredictable operational interference and uncontrolled user behavior remains largely undocumented.

Why Certain Approaches Excel for Specific Dimensions: CNNs succeed at cross-transmission because convolutional kernels learn spatially-invariant features that partially transfer across frequencies and modulation schemes, though this invariance is incomplete. LSTMs address temporal challenges through gating mechanisms enabling selective memory of stable hardware patterns while filtering transient effects, but finite memory limits effectiveness for very long-term drift. Attention mechanisms address mobility by identifying temporal segments where hardware characteristics dominate channel effects, though this requires sufficient SNR for reliable weight computation [69]. Autoencoders enable open-set recognition through reconstruction error that naturally differentiates learned from unlearned patterns, but temporal drift increases reconstruction errors even for known devices, degrading rejection accuracy over time.

Fundamental Trade-offs: Several inescapable trade-offs constrain achievable solutions. The performance–computational cost relationship is consistently inverse: the most accurate approaches (transformers, diffusion models, hybrid architectures) require resources that preclude edge deployment, while edge-feasible approaches (CNNs, autoencoders) sacrifice accuracy under challenging conditions. Approaches optimized for specific generalization dimensions sacrifice performance on others, requiring careful matching to deployment scenarios. Continual learning systems face an adaptation speed versus stability tension: fast adaptation risks catastrophic forgetting of known devices while stable systems respond slowly to new conditions [84], [85]. Synthetic data augmentation faces a quality versus distribution shift limit: higher-fidelity generative models improve realism at exponentially higher cost, while efficient augmentation introduces artifacts constraining synthetic data to approximately one-third of training sets [80]. These trade-offs collectively indicate that no single approach will solve all generalization challenges, practical systems will require principled combinations of complementary techniques matched to specific deployment constraints.

V. Future Directions and Research Gaps

While Sections III and IV have established the fundamental generalization challenges and current solution approaches, significant research gaps remain that prevent practical deployment of RF fingerprinting systems in real-world heterogeneous wireless networks. This section identifies critical future directions by examining publicly available datasets suitable for generalization research, exploring potential advances in foundation models and large-scale pre-training, synthesizing persistent research gaps from our analysis, and discussing integration with emerging wireless technologies. Understanding what resources exist, what remains technically unsolved, and what new opportunities emerging technologies present is essential for guiding future research toward practical RF fingerprinting deployment.

A. Public Datasets for Generalization Research

Generalization research relies critically on diverse, high-quality datasets that capture the operational complexity of real-world wireless deployments, yet dataset availability has historically limited progress in RF fingerprinting research.

Table 10 provides an overview of publicly available RF fingerprinting datasets suitable for generalization research, supporting benchmarking and comparison across studies.

WiSIG - Large-Scale WiFi Dataset: The WiSIG dataset represents the largest publicly available collection for WiFi RF fingerprinting, encompassing 174 off-the-shelf transmitters and 41 USRP receivers with 10 million captured packets across four temporal sessions over one month [48]. The dataset's unprecedented scale (1.4 TB raw, 76.9 GB processed) enables evaluation of unknown device detection and cross-receiver generalization under realistic deployment

TABLE 9. Comparison of Deep Learning Approaches to RF Fingerprinting Generalization

Approach	Addresses	Best Performance	Key Strength	Main Limitation	Compute	Edge?
Multi-Channel CNN	Cross-transmission	99.53% [12]	Frequency-invariant features	Large parameter gaps fail	Medium	Yes
CNN-LSTM Hybrid	Temporal + Mobility	99.6%; 85.8% [39]	Spatial + temporal modeling	High computational cost	High	Difficult
Attention Mechanisms	Dynamic scenarios	95.6% [67]	Selective discriminative focus	Needs clean signals (SNR >5 dB)	Medium	Yes
Transformer based	Scalability + Cross-day	92.9% cross-day [73]	Parameter-efficient encoding, open-set capable	Edge compression required	Medium	Partial
Autoencoders	Open-set recognition	98% [49]	Unknown device detection via reconstruction error	Cross-day temporal drift	Medium	Yes
Statistical Features	Controlled accuracy	99.6% [16]	Interpretable, efficient	52.1% cross-day degradation	Low	Yes
Domain Adaptation	Cross-scenario (G2G & A2G)	89.1% [14]; 60–70% cross-location [15]	Leverages prior knowledge	Requires shared hardware characteristics	High	Difficult
Data Augmentation	All dimensions	91–99% cross-day [37]	Low additional cost; no target domain labels needed	Over-reliance on synthetic signals	Low–Med	Yes
Diffusion Models	Data scarcity	SSIM >0.9 [80]	Highest synthesis quality	100+ GPU-hrs, inference 0.5–10s	Very High	No

scenarios. The multi-receiver structure addresses critical limitations in existing datasets that typically employ single receiver configurations, enabling assessment of receiver-agnostic fingerprinting approaches essential for practical deployment.

LoRa IoT Dataset: The LoRa-60 dataset provides temporal coverage across 60 commercial LoRa devices representing four device categories and three chipset types (SX1272, SX1276, SX1261) [58]. Collections spanning multiple time scales from same-day variations through monthly aging effects enable evaluation of temporal degradation patterns and long-term stability challenges. Environmental diversity including indoor/outdoor scenarios, line-of-sight and non-line-of-sight conditions, and interference scenarios provides realistic assessment of fingerprinting robustness under operational IoT deployment conditions.

UAV Controller Datasets: Two complementary UAV datasets provide controlled evaluation of cross-transmission generalization through parameter variations. The 8-device dataset employs semi-anechoic chamber conditions eliminating multipath effects for controlled baseline assessment [87], while the 17-device dataset provides high-resolution capture (20 GSa/s) across diverse manufacturers in realistic indoor environments [86]. Both datasets enable evaluation of transmission parameter variations including carrier frequency, power levels, and modulation configurations essential for assessing cross-transmission robustness.

WiFi and Bluetooth Combo Datasets: The GLOBECOM22 dataset addresses multi-protocol fingerprinting evaluation through WiFi-Bluetooth combo chipsets where protocols share RF circuitry [21]. Cross-day temporal collections enable evaluation of temporal degradation in multi-protocol contexts. The ACMWiSec21 dataset provides cross-day evaluation with controlled two-day collection structure enabling fine-tuning recovery assessment [77].

Critical Gaps in Current Dataset Availability: Despite these valuable public resources, significant gaps remain that limit generalization research.

Scale and Diversity: While WiSIG provides 174 devices, enterprise and metropolitan deployments require evaluation at 500–10,000+ device scales. The DARPA RFMLS dataset addresses scale but has limited public accessibility, creating a barrier to reproducible large-scale research.

Long-Term Temporal Coverage: Existing public datasets provide temporal coverage from hours to one month. Evaluation of hardware aging effects and environmental drift over 6–12+ months remains unavailable in public datasets, limiting assessment of long-term stability approaches.

TABLE 10. Public RF Fingerprinting Datasets for Generalization Research

Dataset	Devices	Protocol	Availability	Generalization Challenges	Reference
WiSIG	174 Tx, 41 Rx	WiFi 802.11	Public	Unknown device detection, Cross-receiver, Large-scale open-set	Hanna et al. 2022 [48]
LoRa-60	60	LoRa IoT	Public	Temporal drift, Long-term aging, Environmental variations	Shen et al. 2022 [58]
UAV-17	17	UAV Control	Public	Cross-transmission (Tx1–Tx19), Scalability, High-resolution	Ezuma et al. 2020 [86]
UAV-8	8	UAV Control	Public	Cross-transmission (Tx1–Tx19), Controlled environment	Basak et al. 2022 [87]
GLOBECOM22	10	WiFi + Bluetooth	Public	Multi-protocol combo chipsets, Cross-day temporal	Jagannath et al. 2022 [21]
ACMWiSec21	5	WiFi 802.11	Public	Cross-day temporal, Fine-tuning recovery	Li et al. 2021 [77]
DARPA RFMLS	10,000+	Multiple	Limited	Large-scale population, Protocol diversity	Jian et al. 2020 [51]

True Mobile Scenarios: While the LoRa dataset provides channel variations, datasets capturing signals under actual device mobility with verified Doppler characteristics are limited. Most mobility assessments rely on static collection with post-processing simulation rather than true mobile capture.

Multi-Factor Compound Challenges: No publicly available dataset captures compound generalization challenges where temporal drift, environmental variations, and mobility occur simultaneously under controlled experimental conditions. This gap prevents evaluation of holistic adaptation frameworks addressing multi-factor variations.

Cross-Location with Device Identity: Datasets typically collect signals at a single physical location. Evaluation of spatial generalization requires identical devices captured at multiple geographic locations under varying propagation conditions, currently unavailable in public datasets.

These gaps represent critical research infrastructure needs that require community-wide data collection efforts to enable generalization solution development and validation.

B. Foundation Models and Large-Scale Pre-Training

Section IV.C established that current foundation model applications in RF fingerprinting focus primarily on generative augmentation, with diffusion models achieving high-quality synthesis (SSIM >0.9) but remaining protocol-specific and single-purpose [80]. These models learn to generate signals rather than to extract transferable device representations, limiting their utility for generalization beyond the training protocol. Future foundation models for RFF would need to learn protocol-agnostic hardware characteristic representations from large unlabeled signal corpora, analogous to how large language models learn transferable linguistic representations before task-specific fine-tuning.

Fine-Tuning Pathways for Large Signal Models: Two credible fine-tuning paradigms emerge from adjacent fields that warrant direct investigation for RFF.

LLM-Style Tokenization of RF Signals: I/Q sample sequences can be segmented into fixed-length signal tokens analogous to text tokens, enabling transformer-based architectures to process RF bursts autoregressively. A large signal model pre-trained on diverse unlabeled RF transmissions, predicting the next signal token from preceding context, would learn general representations of signal structure and hardware imperfections without requiring device labels. Fine-tuning for RFF would then require only a small labeled dataset per deployment scenario, analogous to few-shot fine-tuning of LLMs for downstream NLP tasks. The hypothesis is that the autoregressive pre-training objective forces the model to internalize hardware-induced distortion patterns as part of its signal prediction task, making these patterns available for device discrimination after fine-tuning with minimal labeled data.

Masked Signal Modeling for Protocol-Agnostic Features: Borrowing from BERT-style masked language modeling, a foundation model pre-trained by masking random time-frequency segments of RF signals and reconstructing them from surrounding context would be forced to learn the statistical regularities of hardware imperfections across protocols [75], [80], [88]. Since hardware effects manifest in every signal segment, the reconstruction task cannot be solved without implicitly modeling them. Fine-tuning on labeled device data would then specialize these general hardware representations for specific identification tasks. This approach is directly analogous to how BERT representations transfer across NLP tasks, and recent success of masked modeling for time-series signals suggests RF applicability [89], though RF's complex-valued structure

and hardware-specific distortion patterns introduce domain-specific challenges requiring systematic investigation.

Table 11 summarizes these fine-tuning pathways alongside current generative approaches, providing a structured comparison of pre-training objectives, data requirements, and anticipated RFF generalization benefits.

Data Requirements versus Current Availability: Effective multi-protocol foundation models would require over 100,000 device samples spanning diverse protocols, manufacturers, and environmental conditions. As Table 10 demonstrates, current public datasets provide approximately 274 unique devices across protocols, a gap of 2–3 orders of magnitude from ideal foundation model data requirements. The DARPA RFMLS dataset's 10,000+ device scale demonstrates feasibility but limited public access prevents community-wide foundation model development. Bridging this gap requires coordinated community data collection efforts at a scale not yet undertaken in the RFF field.

Fundamental Open Questions: Several questions must be resolved before multi-protocol foundation models become practical for RFF deployment.

Cross-Protocol Transfer Effectiveness: Can pre-training on WiFi signals genuinely improve LoRa fingerprinting, or are protocol-specific characteristics too dominant for effective transfer? Preliminary evidence suggests limited transfer between fundamentally different modulation schemes (OFDM vs. FSK), but systematic investigation is needed to determine transfer boundaries.

Input Representation Challenges: Handling varying sample rates (WiFi 100 MSps vs. LoRa $<$1 MSps), different signal structures (OFDM vs. spread spectrum), and protocol-specific timing creates substantial architectural challenges for unified model input processing that current architectures have not addressed.

Minimum Data Requirements: What dataset scale enables effective foundation model pre-training for RFF? Can data-efficient pre-training techniques from NLP translate to RF fingerprinting given its fundamentally different data characteristics and complex-valued signal structure?

These questions represent critical research directions that will determine whether foundation models become practical tools for RF fingerprinting generalization or remain computationally infeasible given current data and resource constraints.

C. Critical Research Gaps and Priorities

Having examined available datasets and future technology directions, we now synthesize the critical research gaps that prevent practical RF fingerprinting deployment, drawing from analysis across Sections III and IV to identify fundamental versus solvable limitations.

The analysis presented in Sections III and IV reveals limitations that cannot be addressed through incremental improvements to existing approaches. These gaps span multiple technical dimensions and represent barriers to achieving the practical deployment capabilities required for real-world security applications in heterogeneous wireless networks.

Multi-Factor Generalization [Critical Priority]: No current approach handles simultaneous compound variations, the interaction patterns documented in Section III.F produce degradation that substantially exceeds individual factor impacts. Addressing this requires physics-informed machine learning frameworks incorporating cyclostationary signal analysis to extract transmission-invariant hardware features [73], causal inference mechanisms that decompose compound environmental changes into manageable components, and hierarchical adaptation strategies that address temporal, spatial, and mobility factors at their respective time scales rather than treating them as a unified distribution shift problem.

Real-World Deployment Validation [Critical Priority]: Section IV.D identified the persistent gap between controlled laboratory results and operational performance, yet no standardized framework exists for real-world RFF validation. Progress requires long-term field studies documenting performance across seasonal and annual cycles, standardized benchmark protocols enabling reproducible operational environment assessment, and collaborative multi-institution deployment studies that capture the environmental diversity absent from current laboratory datasets.

Scalability and Open-Set Detection [High Priority]: Section III.E documented accuracy declining from over 90% for small device sets to 70–80% for populations exceeding 100 devices, with open-set detection compounding this challenge in dynamic environments. Required advances include autoencoder architectures with explicit unknown device modeling building on the reconstruction-error mechanism established in Section IV.B, efficient hierarchical classification strategies mitigating the substantial accuracy degradation observed in large-scale deployments, and robust confidence estimation enabling principled open-set rejection without per-deployment threshold tuning.

Adaptation and Evolution [High Priority]: Static learning approaches cannot accommodate hardware aging, interference evolution, or new device enrollment without disruptive retraining, a fundamental operational limitation identified in Sections III.C and IV.C. Research priorities include continual learning frameworks with catastrophic forgetting mitigation that scale beyond thousands of devices [84], [85], [90], aging-aware models that explicitly incorporate hardware drift trajectory patterns, and online adaptation mechanisms with computational overhead compatible with edge and IoT deployment constraints.

Integration and Computational Efficiency [Medium Priority]: Section IV.D trade-off analysis demonstrated that the most generalization-capable approaches (transformers, hybrid architectures) require resources incompatible with edge deployment, while edge-feasible approaches sacrifice generalization performance. Resolving this requires multi-objective optimization frameworks balancing accuracy and

TABLE 11. Foundation Model Paradigms for RF Fingerprinting: Current State and Future Directions

Paradigm	Pre-Training Objective	Data Requirement	RFF Fine-Tuning Benefit	Status
Diffusion Models [80]	Signal denoising	10K+ samples	Data augmentation only	Current
Autoregressive Signal Model	Next-token prediction	100K+ samples	Few-shot device adaptation	Hypothesized
Masked Signal Model	Segment reconstruction	50K–100K samples	Protocol-agnostic hardware features	Hypothesized
Contrastive Pre-Training	Device pair discrimination	50K+ samples (unlabeled)	Cross-protocol transfer	Hypothesized

computational cost, hardware-software co-design enabling sophisticated generalization techniques on constrained platforms, and integrated testing frameworks that validate compound solution effectiveness rather than optimizing individual dimensions in isolation.

VI. Conclusions

Securing communications in modern UAV-IoT ecosystems demands a paradigmatic shift from conventional protocol-dependent security frameworks to versatile, resource-conscious, and environment-resilient alternatives. The integration of these two transformative technologies introduces operational complexities characterized by dynamic topologies, intermittent connectivity, protocol heterogeneity, and extreme resource asymmetry. As demonstrated throughout this work, traditional security primitives, reliant on static assumptions and cryptographic infrastructure, struggle to meet the demands posed by these multifactorial and decentralized systems.

This survey positions RF fingerprinting as a critical physical-layer security solution capable of transcending protocol boundaries through the exploitation of hardware-induced signal imperfections. Its ability to deliver lightweight, scalable authentication based on non-replicable RF signatures makes it especially suited to dynamic deployments involving mobile aerial platforms and constrained terrestrial sensors. The unique cross-protocol applicability, resilience to spoofing, and potential for on-device inference make RF fingerprinting a promising foundational layer for future communication architectures.

However, the transition from theoretical viability to practical deployment hinges on addressing the generalization gap: the vulnerability of current deep learning-based fingerprinting models to performance degradation in previously unseen transmission scenarios, environmental conditions, and device configurations. Existing efforts often prioritize accuracy in controlled settings while overlooking real-world adaptability and scalability. This paper bridges that divide by reframing generalization not as a peripheral metric, but as a strategic priority, detailing its relationship with model capacity, data diversity, architectural design, and training methodology

Furthermore, this work advances the field through a structured review of RF fingerprinting evolution, from hand-crafted features to CNN, LSTM, attention mechanism, transformer, and autoencoder architectures, while identifying persistent limitations and proposing concrete future research directions. By consolidating technical insights across wireless security domains, this survey lays the foundation for robust, generalizable RF fingerprinting systems aligned with emerging trends in 6G, Non-Terrestrial Networks, and AI-native infrastructures, where hardware-level authentication at scale will become a core security primitive rather than an optional enhancement. The path forward necessitates interdisciplinary collaboration integrating wireless engineering, machine learning, cybersecurity, and hardware systems design to ensure that RF-based authentication evolves into a deployable solution that enhances the resilience, interoperability, and security of connected environments.

ACKNOWLEDGMENT

This work was supported by Thales Digital Identity and Security and the Mitacs Accelerate Fellowship program.

REFERENCES

[1] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019.

[2] N. Neshenko, E. Bou-Harb, J. Crichigno, G. Kaddoum, and N. Ghani, "Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2702–2733, 2019.

[3] I. Stellios, P. Kotzanikolaou, M. Psarakis, C. Alcaraz, and J. Lopez, "A survey of IoT-enabled cyberattacks: assessing attack paths to critical infrastructures and services," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3453–3495, 2018.

[4] M. Adil, H. Song, S. Mastorakis, H. Abulkasim, A. Farouk, and Z. Jin, "UAV-assisted IoT applications, cybersecurity threats, AI-enabled solutions, open challenges with future research directions," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 4, pp. 4583–4605, 2024.

[5] 3rd Generation Partnership Project (3GPP), "SA Rel-20: advancing 5G-Advanced and preparing for 6G." https://www.3gpp.org/news-events/3gpp-news/sa-rel20, 2025. Accessed: July 24, 2025.

[6] V.-L. Nguyen, P.-C. Lin, B.-C. Cheng, R.-H. Hwang, and Y.-D. Lin, "Security and privacy for 6G: A survey on prospective technologies and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 4, pp. 2384–2428, 2021.

[7] A. Kumar, R. Gupta, S. Tanwar, and N. Kumar, "A taxonomy of blockchain-enabled softwarization for secure UAV network," *Computer Communications*, vol. 161, pp. 304–323, 2020.

[8] N. H. Motlagh, T. Taleb, and O. Arouk, "Low-altitude unmanned aerial vehicles-based Internet of Things services: comprehensive survey and future perspectives," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 899–922, 2016.

[9] Y. Zou, J. Zhu, X. Wang, and V. C. M. Leung, "Physical layer security in wireless communications: a tutorial," *IEEE Wireless Communications*, vol. 23, no. 4, pp. 16–26, 2016.

[10] S. Li, Y. Zhu, J. Jiang, and J. Zhang, "A review of physical layer security techniques for Internet of Things: challenges and solutions," *IEEE Access*, vol. 8, pp. 20341–20356, 2020.

[11] Keyfactor and Vanson Bourne, "Digital trust in a connected world: navigating the state of IoT security." https://keyfactor.com/state-of-iot-security-report-2023/, 2023.

[12] R. Kong and H. Chen, "DeepCRF: deep learning-enhanced CSI-based RF fingerprinting for channel-resilient WiFi device identification," *IEEE Transactions on Information Forensics and Security*, 2025.

[13] S. Kuzdeba, J. Carmack, and J. Robinson, "RF fingerprinting with dilated causal convolutions—an inherently explainable architecture," in *Proc. 55th Asilomar Conference on Signals, Systems, and Computers*, (Pacific Grove, CA, USA), pp. 1–8, 2021.

[14] J. Xiao, H. Zhang, Z. Shao, Y. Zheng, and W. Ding, "Progressive unsupervised domain adaptation for radio frequency signal attribute recognition across communication scenarios," *Remote Sensing*, vol. 16, no. 19, p. 3696, 2024.

[15] B. Johnson and B. Hamdaoui, "On the domain generalizability of RF fingerprints through multifractal dimension representation," *Sensors*, vol. 22, no. 11, p. 4045, 2022.

[16] N. Quadar, A. Chehri, and B. Debaque, "Advanced security frameworks for UAV and IoT: a deep learning approach," *Internet of Things*, vol. 32, p. 101594, 2025.

[17] R. Altawy and A. M. Youssef, "Security, privacy, and safety aspects of civilian drones: a survey," *ACM Transactions on Cyber-Physical Systems*, vol. 1, no. 2, pp. 7:1–7:25, 2016.

[18] S. Challa, M. Wazid, A. K. Das, V. Odelu, N. Kumar, and P. Gope, "Lightweight public key infrastructure for the Internet of Things: A systematic literature review," *Internet of Things and Cyber-Physical Systems*, vol. 4, pp. 97–115, 2024.

[19] A. Zohourian, S. Dadkhah, E. C. P. Neto, H. Mahdikhani, P. K. Danso, H. Molyneaux, and A. A. Ghorbani, "IoT Zigbee device security: A comprehensive review," *Internet of Things*, vol. 22, p. 100791, 2023.

[20] G. Baldini and G. Steri, "A survey of techniques for the identification of mobile phones using the physical fingerprints of the built-in components," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1761–1789, 2017.

[21] A. Jagannath, Z. Kane, and J. Jagannath, "RF fingerprinting needs attention: multi-task approach for real-world WiFi and Bluetooth," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, (Rio de Janeiro, Brazil), 2022.

[22] S. Abbas, M. A. Talib, Q. Nasir, S. Idhis, M. Alaboudi, and A. Mohamed, "Radio frequency fingerprinting techniques for device identification: A survey," *International Journal of Information Security*, 2024.

[23] A. Ahmed, B. Quoitin, A. Gros, and V. Moeyaert, "A comprehensive survey on deep learning-based LoRa radio frequency fingerprinting identification," *Sensors*, vol. 24, no. 13, p. 4411, 2024.

[24] L. Xie, L. Peng, J. Zhang, and A. Hu, "Radio frequency fingerprint identification for Internet of Things: a survey," *Security and Safety*, 2024.

[25] G. Yan, X. Fu, Y. Wang, Q. Zhang, and G. Gui, "Radio frequency fingerprint identification towards statistical and deep learning features," *Peer-to-Peer Networking and Applications*, vol. 18, no. 116, 2025.

[26] S. Al-Hazbi, A. Hussain, S. Sciancalepore, G. Oligeri, and P. Papadimitratos, "Radio frequency fingerprinting via deep learning," in *20th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 0824–0829, 2024.

[27] L. Morge-Rollet, F. Le Roy, D. Le Jeune, C. Canaff, and R. Gautier, "RF eigenfingerprints: an efficient RF fingerprinting method in IoT context," *Sensors*, vol. 22, no. 11, p. 4291, 2022.

[28] S. Deng, Z. Huang, X. Wang, and G. Huang, "Radio frequency fingerprint extraction based on multidimension permutation entropy," *International Journal of Antennas and Propagation*, vol. 2017, pp. 1–6, 2017.

[29] K. Sankhe, M. Belgiovine, F. Zhou, L. Angioloni, F. Restuccia, S. D'Oro, S. Ioannidis, and K. Chowdhury, "ORACLE: optimized radio classification through convolutional neural networks," in *Proc. IEEE INFOCOM 2019 – IEEE Conference on Computer Communications*, (Paris, France), pp. 370–378, 2019.

[30] A. M. Ali, E. Uzundurukan, and A. Kara, "Assessment of features and classifiers for Bluetooth RF fingerprinting," *IEEE Access*, vol. 7, pp. 50524–50535, 2019.

[31] H. J. Patel, "Non-parametric feature generation for RF-fingerprinting on ZigBee devices," in *Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA)*, (Verona, NY, USA), pp. 1–5, 2015.

[32] W. Feng, Y. Li, C. Wu, and J. Zhang, "RF fingerprint extraction and device recognition algorithm based on multi-scale fractal features and APWOA-LSSVM," *EURASIP Journal on Advances in Signal Processing*, vol. 2023, no. 1, pp. 1–27, 2023.

[33] J. Yu, A. Hu, G. Li, and L. Peng, "A robust RF fingerprinting approach using multisampling convolutional neural network," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6786–6799, 2019.

[34] L. Xie, L. Peng, and J. Zhang, "Towards robust RF fingerprint identification using spectral regrowth and carrier frequency offset." arXiv preprint arXiv:2412.07269, 2024.

[35] E. Helluy-Lafont, A. Boé, G. Grimaud, and M. Hauspie, "Bluetooth devices fingerprinting using low cost SDR," in *Proc. Sixth International Workshop on Internet of Things: Networking Applications and Technologies*, 2020.

[36] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, "Radio frequency fingerprint identification for LoRa using deep learning," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2604–2616, 2021.

[37] 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 *Proc. 22nd International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc)*, pp. 251–260, 2021.

[38] M. Wang, L. Peng, L. Xie, J. Zhang, M. Liu, and H. Fu, "Design of noise robust open-set radio frequency fingerprint identification method," in *Proc. IEEE INFOCOM 2024 Workshops (DeepWireless)*, 2024.

[39] N. Quadar, A. Chehri, and B. Debaque, "Robust RF fingerprinting for LoRa IoT devices in mobile scenarios using CNN-LSTM-Attention," in *Proc. 101st IEEE Vehicular Technology Conference (VTC2025-Spring)*, pp. 1–5, 2025.

[40] I. Nemer, T. Sheltami, I. Ahmad, and A. U.-H. Yasar, "RF-based UAV detection and identification using hierarchical learning approach," *Sensors*, vol. 21, no. 6, p. 1947, 2021.

[41] S. Taşcıoğlu, A. Kalaycioğlu, M. Köse, and G. Soysal, "RF fingerprinting using transient-based identification signals at sampling rates close to the Nyquist limit," *Electronics*, vol. 14, no. 1, p. 4, 2025.

[42] N. Yuan, J. Zhang, Y. Ding, and S. L. Cotton, "Robust radio frequency fingerprint identification for Bluetooth Low Energy under low SNR and channel variations," in *Proc. 2025 IEEE Wireless Communications and Networking Conference (WCNC)*, 2025.

[43] G. Qing, H. Wang, and T. Zhang, "Radio frequency fingerprinting identification for ZigBee via lightweight CNN," *Physical Communication*, vol. 44, p. 101250, 2021.

[44] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, "Detection and classification of UAVs using RF fingerprints in the presence of Wi-Fi and Bluetooth interference," *IEEE Access*, vol. 8, pp. 166752–166765, 2020.

[45] R. Meng, B. Xu, X. Xu, M. Sun, B. Wang, S. Han, S. Lv, and P. Zhang, "A survey of machine learning-based physical-layer authentication in wireless communications," *Journal of Network and Computer Applications*, vol. 235, p. 104085, 2025.

[46] Z. Lai, Z. Chang, M. Sha, Q. Zhang, N. Xie, C. Chen, and D. Niyato, "A comprehensive survey on physical layer authentication techniques: Categorization and analysis of model-driven and data-driven approaches," *ACM Computing Surveys*, vol. 57, no. 5, p. 117, 2025.

[47] M. Shakiba-Herfeh, A. Chorti, and H. V. Poor, "Physical layer security: Authentication, integrity, and confidentiality," in *Physical Layer Security*, pp. 129–150, 2021.

[48] S. Hanna *et al.*, "WiSig: large-scale WiFi signal dataset for receiver and channel agnostic RF fingerprinting," *IEEE Access*, vol. 10, pp. 22809–22831, 2022. Dataset: https://cores.ee.ucla.edu/downloads/datasets/wisig/.

[49] S. Hanna, S. Yan, and D. Cabric, "Open set wireless transmitter authorization: deep learning approaches and dataset considerations,"

IEEE Transactions on Cognitive Communications and Networking, vol. 7, no. 2, pp. 59–72, 2021.

[50] A. Jagannath, J. Jagannath, and P. S. P. V. Kumar, "A comprehensive survey on radio frequency (rf) fingerprinting: Traditional approaches, deep learning, and open challenges, journal = Computer Networks, volume = 219, pages = 109455, year = 2022,"

[51] T. Jian, B. C. Rendon, E. Ojuba, N. Y. Soltani, Z. Wang, K. Sankhe, A. Gritsenko, J. Dy, K. Chowdhury, and S. Ioannidis, "Deep learning for RF fingerprinting: a massive experimental study," *IEEE Internet of Things Magazine*, vol. 3, no. 1, pp. 50–57, 2020.

[52] J. Chen, H. Yin, and Y. Yang, "A cross-domain few-shot learning method based on domain knowledge mapping." arXiv preprint arXiv:2504.06608, 2025.

[53] M. O. Zeeshan, M. Pedersoli, A. L. Koerich, and E. Grange, "Progressive multi-source domain adaptation for personalized facial expression recognition." arXiv preprint arXiv:2504.04252, 2025.

[54] Y. Zhang, Z. Zhou, and X. Li, "Specific emitter identification handling modulation variation with margin disparity discrepancy," *arXiv preprint arXiv:2403.11531*, 2024.

[55] J. A. Gutierrez del Arroyo, B. J. Borghetti, and M. A. Temple, "Considerations for radio frequency fingerprinting across multiple frequency channels," *Sensors*, vol. 22, no. 6, p. 2111, 2022.

[56] Y. Feng *et al.*, "Radio frequency fingerprint recognition method based on prior information," *Digital Signal Processing*, 2024.

[57] 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*, pp. 439–450, 2023.

[58] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro, "Towards scalable and channel-robust radio frequency fingerprint identification for LoRa," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 774–787, 2022.

[59] D. Luan, S. Thompson, and J. Jiang, "Attention based neural networks for wireless channel estimation," *IEEE Communications Letters*, vol. 26, pp. 1424–1428, June 2022. Application of attention mechanisms to wireless signal processing and channel estimation.

[60] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, "Deep learning for rf device fingerprinting in cognitive communication systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, 2018.

[61] O. O. Medaiyese, A. P. Lauf, and I. Guvenc, "A low complexity feature extraction for the rf fingerprinting process," in *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pp. 1–6, IEEE, 2020.

[62] E. G. Li, X. Geng, J. Lee, Y. Jung, and D. Lee, "A review on deep learning for edge intelligence in 6g networks," *IEEE Access*, vol. 9, pp. 79366–79396, 2021.

[63] M. Raissi, P. Perdikaris, N. Ahmadi, and G. E. Karniadakis, "Physics-informed neural networks and extensions," *arXiv preprint arXiv:2408.16806*, 2024.

[64] J. Dusenka, "A robust rf fingerprinting approach using physics-informed neural networks," 2024.

[65] X. Chen and X. Zhang, "RF genesis: Zero-shot generalization of mmWave sensing through simulation-based data synthesis and generative diffusion models," in *Proceedings of the 21st ACM Conference on Embedded Networked Sensor Systems*, SenSys '23, (Istanbul, Turkiye), pp. 1–14, ACM, 2023.

[66] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury, "Deep learning convolutional neural networks for radio identification," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 146–152, 2018.

[67] H. Gu, L. Su, W. Zhang, and C. Ran, "Attention is needed for RF fingerprinting," *IEEE Access*, vol. 11, pp. 87316–87329, 2023.

[68] H. Kulhandjian, M. Kulhandjian, C. D'Amours, C. Ellement, and E. Kermorvant, "AI-based RF-Fingerprinting Framework and Experimental Results Using ResNet and GoogLeNet," in *2023 International Conference on Computing, Networking and Communications (ICNC)*, pp. 407–411, IEEE, 2023.

[69] X. Zhang, J. Zheng, W. Ding, S. Wang, W. Wu, and Z. Wang, "Fine-grained radio frequency fingerprint recognition network based on attention mechanism," *Electronics*, vol. 13, no. 3, p. 590, 2024.

[70] L. Wu, Y. Zhao, M. Feng, and A. Abdalla, "Deep learning based RF fingerprinting for device identification and wireless security," *Electronics Letters*, vol. 54, no. 24, pp. 1405–1407, 2018.

[71] D. Roy, T. Mukherjee, M. Chatterjee, and E. Pasiliao, "Rf transmitter fingerprinting exploiting spatio-temporal properties in raw signal data," in *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pp. 89–96, IEEE, 2019.

[72] Y. Hu, Y. Fu, and Y. Chen

[73] N. Quadar, A. Chehri, and B. Debaque, "Scalable deep learning for rf fingerprinting: The made architecture for robust physical-layer device identification," *IEEE Open Journal of the Communications Society*, vol. 7, pp. 1973–1993, 2026.

[74] J. Yu, A. Hu, F. Zhou, Y. Xing, Y. Yu, G. Li, and L. Peng, "Radio frequency fingerprint identification based on denoising autoencoders," in *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pp. 1–6, IEEE, 2019.

[75] G. Parpart, J. H. Tu, B. Clymer, J. Lee, and J. Babcock, "Transformer masked autoencoders for rf device fingerprinting," in *MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM)*, pp. 859–862, IEEE, 2024.

[76] Z. Wang, J. Li, W. Wang, Z. Dong, Q. Zhang, and Y. Guo, "Review of few-shot learning application in csi human sensing," *Artificial Intelligence Review*, vol. 57, no. 2, pp. 1–45, 2024.

[77] H. Li, C. Wang, N. Ghose, and B. Wang, "POSTER: robust deep-learning-based radio fingerprinting with fine-tuning," in *Proc. ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec)*, 2021. Dataset: https://github.com/SmartHomePrivacyProject/RadioFingerprinting.

[78] N. Houlsby, A. Giampouris, A. Ratner, S. Shen, A. Taylor, D. Eck, K. Cho, and O. Firat, "Parameter-efficient transfer learning for nlp," pp. 2790–2799, 2019.

[79] J. Wen, J. Kang, D. Niyato, Y. Zhang, J. Wang, B. Sikdar, and P. Zhang, "Generative AI for data augmentation in wireless networks: analysis, applications, and case study." arXiv preprint arXiv:2411.08341, 2024.

[80] G. Chi, Z. Yang, C. Wu, J. Xu, Y. Gao, Y. Liu, and T. X. Han, "RF-Diffusion: radio signal generation via time-frequency diffusion," in *Proc. 30th Annual International Conference on Mobile Computing and Networking (ACM MobiCom)*, (Washington, D.C., USA), pp. 77–92, 2024.

[81] J. Sun, M. Alazab, W. Xu, H. Chen, L. Khan, and M. S. I. Pathan, "Edge AI-based radio frequency fingerprinting for IoT networks," *arXiv preprint arXiv:2412.10553*, 2024.

[82] S. Al-Hazbi, A. Hussain, S. Sciancalepore, G. Oligeri, and P. Papadimitratos, "Radio frequency fingerprinting via deep learning: Challenges and opportunities." arXiv preprint arXiv:2310.16406, 2023.

[83] F. Restuccia, S. D'Oro, A. Al-Shawabka, and T. Melodia, "Securing the internet of things in the age of machine learning and software-defined networking," *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4829–4842, 2018.

[84] H. Shin, J. K. Lee, J. Kim, and J. Kim, "Continual learning with deep generative replay," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[85] B. Zhao, M. Alfarra, and P. H. S. Torr, "Memory efficient continual learning with transformers," vol. 35, pp. 29996–30009, 2022.

[86] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, "Drone remote controller RF signal dataset," 2020. 124 GB; 17 controllers; ∼1000 signals per controller.

[87] S. Basak, S. Rajendran, S. Pollin, and B. Scheers, "Combined RF-based drone detection and classification," *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 1, pp. 111–120, 2022.

[88] Z. Zhou and X. Liu, "Masked autoencoders in computer vision: A comprehensive survey," *IEEE Access*, vol. 11, pp. 113560–113579, 2023.

[89] W. A. Gardner, "The spectral correlation theory of cyclostationary time-series," *Signal Processing*, vol. 11, no. 1, pp. 13–36, 1986.

[90] J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness, G. Desjardins, A. A. Rusu, K. Milan, J. Quan, T. Ramalho, A. Grabska-Barwinska, D. Hassabis, C. Clopath, D. Kumaran, and R. Hadsell, "Overcoming catastrophic forgetting in neural networks," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.

Nordine Quadar(Member, IEEE) completed his B.Sc. and M.Sc. at the University of Ottawa. He is currently a Ph.D. candidate in Computer Science at the Royal Military College of Canada. His research focuses on deep learning, RF fingerprinting, physical-layer security, Internet of Things, embedded AI, Cybersecurity and Wireless Communication.

Abdellah Chehri(Senior Member, IEEE) is a Full Professor in the Department of Mathematics and Computer Science at the Royal Military College of Canada (RMC) in Kingston, Ontario. He also holds affiliate professorships at the Université du Québec en Outaouais (UQO) and the Université du Québec à Chicoutimi (UQAC), as well as an adjunct professorship at the University of Ottawa. His research spans a wide spectrum of advanced technologies, including intelligent mobility systems, sustainable transportation infrastructure, and vehicular communications. Dr. Chehri is internationally recognized for his contributions to V2X communications, 5G/6G and next-generation wireless networks, and AI-driven edge computing. His work integrates the Internet of Things (IoT), RF sensing, and signal processing to enable real-time decision-making in autonomous systems and smart environments. Dr. Chehri has served as a guest and associate editor for numerous highly regarded academic journals. He is actively involved in several professional societies, including the IEEE Public Safety Technology (PST) Initiative, where he serves as co-chair of the Transportation Committee, as well as the IEEE Communications Society (ComSoc), IEEE Vehicular Technology Society (VTS), and IEEE Photonics Society.

Benoit Debaque(Member, IEEE) Benoit Debaque has 25 years of work experience as a Researcher of computer vision and artificial intelligence. He has been an Artificial Intelligence Specialist with Thales Digital Solutions Canada, since 2019. His research interest includes sensor fusion to embedded AI in highly constrained environments.

Halim Yanikomeroglu(Fellow, IEEE) is a Chancellor's Professor in the Department of Systems and Computer Engineering at Carleton University, Canada. A globally recognized thought leader in wireless communications, he is a Fellow of the IEEE, the Engineering Institute of Canada (EIC), and the Canadian Academy of Engineering (CAE). His research has profoundly shaped the evolution of cellular networks, including 5G and emerging 6G technologies, with a focus on aerial networks, non-terrestrial architectures, and advanced radio access systems. With over 700 peer-reviewed publications and 39 granted patents, Dr. Yanikomeroglu's work bridges academia and industry, supported by collaborations with major tech firms such as Huawei, Samsung, and Telus. He has delivered numerous keynotes and tutorials worldwide and played pivotal roles in IEEE governance, including chairing the WCNC Steering Committee and serving as a Distinguished Speaker for both the IEEE Communications Society and Vehicular Technology Society. His accolades include the IEEE Wireless Communications Technical Committee Recognition Award and the IEEE Vehicular Technology Society Stuart Meyer Memorial Award, among others.

Gunes Karabulut Kurt(Senior Member, IEEE) is a Canada Research Chair (Tier 1) in New Frontiers in Space Communications and a Full Professor in the Department of Electrical Engineering at Polytechnique Montréal. She also serves as an Adjunct Research Professor at Carleton University. Her research spans space security, satellite networks, and wireless testbeds, with a strong emphasis on multi-functional space systems and autonomous spectrum management. Dr. Kurt is the Director of the Poly-Grames Research Center and co-founder of ASTROLITH, a transdisciplinary initiative focused on space resource and infrastructure engineering. A Marie Curie Fellow and recipient of the TÜBA-GEBIP Outstanding Young Scientist Award, she has held leadership roles across IEEE, including chairing the Special Interest Group on Satellite Mega-constellations and serving on the WCNC Steering Board. Her editorial contributions span six IEEE journals, and her work continues to shape the future of secure, resilient space communications.

2.3 Chapter Summary

The comprehensive analysis presented in the preceding survey establishes the current state of RF fingerprinting technology while identifying the critical research gaps that motivate the technical investigations of this thesis. We note here that while the literature demonstrates impressive accuracy metrics, usually ranging between 95% and 99%, these results are almost exclusively obtained under tightly controlled conditions. Consequently, fundamental generalization challenges remain the primary barrier to practical deployment.

Our literature analysis reveals consistent and troubling performance degradation patterns across the field: cross-transmission accuracy drops of 20-70%, temporal degradation between 20-60%, and mobility-induced reductions of 15-40%. We argue that these failures across multiple independent studies indicate fundamental architectural limitations rather than implementation issues. This insight, therefore, establishes a clear need for the new approaches put forth in the following chapters. To this end, we define the core challenge as the development of RF fingerprinting frameworks that can maintain robust performance across diverse operational axes while simultaneously meeting the scalability and efficiency requirements inherent to UAV and IoT deployments. This multifaceted challenge, including cross-transmission generalization, temporal adaptation, mobility handling, and computational efficiency, must be addressed in a unified manner to achieve practical viability.

The research gaps identified through this chapter directly inform the four technical contributions of this work as follow: Statistical feature enhancement (Chapter 4) addresses the necessity for temporal robustness. Temporal modeling (Chapter 5) provides the mechanisms to handle mobility scenarios. Cyclostationary features (Chapter 6) tackles the complexities of cross-transmission challenges. And the MADE (Chapter 7) offers a scalable solution and foundation for unknown device detection. These subsequent chapters address these challenges using complementary approaches that collectively advance RF fingerprinting toward practical deployment readiness.

Chapter 3

3 Datasets and Experimental Framework

3.1 Introduction

Evaluating RF fingerprinting under realistic conditions requires datasets that capture the operational diversity of UAV and IoT deployments. As we established in Chapter 2, current RF fingerprinting approaches face significant performance degradation when confronted with cross-transmission parameter variations, temporal drift, and mobile scenario effects. These generalization challenges produce accuracy drops ranging from 20–70% in cross-transmission scenarios and 20–60% under temporal shift, while mobility-induced conditions cause 15–40% degradation, which limits the practical viability of existing fingerprinting systems [42].

RF fingerprinting research has traditionally been constrained by limited dataset availability and protocol-specific evaluations that hinder meaningful comparison between approaches [19, 18]. The absence of standardized evaluation frameworks resulted in fragmented research progress where technical contributions cannot be objectively compared [43]. We address these limitations in this chapter by establishing an experimental framework built upon six carefully selected datasets that collectively span the operational diversity of contemporary UAV and IoT systems.

Chapter Purpose and Scope

This chapter serves as the methodological foundation for the technical contributions we present in Chapters 4 through 7. We provide three essential components:

Dataset Portfolio: We present six datasets spanning WiFi (IEEE 802.11), UAV remote control links, LoRa IoT communications, and Bluetooth emis-

sions. We selected each dataset to address specific generalization challenges while maintaining experimental rigor and reproducibility.

Unified Preprocessing Framework: We define standardized signal processing pipelines and quality assurance protocols that ensure fair comparison across different technical approaches while preserving protocol-specific characteristics essential for fingerprinting.

Evaluation Methodology: We establish consistent experimental design principles and performance metrics that allow rigorous assessment of our technical contributions under standardized conditions.

Relationship to Technical Chapters

We strategically integrated the datasets and methodologies presented here with the technical contributions in subsequent chapters. Each technical chapter contains detailed descriptions of dataset usage, experimental protocols, and application-specific preprocessing relevant to that particular contribution. This chapter provides the *complementary overview* that establishes the unified framework connecting these individual evaluations.

The integration follows a structured mapping: Chapter 4 examines statistical enhancement methods using WiFi and Bluetooth datasets with cross-day collection structures; Chapter 5 assesses temporal modeling and mobile scenario adaptation using the LoRa IoT dataset's extensive temporal coverage; Chapter 6 evaluates progressive learning approaches using UAV controller datasets with controlled transmission parameter variations; and Chapter 7 validates unknown device detection capabilities using WiSIG dataset's scale and receiver diversity.

Chapter Organization

Section 3.2 presents the dataset selection criteria and quality assessment framework that guided our portfolio construction. Section 3.3 presents the six-dataset portfolio with emphasis on unique characteristics and challenge-specific capabilities. Section 3.4 describes the standardized experimental methodology, and Section 3.6 summarizes how the integrated framework supports the technical evaluations in subsequent chapters.

3.2 Dataset Selection and Quality Framework

We designed the dataset selection process to address the generalization challenges that limit practical deployment of RF fingerprinting systems. As we

discussed in Chapter 2, current approaches suffer from performance degradation when confronted with variations in transmission parameters, temporal conditions, and mobility scenarios. To evaluate solutions to these challenges, we selected datasets based on criteria that ensure experimental validity, operational realism, and reproducibility.

3.2.1 Selection Criteria

We constructed the dataset portfolio following four primary selection criteria that collectively ensure evaluation coverage across the identified generalization challenges:

Transmission Parameter Diversity: Selected datasets must contain signals captured under varying transmission configurations to enable cross-transmission generalization assessment. This includes variations in carrier frequency, transmission power, modulation schemes, and protocol-specific parameters that reflect real-world operational diversity. We excluded datasets lacking transmission parameter variations as they cannot adequately evaluate cross-transmission generalization capabilities.

Temporal Coverage: Datasets must include signal collections across multiple time periods to evaluate temporal stability and aging effects. This encompasses same-day variations, cross-day stability, and long-term degradation patterns that affect fingerprinting accuracy in deployed systems. Temporal coverage is essential because fingerprinting approaches must maintain performance as device hardware ages and environmental conditions change over time.

Environmental Realism: We required datasets that represent realistic deployment environments rather than idealized laboratory conditions. This includes indoor/outdoor variations, multipath effects, and interference scenarios typical of UAV and IoT operational environments. While controlled laboratory collections provide valuable baseline assessments, practical deployment requires evaluation under realistic channel conditions that cannot be replicated in controlled settings.

Protocol Diversity: To address the heterogeneous nature of UAV/IoT ecosystems, we ensured that datasets span multiple wireless protocols including WiFi (IEEE 802.11), UAV control links, LoRa communications, and Bluetooth emissions. Single-protocol datasets, while valuable for protocol-specific optimization, cannot assess how well fingerprinting approaches generalize across the diverse wireless technologies found in integrated deployments.

3.2.2 Quality Assessment Framework

Each selected dataset undergoes quality assessment to ensure experimental validity and reproducibility. We evaluate datasets across multiple dimensions that directly impact the reliability of experimental results:

Signal Quality Metrics: We assess signal-to-noise ratio characteristics, dynamic range, and sampling adequacy to ensure sufficient signal fidelity for fingerprinting analysis. Datasets should demonstrate adequate SNR levels that enable extraction of hardware-specific characteristics while still representing realistic operational conditions. We exclude datasets with inadequate SNR or sampling artifacts that could bias fingerprinting results. The assessment also includes verification of sample rate sufficiency relative to signal bandwidth, ensuring that hardware-induced transients are adequately captured.

Device Diversity: We evaluate the number and types of devices included in each dataset to ensure sufficient diversity for robust fingerprinting evaluation. Datasets must include multiple device manufacturers, chipset variations, and hardware generations to avoid overfitting to specific device characteristics. Single-manufacturer or single-chipset datasets may achieve artificially high performance that does not generalize to the heterogeneous device populations encountered in real deployments. Our assessment considers both inter-manufacturer diversity (different brands) and intra-manufacturer diversity (different models from the same brand).

Collection Methodology: We assess the experimental setup, collection procedures, and environmental controls for reproducibility and to minimize confounding factors. Key factors include receiver calibration procedures, transmitter placement consistency, environmental condition monitoring, and temporal organization of collection sessions. Datasets must provide sufficient metadata for independent verification and extension of experimental results, and we excluded those with inadequate documentation.

Data Format Standardization: We evaluate compatibility with standardized signal processing pipelines and the availability of proper metadata to ensure seamless integration into our experimental framework. Datasets must provide clear documentation of signal format, sampling parameters, device labels, and collection conditions. We prioritize datasets with open formats that facilitate automated processing and quality validation rather than proprietary formats that hinder reproducibility.

3.3 Dataset Portfolio

Our selected portfolio comprises six complementary datasets that collectively address the generalization challenges identified in Chapter 2. Each dataset provides specific evaluation capabilities while contributing to the broader framework for assessing RF fingerprinting performance across diverse operational conditions. We present the portfolio overview first, followed by detailed characterization of each dataset.

3.3.1 Portfolio Overview and Integration

Tables 3.1 and 3.2 provide an overview of the dataset portfolio, including source citations, scale characteristics, protocol coverage, and integration with technical chapters. The portfolio spans device populations from 5 to 174 transmitters, encompasses four distinct wireless protocols, and provides temporal coverage ranging from same-day collections to multi-month longitudinal studies.

Table 3.1: Dataset characteristics and sources

Dataset	**Scale**	**Protocol**	**Source**	**Primary Challenge**
GLOBECOM22	10 devices	WiFi + Bluetooth	Jagannath et al. [11]	Multi-Protocol
ACMWiSec21	5 devices	WiFi (802.11)	Li et al. [44]	Cross-Day
UAV-8	8 devices	UAV Control	Basak et al. [45]	Cross-Transmission
UAV-17	17 devices	UAV Control	Ezuma et al. [46]	Cross-Transmission
LoRa-60	60 devices	LoRa IoT	Shen et al. [47]	Temporal/Mobile
WiSIG	174 Tx, 41 Rx	WiFi (802.11)	Hanna et al. [48]	Unknown Device Detection

The portfolio provides complementary evaluation capabilities across multiple dimensions. WiSIG's scale enables assessment of unknown device detection and cross-receiver generalization with 174 transmitters and 41 receivers. The UAV controller datasets provide controlled evaluation of cross-transmission generalization through systematic parameter variations across both 8-device

Table 3.2: Dataset technical integration and evaluation capabilities

Dataset	**Technical Integration**	**Evaluation Capability**
GLOBECOM22	Chapter 4	Attention mechanisms, cross-day generalization, multi-protocol
ACMWiSec21	Chapter 4	Statistical features, temporal adaptation, fine-tuning
LoRa-60	Chapter 5	Aging effects, mobility adaptation, channel robustness
UAV-8	Chapter 6	Progressive learning, controlled parameter variations
UAV-17	Chapter 6	Scalability assessment, continual adaptation
WiSIG	Chapter 7	Large-scale open-set recognition, cross-receiver generalization

and 17-device populations. LoRa IoT dataset's temporal coverage supports evaluation of aging effects and mobile scenario adaptation across same-day, cross-day, and multi-week collection periods. WiFi and Bluetooth datasets enable assessment of statistical enhancement methods and multi-protocol fingerprinting under cross-day temporal variations.

We present detailed characterization of each dataset in the following subsections, emphasizing the unique characteristics and specific contributions to our evaluation framework.

3.3.2 WiFi and Bluetooth Datasets

The proliferation of WiFi-Bluetooth combo chipsets in contemporary IoT devices presents unique challenges for RF fingerprinting, as approaches must distinguish devices across heterogeneous wireless protocols while maintaining robustness to temporal variations. We use two complementary datasets to address multi-protocol fingerprinting evaluation and cross-day generalization: the GLOBECOM22 dataset for real-world combo chipset evaluation [11] and the ACMWiSec21 dataset for controlled cross-day WiFi evaluation [44].

GLOBECOM22 Dataset: Multi-Protocol Combo Chipsets

The GLOBECOM22 dataset addresses the challenge of fingerprinting devices that share RF circuitry across multiple wireless protocols. It comprises 8 Raspberry Pi 4B devices equipped with Cypress CYW43455 combo chipsets and 2 Lenovo laptops with Intel AC 7260 chipsets. Signal collection employs a USRP X300 with UBX160 daughterboard and VERT2450 antenna, operating at 66.67 MHz bandwidth centered at 2.414 GHz.

Key Specifications:

- **Devices:** 8 Raspberry Pi 4B (Cypress CYW43455), 2 Lenovo laptops (Intel AC 7260)
- **Protocols:** WiFi (IEEE 802.11g, Channel 8) and Bluetooth (1600 hops/second)
- **Temporal Structure:** Day 1/Day 2 collections separated by months
- **Format:** 40 Mega Samples per capture, SigMF compliance with JSON metadata
- **Collection Conditions:** Multiple power levels, line-of-sight, distance variations (0.5–3.0m)

What makes this dataset particularly useful is its combo chipset focus, where WiFi and Bluetooth emissions share antenna and RF front-end circuitry. This creates subtle hardware variations that enable cross-protocol fingerprinting evaluation. The months-long gap between Day 1 and Day 2 collections also provides a realistic temporal degradation scenario while maintaining controlled experimental conditions.

ACMWiSec21 Dataset: Controlled Cross-Day Evaluation

The ACMWiSec21 dataset provides evaluation of cross-day generalization and fine-tuning recovery under controlled laboratory conditions. It employs 5 HackRF One transmitters with ANT500 antennas, implementing IEEE 802.11 a/g protocol with BPSK 1/2 modulation using GNU Radio.

Key Specifications:

- **Devices:** 5 HackRF One transmitters with ANT500 antennas
- **Protocol:** IEEE 802.11 a/g, BPSK 1/2, GNU Radio implementation
- **Parameters:** 2.45 GHz, 2 MHz bandwidth, 288 IQ samples per trace
- **Organization:** 2-day collection, 100,000 traces per device per day
- **Evaluation Structure:** Same-day baseline, cross-day challenge, fine-tuning assessment

The controlled two-day collection structure enables evaluation of temporal degradation effects and minimal-data adaptation strategies. The dataset provides baseline same-day performance for reference, cross-day generalization

challenges that reflect realistic temporal shifts, fine-tuning recovery assessment with limited adaptation data, and controlled conditions that isolate temporal effects from other confounding factors.

Evaluation Capabilities and Integration

Both datasets enable evaluation across multiple dimensions including protocol diversity (WiFi and Bluetooth in GLOBECOM22), temporal stability (same-day to multi-month variations), and multi-task learning scenarios. The GLOBECOM22 dataset's combo chipset design enables evaluation of shared RF circuitry fingerprinting, while ACMWiSec21 provides controlled cross-day evaluation with fine-tuning assessment.

In Chapter 4, we use these datasets to evaluate statistical enhancement methods and multi-modal attention architectures. Their multi-protocol structure supports evaluation of 14 statistical features including magnitude, phase, power, entropy, and spectral flatness across both temporal and protocol dimensions. The multi-protocol nature also enables assessment of multi-modal attention mechanisms that integrate spatial-temporal features, temporal pattern extraction, and time-frequency attention. We present the detailed experimental protocols, feature extraction procedures, and attention mechanism evaluations in Chapter 4.

3.3.3 UAV Controller Datasets

UAV communication datasets form the basis for evaluating cross-transmission generalization capabilities in our RF fingerprinting framework. We selected two complementary datasets to evaluate progressive learning approaches and scalability under controlled transmission parameter variations: an 8-device collection developed through collaboration between KU Leuven and the Royal Military Academy [45], and a 17-device extended dataset from MPACT Lab at NC State University [46].

Dataset Specifications and Infrastructure

The 8-device dataset provides controlled environment evaluation in a semi-anechoic chamber, which eliminates multipath effects and enables isolation of hardware-specific fingerprinting characteristics. The collection employs an Ettus USRP X310 receiver operating at 100 MSps with center frequency of 2.44 GHz, positioned 7 meters from transmitters. The dataset encompasses 9 UAV/controller categories including commercial DJI systems (Mini 2, Inspire

2, Matrice 300) and hobby-grade RC controllers (Spektrum DX4e, Taranis Q X7, Nine Eagles, WLtoys, Q205, SJRC F11 Pro).

The 17-device dataset extends the device diversity with high-resolution capture in an indoor laboratory environment. Signal collection employs a Keysight MSOS604A oscilloscope operating at 20 GSa/s with 6 GHz bandwidth and 24 dBi antenna. It contains 17 controllers from 8 manufacturers with approximately 1000 signals per controller, each signal containing 5 million samples, totaling 124 GB.

Table 3.3: Comparison of UAV controller dataset specifications

Specification	**8-Device** [45]	**17-Device** [46]
Dataset Characteristics		
Controllers	9 UAV/controller categories	17 controllers (15 unique)
Manufacturers	5 brands	8 brands
Dataset Size	66 GB	124 GB
Signals/Controller	∼1,000	∼1,000
Experimental Configuration		
Environment	Semi-anechoic chamber	Indoor laboratory (1–5 m)
Receiver	Ettus USRP X310	Keysight MSOS604A
Sampling Rate	100 MSps	20 GSa/s
Center Frequency	2.44 GHz	2.4 GHz
Distance	7 m (fixed)	1–5 m (variable)
Evaluation Capabilities		
Primary Use	Controlled environment	High-resolution diversity
Key Advantage	Eliminates multipath	Large-scale evaluation
Transmission Params	Tx1–Tx19 variations	Tx1–Tx19 variations

Cross-Transmission Parameter Framework

A critical feature of both datasets is the systematic variation of transmission parameters that allows controlled evaluation of cross-transmission generalization. Both datasets employ identical transmission parameter notation (Tx1 through Tx19) for unified evaluation, with variations in carrier frequency, transmission power levels, modulation parameters, and frequency hopping sequences for FHSS devices.

Parameter Variation Structure:

- **Training Configuration:** Tx1 and Tx2 for base model establishment

- **Evaluation Configuration:** Tx3–Tx19 for cross-transmission generalization assessment
- **Parameter Diversity:** Frequency variations within 2.4 GHz ISM band, multiple power settings, protocol-specific modulation variations

The datasets capture both remote control (RC) and video transmission RF signals, providing coverage of the main UAV communication links. Signal characteristics vary significantly across manufacturers due to different modulation schemes, frequency hopping patterns, and transmission power levels, which creates natural diversity that is essential for robust fingerprinting evaluation. The complementary nature of these two datasets enables evaluation across different scales and experimental conditions: controlled semi-anechoic chamber versus realistic indoor laboratory, and different instrumentation approaches that offer different trade-offs between flexibility and precision.

Evaluation Capabilities and Integration

We selected both UAV controller datasets specifically to evaluate the progressive learning approach we present in Chapter 6. The transmission parameter variations enable controlled assessment of continual adaptation capabilities when new transmission configurations are encountered. Because the parameter variations are controlled, we can isolate cross-transmission generalization effects from other confounding factors and provide clear assessment of progressive learning effectiveness.

The datasets support our standardized evaluation protocol where we initially train on Tx1–Tx2 configurations followed by progressive evaluation across Tx3–Tx19 variations. Evaluation across both datasets validates performance consistency at both the 8-device and 17-device scales. Chapter 6 presents the detailed experimental protocols including cyclostationary feature extraction procedures, progressive learning implementation, and performance analysis across all transmission parameter variations.

3.3.4 LoRa IoT Dataset

Long Range (LoRa) communication protocols represent an important component of modern IoT ecosystems, especially in applications requiring extended range and low power consumption. The LoRa IoT device dataset provides evaluation capabilities for temporal modeling and mobile scenario adaptation, addressing the challenge of maintaining fingerprinting accuracy across varying environmental conditions and device aging effects [47].

Device Portfolio and Collection Infrastructure

The LoRa dataset encompasses 60 commercial off-the-shelf LoRa devices representing four distinct device categories and three different LoRa chipsets. This ensures evaluation across the range of LoRa hardware commonly deployed in IoT applications. Signal collection employs a USRP N210 software-defined radio platform at 868.1 MHz (European LoRa frequency), focusing on LoRa preamble sequences for optimal fingerprinting effectiveness.

Device Distribution and Chipset Diversity:

- **Pycom LoPy4 (Devices 1–45):** SX1276 high-performance LoRa transceiver
- **mbed SX1261 shield (Devices 46–50):** SX1261 ultra-low power transceiver
- **Pycom FiPy (Devices 51–55):** SX1272 original LoRa chipset architecture
- **Dragino SX1276 shield (Devices 56–60):** SX1276 alternative manufacturer implementation

Technical Specifications:

- **Receiver:** USRP N210 software-defined radio
- **Frequency:** 868.1 MHz (European LoRa band)
- **Structure:** 26 sub-datasets (HDF5 format) with IQ samples and device labels
- **Data Organization:** Concatenated I-branch and Q-branch components with metadata
- **Signal Focus:** LoRa preamble sequences (consistent device-specific characteristics)

Temporal Coverage and Environmental Diversity

One of the main strengths of the LoRa dataset lies in its temporal coverage, which supports evaluation of aging effects and long-term stability patterns that are relevant for operational IoT deployments. The dataset includes collections spanning multiple time scales together with environmental variation analysis to evaluate fingerprinting robustness under diverse conditions.

Temporal Framework:

- **Short-Term Variations:** Same-day environmental changes and operational state variations
- **Medium-Term Patterns:** Cross-day stability and weekly cycling effects
- **Long-Term Aging:** Monthly collections evaluating device aging and hardware drift

Environmental Conditions:

- **Indoor/Outdoor Scenarios:** Controlled and realistic environmental conditions
- **Channel Variations:** Line-of-sight (LOS) and non-line-of-sight (NLOS) evaluation
- **Interference Assessment:** Dense IoT deployment simulation scenarios

The modular HDF5 structure supports flexible experimental design while maintaining clear organizational principles for temporal and environmental variation analysis. Each file contains metadata including collection timestamps, environmental conditions, and quality metrics.

Evaluation Capabilities and Integration

The LoRa dataset's environmental and temporal coverage makes it particularly suitable for evaluating the mobile scenario adaptation approaches we present in Chapter 5. It supports mobility evaluation through channel variation analysis, temporal pattern assessment, and Doppler effect simulation.

We can use this dataset for multiple evaluation scenarios: chipset-level classification among three chipset types (SX1272, SX1276, SX1261), manufacturer-level distinction between Pycom and Dragino devices, individual device identification among all 60 devices, and unknown device detection for devices not seen during training. The device organization supports controlled evaluation of hierarchical classification approaches at different granularity levels.

In Chapter 5, we use the LoRa dataset's temporal structure to evaluate LSTM-based temporal modeling and mobile scenario adaptation techniques. The multi-timescale coverage supports assessment of short-term channel variations, medium-term aging patterns, and long-term stability degradation. We present the detailed experimental protocols and mobile scenario simulation methodologies in that chapter.

3.3.5 WiSIG: Large-Scale WiFi Dataset

The WiSIG (WiFi Signal) dataset represents the largest publicly available collection for WiFi-based RF fingerprinting research, providing scale and experimental diversity for evaluating receiver and channel agnostic fingerprinting [48]. This dataset addresses limitations in existing RF fingerprinting collections by incorporating large numbers of both transmitters and receivers, supporting evaluation of cross-receiver generalization and unknown device detection.

Dataset Scale and Unique Characteristics

WiSIG encompasses 10 million WiFi packets transmitted by 174 off-the-shelf WiFi transmitters and captured by 41 USRP receivers across four distinct capture sessions performed over one month. In terms of both transmitter and receiver diversity, this is the largest publicly available RF fingerprinting dataset, it addresses a gap in the fingerprinting research where most existing datasets consider at most 100 transmitters and rely on single receiver configurations [48].

The experimental infrastructure uses the Orbit testbed [49], consisting of a 20×20 grid of nodes with approximately 1 meter separation. Each node contains a roof-mounted computer equipped with WiFi radios, with selected nodes augmented with USRP receivers. Additional USRP receivers positioned within massive MIMO racks in the same environment provide spatial diversity for signal collection.

Key Technical Specifications:

- **Scale:** 174 transmitters, 41 USRP receivers, 10 million packets
- **Temporal Coverage:** 4 capture sessions over 1 month
- **Data Volume:** 1.4 TB (Raw), 76.9 GB (Processed)
- **Signal Format:** 256 IQ samples per identification signal
- **Processing Stages:** Raw captures, equalized/non-equalized variants, pre-configured subsets

The dataset uses a three-stage processing pipeline that balances accessibility with experimental flexibility. Raw WiSIG provides complete USRP captures (1.4 TB), Full WiSIG contains extracted identification signals including both non-equalized and equalized variants using WiFi preambles (76.9 GB), and packaged subsets provide pre-configured collections for specific applications: ManyTx (150 transmitters, device diversity), ManyRx (32 receivers, cross-receiver evaluation), ManySig (10,000 signals per device, statistical analysis), and SingleDay (temporal consistency assessment).

Evaluation Capabilities and Integration

WiSIG's scale and device coverage make it uniquely suitable for evaluating scalability and unknown device detection. The multi-receiver architecture creates natural domain shift scenarios that challenge traditional closed-set fingerprinting assumptions. We can use the dataset to evaluate cross-receiver generalization across 41 diverse receivers, large-scale device identification with 174 transmitters approaching realistic deployment scenarios, temporal stability assessment across multiple capture sessions, and channel-agnostic evaluation under diverse propagation conditions.

Figure 3.1: WiSIG experimental infrastructure showing Orbit testbed architecture with 20×20 grid configuration and USRP receiver placement

In Chapter 7, we use WiSIG to evaluate the MADE architecture's open-set recognition capabilities and anomaly detection performance. The receiver diversity allows us to assess architectural robustness across different receiver positions, while the device population supports evaluation of scalability limits for our autoencoder-based reconstruction approach. We present the detailed experimental protocols and performance analysis in Chapter 7.

3.4 Standardized Evaluation Methodology

We employ consistent evaluation principles across all datasets to ensure fair comparison and statistical validity. While each technical chapter presents its own detailed experimental protocols, preprocessing procedures, and performance analyses, this section describes the overarching methodological principles that guide all our evaluations.

3.4.1 Signal Preprocessing and Quality Assurance

Signal preprocessing follows protocol-specific procedures appropriate to each dataset while maintaining consistency in quality assurance standards. Our preprocessing pipeline incorporates the following components:

Amplitude Normalization: We normalize signals to remove variations arising from transmission power differences and propagation path loss. The normalization procedures preserve relative amplitude characteristics while standardizing signal magnitude across devices and collection sessions [50, 17].

Sequence Extraction: Protocol-specific sequence extraction procedures isolate device-specific transmission characteristics. For WiFi signals, preamble-based extraction captures hardware transients. For UAV controllers, burst extraction isolates individual transmission events. For LoRa devices, preamble sequences provide consistent fingerprinting characteristics. We document the detailed extraction procedures specific to each protocol in the relevant technical chapters.

Quality Filtering: We apply quality assurance procedures to filter signals based on signal quality assessment. Figure 3.2 illustrates the quality checkpoint framework we apply across all datasets. Signals undergo SNR assessment, corruption detection to identify and remove artifacts, and protocol conformance validation to verify proper signal structure. We exclude signals failing quality thresholds from both training and evaluation sets.

The specific preprocessing parameters, quality thresholds, and sequence extraction procedures vary by protocol and we detail them in Chapters 4 through 7, where each technical contribution describes the preprocessing applied to its evaluation datasets.

3.4.2 Experimental Design and Validation

We follow rigorous experimental design principles to ensure reliable and reproducible evaluation results across all technical contributions:

Train/Validation/Test Splits: We partition datasets following procedures that respect temporal relationships and device distributions. For datasets with temporal structure (LoRa, GLOBECOM22, ACMWiSec21), our splits maintain temporal ordering to avoid data leakage. For cross-transmission evaluations (UAV datasets), we train on limited transmission parameters (Tx1–Tx2) and evaluate on unseen parameters (Tx3–Tx19). We document split ratios and specific partitioning strategies in each technical chapter.

Cross-Validation Protocols: Our cross-validation strategies account for dataset-specific characteristics including device distribution, temporal struc-

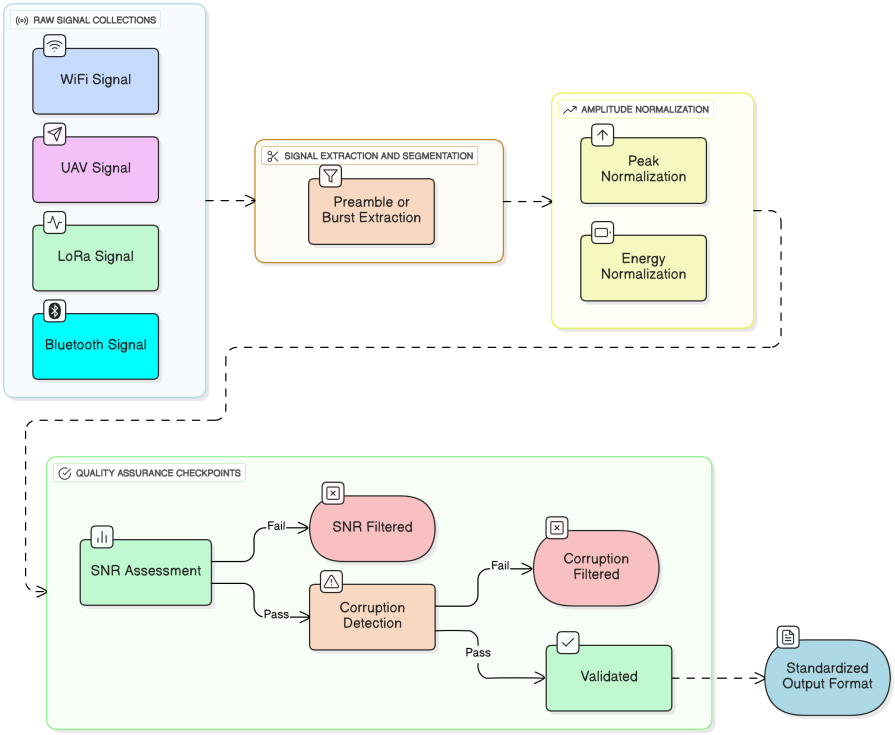


Figure 3.2: Quality assurance checkpoint framework applied across all datasets, showing protocol-specific signal inputs, preprocessing stages, and quality validation procedures leading to standardized output format

ture, and transmission parameter variations. The choice of validation strategy depends on dataset scale and experimental objectives.

Performance Metrics: We evaluate using standardized metrics including classification accuracy, precision, recall, and F1-scores.

Statistical Validation: We employ appropriate statistical significance testing for performance comparisons to ensure that observed differences represent genuine improvements rather than random variation [51]. We apply confidence intervals and statistical tests according to experimental design and sample sizes available in each evaluation.

3.4.3 Reproducibility Standards

We ensure reproducibility through documentation and standardized reporting:

Dataset Documentation: All datasets we employ are publicly available. We document dataset citations, access procedures, and preprocessing code to support independent verification.

Implementation Details: We document software frameworks, library versions, and hyperparameter configurations in each technical chapter. Where feasible, code repositories provide reference implementations for reproduction of experimental results.

Experimental Conditions: We document hardware specifications, computational resources, and training procedures to provide complete experimental context, including GPU/CPU specifications, memory requirements, and training duration for each approach.

The standardized evaluation methodology ensures that our technical contributions in Chapters 4 through 7 undergo rigorous assessment under consistent conditions while accommodating the specific requirements of each approach and dataset combination.

3.5 Dataset Acquisition Challenges and Community Recommendations

The construction of the six-dataset portfolio for this thesis required navigating substantial practical barriers that are rarely discussed in the RF fingerprinting literature but represent significant obstacles to research progress in this field. This section documents the challenges encountered during dataset acquisition and integration, and proposes recommendations to the research community based on this experience. These observations complement the published analysis in our survey work [39, 34], where the scarcity and limited diversity of publicly available RF datasets was identified as one of the primary barriers to generalization research.

3.5.1 Acquisition Barriers

Specialized Equipment Requirements. Capturing RF signals suitable for fingerprinting research requires equipment and expertise that most machine learning researchers do not possess. A Software-Defined Radio (SDR) platform capable of capturing wideband I/Q samples with sufficient dynamic range and phase coherence, such as the USRP N210 used in the LoRa dataset collection or the high-resolution oscilloscopes used in UAV controller capture, represents a significant capital investment. Beyond the hardware itself, operating these platforms requires expertise in RF system calibration, carrier frequency offset

correction, gain control, and protocol-specific timing synchronization. Errors in any of these steps introduce systematic biases that corrupt the hardware-specific signatures the research aims to capture, making uncalibrated or poorly configured collection setups unsuitable regardless of the volume of data collected. This barrier means that high-quality RF fingerprinting datasets can only be produced by research groups with both RF engineering expertise and machine learning research objectives, a combination that remains relatively rare.

Domain Expertise for Collection Protocol Design. Beyond equipment operation, designing a collection protocol that supports the intended generalization evaluation requires deep understanding of both the wireless protocol being captured and the fingerprinting challenge being studied. The UAV controller datasets used in Chapter 6, for example, require systematic variation of transmission parameters across 19 controlled configurations to enable cross-transmission generalization evaluation. Designing this protocol requires knowledge of which parameters to vary, the range of operationally realistic values, how to control confounding factors, and how to ensure that parameter variations are isolated rather than correlated with environmental changes. Without this domain knowledge, collected datasets may fail to capture the generalization challenge they were intended to evaluate, producing results that do not reflect real deployment conditions despite significant collection effort.

3.5.2 Data Sensitivity and Sharing Constraints

The most significant barrier to expanding the public RF fingerprinting dataset ecosystem is not technical but organizational: the entities that collect the most operationally relevant RF data are precisely those with the strongest reasons not to share it publicly.

Defense organizations, spectrum regulatory agencies, telecommunications operators, and critical infrastructure operators collect RF signal data from operational networks as part of their monitoring and security activities. This data would be invaluable for fingerprinting research because it reflects real operational conditions, genuine device diversity, and authentic threat scenarios. However, sharing such data publicly risks exposing sensitive information about network topology, device capabilities, operational patterns, and security posture. The UAV controller fingerprinting domain is particularly affected by this constraint: military and security applications generate the most relevant operational data, but classification and export control concerns prevent public release. This sensitivity paradox, the most useful data being the least share-

able, means that public datasets necessarily represent a biased sample of the operational landscape, skewed toward academic laboratory collections with consumer-grade devices in controlled environments. The lab-to-field performance gap documented throughout this thesis is, in part, a direct consequence of this data availability gap: models trained on laboratory data that does not reflect operational conditions cannot be expected to generalize to those conditions.

3.5.3 Scale and Protocol Diversity Limitations

The six datasets used in this thesis represent the most suitable publicly available options at the time of collection. Yet even this carefully selected portfolio reveals fundamental scale and diversity limitations. WiSIG, the largest publicly available WiFi fingerprinting dataset, contains 174 transmitters, an unprecedented scale for academic research but modest compared to real IoT deployments that may involve thousands of devices from hundreds of manufacturers. The UAV controller datasets cover eight to seventeen consumer drone models, while the operational UAV market includes hundreds of distinct platforms with diverse RF characteristics. The LoRa-60 dataset covers a single European frequency band (868.1 MHz), leaving the North American 915 MHz band and Asian 433 MHz band without comparable public fingerprinting datasets. Cellular protocols (4G LTE, 5G NR) and emerging standards (Wi-Fi 7, Bluetooth 5.3) lack public fingerprinting datasets entirely despite representing the dominant wireless protocols in deployed IoT systems.

3.5.4 Recommendations for the Research Community

Based on the experience accumulated through this thesis and the broader dataset survey work [39], we offer the following recommendations to the RF fingerprinting research community:

Standardized collection protocols. Adopting shared protocols for signal capture, annotation, and format would enable datasets from different research groups to be combined into larger, more diverse collections. A reference protocol specifying minimum SDR requirements, calibration procedures, metadata standards, and quality validation thresholds would significantly reduce the barriers to contributing new datasets and comparing results across research groups.

Federated and privacy-preserving sharing mechanisms. Organizations holding operationally sensitive RF data should be enabled to contribute to research without exposing raw signals. Federated learning frameworks that

train models on locally held data without centralizing signals, and differential privacy mechanisms that add calibrated noise to protect individual transmission characteristics, represent technically feasible paths toward extracting research value from sensitive datasets while managing disclosure risk.

Synthetic data generation with hardware validation. Physically grounded simulation tools that model hardware imperfection sources, oscillator phase noise, power amplifier nonlinearity, I/Q imbalance, at sufficient fidelity to support fingerprinting research could partially compensate for dataset scarcity. Such tools must be validated against real hardware measurements to ensure that synthetic fingerprints reflect genuine manufacturing variation rather than simplified statistical models that do not generalize.

A community benchmark dataset. The RF fingerprinting field would benefit significantly from a large-scale, multi-protocol, multi-condition reference dataset analogous to ImageNet in computer vision, serving as a standardized evaluation benchmark against which new methods can be consistently compared. Such a benchmark should cover at minimum WiFi, Bluetooth, LoRa, and one cellular standard, include temporal collections spanning multiple months, incorporate both controlled and uncontrolled environments, and scale to at least several hundred devices. Coordinated collection by a consortium of academic and industry partners, following standardized protocols, represents the most realistic path to achieving this benchmark.

3.6 Chapter Summary

In this chapter, we established the experimental foundation for evaluating RF fingerprinting approaches across the generalization challenges that limit practical deployment. Our six-dataset portfolio covers WiFi, UAV control links, LoRa, and Bluetooth protocols with device scales ranging from 5 to 174 transmitters, and operational conditions spanning controlled laboratory settings to realistic deployment scenarios.

The portfolio addresses complementary evaluation dimensions: WiFi and Bluetooth datasets enable assessment of statistical enhancement methods and cross-day generalization (Chapter 4); UAV controller datasets provide controlled cross-transmission parameter evaluation for progressive learning approaches (Chapter 6); the LoRa IoT dataset supports temporal modeling and mobile scenario adaptation assessment (Chapter 5); and WiSIG enables large-scale unknown device detection and cross-receiver generalization evaluation (Chapter 7).

Our standardized evaluation methodology, with consistent preprocessing

procedures, quality assurance protocols, and experimental design, ensures fair comparison across the different approaches while maintaining rigor. Each technical chapter presents detailed experimental protocols specific to its contributions, while this chapter provides the unified framework that connects these evaluations and enables comparison across approaches and datasets.

Chapter 4

4 Statistical Feature Enhancement for RF Fingerprinting

4.1 Introduction

The generalization challenges we identified in Chapter 2 represent major barriers to practical RF fingerprinting deployment in UAV and IoT environments. Cross-transmission parameter variations cause accuracy degradation of 20–70%, while temporal variations introduce cross-day performance drops from over 95% to below 60% [30]. These failures prevent current approaches from meeting reliability requirements for security-critical applications.

In this chapter, we present statistical feature enhancement as our foundational solution to temporal generalization challenges. The approach uses multi-domain feature extraction across time, frequency, statistical, and fundamental characteristics domains, combined with sliding window temporal analysis and summary statistics aggregation. Unlike approaches that rely on single-domain characteristics or raw signal processing, statistical enhancement provides broad signal characterization that maintains discriminative capability while achieving better temporal robustness.

We evaluate the approach using the GLOBECOM22 and ACMWiSec21 datasets established in Chapter 3 (Section 3.3.2), with systematic assessment across same-day (Train Test Same time frame Same Day - TTSS), cross-time-frame same-day (Train Test Different time frame Same day - TTDS), and cross-day (Train Test Different time frame Different day - TTDD) scenarios. The approach achieves 99.6% accuracy in controlled conditions with meaningful improvements in challenging cross-day scenarios (52.1% versus 37.5% for traditional methods), which establishes the foundation for the advanced techniques we present in subsequent chapters.

The work presented in this chapter has been published as "Advanced security frameworks for UAV and IoT: A deep learning approach" in *Internet of Things* (Elsevier) [30]. Section 4.2 presents the complete published work. Section 4.3 provides our critical analysis examining the approach's contributions, its integration with the experimental framework from Chapter 3, and the limitations that motivate the advanced techniques we developed in Chapters 6 through 7. Section 4.4 summarizes the key contributions and positions statistical enhancement within the generalization framework developed across the thesis.

4.2 Statistical Feature Enhancement

This section presents the complete published work presenting our statistical feature enhancement methodology for addressing cross-day temporal generalization in RF fingerprinting. We demonstrate that multi-domain feature extraction combined with sliding window temporal analysis and summary statistics aggregation yields marked performance improvements over traditional approaches, while maintaining computational efficiency suitable for resource-constrained UAV and IoT deployments.

The methodology encompasses systematic extraction of 14 features across time, frequency, statistical, and fundamental characteristics domains; sliding window segmentation with non-overlapping regions of interest that capture temporal patterns; and summary statistics vectors that aggregate mean, variance, standard deviation, extreme values, and percentiles to characterize feature distributions across varying operational conditions.

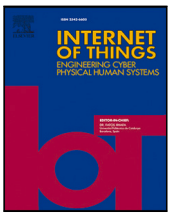
We validated the approach using the GLOBECOM22 and ACMWiSec21 datasets across multiple evaluation scenarios. The recursive feature elimination process identifies magnitude, phase, power, entropy, and spectral flatness as the most discriminative features, achieving substantial dimensionality reduction while maintaining superior classification performance. The complete experimental protocols, mathematical formulations, feature definitions, and performance analyses are presented in the published work below.

Internet of Things 32 (2025) 101594

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.elsevier.com/locate/iot



Advanced security frameworks for UAV and IoT: A deep learning approach

Nordine Quadar [a,1], Abdellah Chehri [a,1,*], Benoit Debaque [b,1]

[a] Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, K7K 7B4, Canada
[b] Thales Research and Technology, Quebec, G1P 4P5, Canada

ARTICLE INFO

Keywords:
Advanced Security
IoT
Deep learning
RF fingerprinting

ABSTRACT

The integration of unmanned aerial vehicles (UAVs) has opened new avenues for enhanced security and functionality. The security of UAVs through the detection and analysis of unique signal patterns is a critical aspect of this technological advancement. This approach leverages intrinsic signal characteristics to distinguish between UAVs of identical models, providing a robust layer of security at the communication level. The application of artificial intelligence in UAV signal analysis has shown significant potential in improving UAV identification and authentication. Recent advancements utilize deep learning techniques with raw In-phase and Quadrature (I/Q) data to achieve high-precision UAV signal recognition. However, existing deep learning models face challenges with unfamiliar data scenarios involving I/Q data. This work explores alternative transformations of I/Q data and investigates the integration of statistical features such as mean, median, and mode across these transformations. It also evaluates the generalization capability of the proposed methods in various environments and examines the impact of signal-to-noise ratio (SNR) on recognition accuracy. Experimental results underscore the promise of our approach, establishing a solid foundation for practical deep-learning-based UAV security solutions and contributing to the field of IoT.

1. Introduction

The rapid advancement of smart devices and consumer electronics, such as those equipped with Radio Frequency (RF) capabilities, has ushered in an era of unparalleled connectivity and convenience, presenting myriad opportunities. However, this proliferation also brings forth significant cybersecurity concerns for embedded systems. As these devices become crucial components of smart cities, vehicular networks, and various industrial applications, their vulnerabilities have become increasingly evident.

The susceptibility to passive attacks, such as unauthorized eavesdropping, alongside active threats like information manipulation, spoofing, and denial of service attacks, raises serious alarms regarding the security and integrity of these networks [1]. These concerns extend to emerging technologies like Unmanned Aerial Vehicles (UAVs). Integrating artificial intelligence into RF-based UAV signal analysis has shown significant potential in enhancing UAV identification and authentication [2].

The rapid evolution of IoT technologies has amplified these security challenges. As highlighted by Rangarajan and Al-Quraishi in [3], the convergence of emerging technologies within IoT ecosystems requires robust security frameworks that can adapt to evolving threats. This is particularly crucial in UAV systems, where security breaches could have significant implications for both operational safety and data integrity.

* Corresponding author.
E-mail addresses: quadar@rmc.ca (N. Quadar), chehri@rmc.ca (A. Chehri), benoit.debaque@ca.thalesgroup.com (B. Debaque).
[1] All authors contributed equally to this work.

https://doi.org/10.1016/j.iot.2025.101594

Received 12 January 2025; Received in revised form 16 February 2025; Accepted 23 March 2025
Available online 22 April 2025

In this interconnected landscape, a secure UAV identification becomes paramount. When considering the network requirements, it is important to prioritize the security aspect to select the most secure route and path for link transmission. Despite the use of access authentication and encrypted protocols as traditional security measures, recent studies reveal that malicious users can exploit vulnerabilities by stealing or fabricating identifying information, thereby underscoring the pressing need for enhanced security measures at the foundational layers such as the physical layer [4].

Therefore, exploring innovative approaches to enhance the security of unmanned aerial vehicles (UAVs) becomes imperative. In this context, the adoption of RF fingerprinting emerges as a promising solution, offering a new and efficient identification method that holds significant potential for reinforcing the resilience of these interconnected systems. The main objective of RF fingerprinting is identifying transmitters by analyzing inherent characteristics in the radio waveform, known as RF fingerprints [5].

These distinctive characteristics emerge from the devices' internal RF circuitry. Imperfections inherent in the manufacturing process introduce variations in components such as the power amplifier (PA), low noise amplifier (LNA), clock circuits, and local oscillators (LO), among others. Elements like IQ imbalance, clock skew, and out-of-band (OOB) spurious leakage exhibit differences even among devices from the same manufacturer. Despite being minuscule, these unique features collectively form the RF signature of a wireless device. Through meticulous signal processing techniques, this RF signature can be extracted without prior knowledge or induced impairments.

This research explores alternative transformations of I/Q data and examines the incorporation of statistical features such as mean, median, and mode across these transformations. The study evaluates the generalization capability of the proposed methods in various environments and investigates the influence of signal-to-noise ratio (SNR) on classification accuracy. Experimental results underscore the promise of our approach, laying a solid foundation for practical deep-learning-based UAV security solutions, contributing significantly to the field of UAV-assisted consumer electronics [2].

These fingerprints pose a tough challenge for malicious users, making them a good solution to this cybersecurity issue. [6]. The process of RF fingerprinting involves several steps. Firstly, the radio waveform is received and preprocessed; subtle characteristics are extracted and fused from the received waveform; and finally, the identification of the device sending the radio waveform takes place. Thus, RF fingerprinting aligns itself as a typical classification task, where the challenging aspect lies in extracting features that accurately reflect the subtle hardware differences or defects from the signal waveform. This extraction process becomes tough when dealing with blind extraction from commercial-off-the-shelf (COTS) Internet of Things (IoT) devices of the same family. The recent studies in deep learning (DP), however, have demonstrated a significant impact on the wireless security field [7–9].

RF fingerprinting, situated within the broader domain of signal intelligence applications [10], such as modulation and signal classification, faces distinctive challenges related to the inherent minute characteristics of hardware. These challenges include the aging of hardware components, the impact of wireless standards on fingerprint features, multipath propagation effects due to obstacles, walls, environmental changes, location alterations, and noise and interference. For instance, capturing the hardware-intrinsic features becomes especially demanding with the frequency-hopping nature of some waveforms, such as Bluetooth. The performance in such cases relies significantly on factors like input preprocessing, input sample length, features selection, and the underlying DL architecture [11].

Conventional RF fingerprinting approaches require the manual design of complex feature engineering customized for specific communication systems. For example, the use of transmitter hardware imperfections, such as carrier frequency offset (CFO), instantaneous frequency in-phase and quadrature (I/Q) offset, and I/Q imbalance, have been discussed in [12,13] to authenticate the user identity of radio devices.

Unfortunately, these approaches require a comprehensive understanding of the communication protocol. They may demonstrate limited efficacy in the presence of environmental variations, such as fluctuations in temperature and electromagnetic interference [13]. In contrast, recent RF fingerprinting approaches leverage deep convolutional neural networks (CNN), which automatically learn features from the data. This shift has proven some effectiveness, as CNN demonstrates the ability to automatically learn more effective and universal RF fingerprints compared to traditional feature engineering methods [14–16].

Despite these advancements, challenges with model generalization persist. For instance, even with attention mechanisms, CNN models only reached 27.5% accuracy when dealing with unseen data in Bluetooth RF fingerprinting scenarios [16]. Addressing these challenges is crucial to ensure the practical operational efficiency of DL models. Hence, enhancing neural network structures and refining feature selection is essential for obtaining more distinctive and reliable RF fingerprints.

In this study, we explore recent advancements in applying DL techniques to identify and classify diverse devices through passive examination of the RF spectrum. The existing literature has introduced numerous approaches, each using its distinctive dataset and methodologies, challenging direct comparisons.

Some studies, such as work in [17,18], focus on classification while keeping signals in their 1D format by using raw in-phase and quadrature (IQ) samples as inputs for their proposed models. Conversely, the work outlined in [16] tried classification after a 3D transformation, utilizing magnitude, phase information of the decimated signal, and power spectral density (PSD).

Considering these factors, our study aimed to investigate further possible changes to the input signal. We explored potential improvements by including statistical and non-parametric characteristics like average, middle value, and most frequent value among these changes. Simply put, the wireless signal detection and classification field using RF fingerprinting has been started but is still in its early stages. As a result, a series of experiments was carried out to assess the influence of different parameters of the RF signal on the overall system's accuracy. We aim to gain insights by comparing our findings with the methodologies used by various experts in this field.

The remainder of this paper is structured as follows: Section 2 provides a brief overview of recent research progress in RF fingerprinting, focusing on generalization issues. Section 3 introduces our proposed RF fingerprinting feature generation module. Section 4 presents the experimental scenarios and the corresponding experimental results and analysis. Finally, Section 5 outlines the future directions of our work, and Section 6 concludes the paper.

2. Related work

Most of the traditional works have used feature extraction techniques like wavelet transform analysis [19], I/Q imbalance [20], and radio turn-on transient-based methods [21]. These techniques are specifically designed at the physical layer to distinguish wireless devices uniquely.

Numerous approaches have been explored to enhance the security of wireless devices operating on standards like WiFi, Bluetooth, and LoRa, making it a subject of active research. Moreover, with the recent surge in machine learning, DL has emerged as a tool to address some challenges [22]. However, there remains to be a gap in the application of DL, especially a lightweight deployable framework that enhances the generalization capability for fingerprinting real-world wireless devices.

In the study presented by [23], the identification of ZigBee devices was explored through the generation and utilization of non-parametric features, such as mean, median, mode, and linear model coefficients. The research involved capturing complex IQ samples from four Texas Instruments ZigBee CC2420 devices using an Agilent E3238S receiver. The received signals' phase variable was computed, and the preamble region of the phase variables was segmented into 32 equally sized Regions of Interest (ROIs). Subsequently, non-parametric features were calculated for each ROI. A random forest classifier with 1000 trees was employed to classify the signals. The individual use of the four non-parametric features for device classification produced results indicating classification accuracy exceeding 97% for signal-to-noise ratio (SNR) of 10 dB.

Additionally, the study compared the performance of non-parametric features with parametric features, including variance, skewness, and kurtosis, computed for each ROI. The non-parametric features exhibit improvements of up to 9% over their parametric counterparts. In another study [24], researchers focused on extracting modulation-specific features, including differential constellation trace figure, carrier frequency offset, modulation offset, and I/Q offset, to fingerprint ZigBee devices. The reported accuracy of approximately 95% on a 54-radio testbed showcased the method's effectiveness. However, the study acknowledges a limitation by not considering potential channel influences, such as the multipath effect, on the classification scheme.

In [25], the authors introduce a multidimensional permutation entropy-based RF fingerprinting method. Permutation Entropy (PE), measuring the complexity of a given time series, was employed to extract and amplify minuscule changes in the time signal. The method involved capturing radio signals, extracting signal envelopes, and calculating multidimensional permutation entropy for the envelope to create the RF fingerprint feature vector. Classification was performed using a support vector machine (SVM) classifier. Evaluation entails collecting 100 data sets from three AKDS700 radios using a digital receiver and an oscilloscope. Multidimensional permutation entropy was computed for all signals, and the SVM trained on these features achieved an average accuracy of 90% for SNR 10 dB in recognizing the three radio devices.

In another study [26], the authors introduced an RF fingerprinting method designed for identifying IoT devices, employing more entropy-based statistical features, namely PE and Dispersion Entropy (DE). The evaluation of this method involved nine nRF24LU1+ IoT devices, and RF signals from these devices were captured using an N210 USRP with an XCVR2450 front end. The real-valued In-Phase and Quadrature (IQ) samples were captured using the USRP, followed by synchronization and normalization to obtain the burst of traffic associated with the payload. Statistical features were computed for each received payload, including variance, skewness, kurtosis, Shannon entropy, log entropy, PE, and DE. Three classifiers, k-NN, SVM, and decision tree, were trained using a subset of the ten features mentioned above. The authors demonstrate that the classifier trained with PE and DE features and statistical features achieves an accuracy 24% to 30% higher than the classifier trained solely with statistical features (Shannon entropy and log entropy). Notably, using only the PE feature alongside statistical features substantially improves accuracy compared to using Shannon entropy and log entropy alone. Finally, the authors illustrate that all three classifiers exhibit similar classification accuracy when trained with PE and DE features in conjunction with statistical features.

Authors in study [27] proposed a Multi-Fingerprint and Multi-Classifier Fusion (MFMCF) localization method for RF fingerprinting. The technique aims to enhance WiFi Access Points (APs) localization accuracy by constructing composite fingerprints and employing multiple classifiers. A composite fingerprint set (CFS) was formed, incorporating Received Signal Strength (RSS), Signal Strength Difference (SSD), and Hyperbolic Location Fingerprint (HLF) features. The decision structure involved three classifiers: k-NN, SVM, and random forest, contributing to a more accurate location estimate. RSS data from seven APs at 35 points in an indoor location were collected, with 100 RSS data recorded for each AP at each location.

We used the Grubbs method to identify outliers in the RSS data and replaced them with Gaussian random numbers. The SSD and HLF fingerprints were generated using the RSS data that was collected. The dimensions of the CFS were reduced using linear discriminant analysis (LDA), and this reduced dimension was then used to train the three classifiers. During the testing process, we calculated the entropy of each classifier. The classifier with the lowest entropy was then selected for location estimation.

To evaluate the proposed MFMCF method, 12 features from the 49 features in CFS were used, covering over 95% of the information. The reported probability of zero positioning error for MFMCF was 96.5%, representing an increase of 4.2%, 6.4%, and 7.7% compared to RSS, SSD, and HLF, respectively, when used as independent fingerprint features. In comparison with independent classifiers, where CFS trains and tests individual classifiers, the probability of zero positioning errors for MFMCF, RF, k-NN, and SVM was 96.5%, 90.2%, 92.9%, and 94.8%, respectively. The authors highlight the superiority of the proposed MFMCF method over independent classifiers.

In [28], researchers introduced a novel approach known as RF distinct native attribute (RF-DNA) for RF fingerprinting, aiming to identify ultra-wideband noise radar-emitting devices. The method involves extracting key RF-DNA fingerprint features, including variance, skewness, and kurtosis, from the time-domain response of signals. Additionally, the authors derived features like normalized power spectral density (PSD) and discrete Gabor transform from the spectral-domain response. Signal classification was conducted using the multiple discriminant analysis with maximum likelihood (MDA/ML) classifier and the generalized relevance

learning vector quantization improved (GRLVQI) classifier. Notably, the results showcase the method's effectiveness, achieving accuracies above 80% in classifying defective units.

In [29], the study introduces LoRa emitter fingerprinting utilizing a spectrogram-based Convolutional Neural Network (CNN). The focus of this work was on the carrier frequency offset (CFO) of LoRa emitters. Model generalization was evaluated across different days in a wired setup where emitters are cabled over an attenuator to the receiving USRP radio. In [30], the authors employ a triplet loss-based CNN model for fingerprinting base stations transmitting 5G New Radio, LTE, or WiFi waveforms. However, it is important to note that these base stations are software-defined radio-based rather than real-world base stations, emitting synthetically generated waveforms with MATLAB's LTE, WLAN, and 5G toolboxes. The study included a generalization test where training and testing were conducted on different days. However, the emissions faced challenges in terms of multipath effects, fading, and orientation in the Time-Transfer-Difference (TTD) scenario, which closely resembles real-world deployment settings.

The work done in [31] presents a comprehensive set of comparative experiments aimed at systematically analyzing the impact of environmental factors, including signal-to-noise ratio (SNR), dynamic channel conditions, and the number of targets. Through precise control of these factors and assessing the performance of different algorithms in diverse situations, researchers gain valuable insights into the strengths and limitations of various approaches to RF fingerprinting.

In another study [32], authors suggest a multi-sampling network for fingerprinting devices using signals with varying sampling rates. Their study highlights the notable effects of dynamic environments on RF fingerprinting. In efforts to reduce these effects, they commonly employ signal preprocessing techniques, such as denoising, channel estimation, data augmentation, and other methods. Nevertheless, the optimization of neural network structures to extract more discriminative and robust RF fingerprints remains a challenging aspect of the research, underscoring the need for further exploration in this domain.

Although the introduction of innovative techniques in the literature, challenges related to model generalization persist as a primary constraint in attaining optimal performance for trained and deployed DL models within real-world operational environments. Even with integrating state-of-the-art attention mechanisms into DL models, as presented in [16], the model's accuracy under unseen data remains limited, reaching only 27.5%. Consequently, research efforts in this domain must prioritize validating the model's generalization capability, accounting for the impact of confounding factors in the assessment. This study addresses this need by evaluating the utilization of non-parametric features across various scenarios, leveraging two distinct datasets for comprehensive analysis.

3. Methodology

In RF fingerprinting, selecting features is crucial for accurate device identification. The process should involve extracting meaningful features from IQ samples that capture the unique characteristics of each device's RF signal. The literature showed that the direct use of IQ samples as input to the ML models does not provide the best performance, especially when it comes to generalization capability [33]. On the other hand, using statistical data or derived features from the IQ samples improves the performance [5]. Based on this, our work tries to use these two techniques and design a new feature generation method that considers parametric and non-parametric metrics of extracted features.

This section offers a comprehensive analysis of the design of the proposed RF fingerprint generator. A sliding window approach segments each transmission in both the training and test sets, serving as a foundation for further analysis.

Inspired by the approach introduced in [33], we use a sliding window to form equal size and non-overlapping regions of interest (ROIs) representing the transmitted I/Q raw. As shown in 1, given a time-series I/Q sequence k of length M_k, we divide the transmission to generate $Nk = M_k/j$ slices of length j. We generate a list of RF characteristics for each slice, such as power density, entropy, spectral flatness, and magnitude. This new raw data is then divided to generate $N_k - i + 1$ slice of length i by sliding a window with stride 1. Finally, we produce the summary statistics vectors for each new slice that will be used as input for the machine learning model. Each vector is then labeled individually, with the ID of the device that generated the transmission as a label.

More formally, given an input sequence of I/Q samples:

$$S = \{(I_i, Q_i) | i = 1, \ldots, M_k\} \tag{1}$$

The first transformation divides S into N_k slices:

$$S_j = \{(I_i, Q_i) | i = (j-1)w + 1, \ldots, jw\} \tag{2}$$

where $j = 1, \ldots, N_k$ and w is the window size.

For each slice S_j, we compute features across different domains:

- Time domain: directly from I/Q values
- Frequency domain: after FFT transformation
- Statistical domain: using probability distributions
- Fundamental characteristics: using signal properties

76

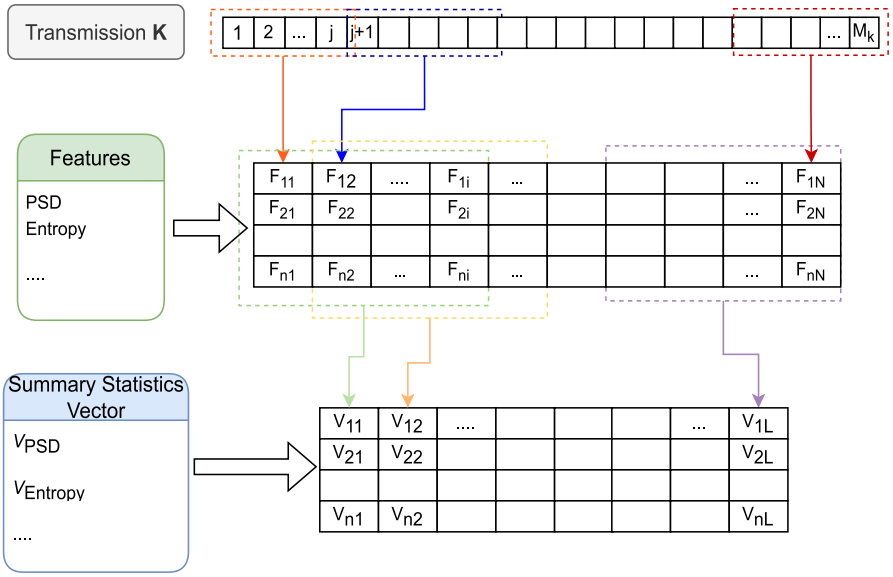


Fig. 1. Feature generation using slicing window.

This multi-domain approach ensures we capture diverse signal characteristics that could be unique to each device.

The feature generation process can be formally defined as:

$$F_{Device_k} = \{f_{ij} | i \in [1, n], j \in [1, N_k]\} \tag{3}$$

where f_{ij} represents the ith feature for the jth slice of the total n features and N_k slices. Each slice contains j consecutive I/Q samples from the original transmission.

For each feature F, we compute the summary statistics vector SV_F as defined in the following equation:

$$SV_F = [\mu(F), \sigma^2(F), \sigma(F), min(F), max(F), P_p(F)] \tag{4}$$

where $\mu(F)$ is the mean value, $\sigma^2(F)$ is the variance, $\sigma(F)$ is the standard deviation, $min(F)$ and $max(F)$ are the minimum and maximum values, and $P_p(F)$ represents the vector of percentiles of the slice F.

For instance, in our experiments with 20 million I/Q samples:

1. First transformation: Split into slices of 200 samples each
2. Feature computation: Calculate 14 features per slice
3. Summary statistics: Compute statistical measures across features
4. Final output: Feature vectors ready for deep learning model input

As discussed in [16], to achieve ubiquity, transparency, and robustness across various capture instances, the fingerprinting method should leverage unprocessed IQ samples obtained through passive signal reception across a range of wireless protocols. Using the same concept, we use unprocessed captured signals. Our chosen features are grouped into distinct categories, each designed to capture specific aspects of the RF signal.

3.1. Time-domain features

In the time-domain category, features such as mean and standard deviation of in-phase (I) and quadrature (Q) components were selected. The mean values provide insights into the average strength of the signal in both the I and Q dimensions. At the same time, the standard deviations quantify the variability or spread of these components. These features are fundamental for characterizing the temporal dynamics of the RF signal, offering valuable information about the central tendencies and fluctuations over time.

For a given slice containing N samples, these features are computed as:

Mean of I/Q components:

$$\text{Mean\_I} = \frac{1}{N}\sum_{i=1}^{N} I_i, \quad \text{Mean\_Q} = \frac{1}{N}\sum_{i=1}^{N} Q_i \tag{5}$$

Standard deviation of I/Q components:

$$\text{Std\_I} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i - \bar{I})^2}, \quad \text{Std\_Q} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Q_i - \bar{Q})^2} \tag{6}$$

These statistical measures capture device-specific characteristics arising from hardware imperfections in the RF chain. I/Q imbalance, a common hardware imperfection, manifests as differences between I and Q component statistics, creating unique patterns that persist across transmissions from the same device.

3.2. Frequency-domain features

Within the frequency-domain category, the inclusion of features like frequency centroid and spectral flatness aims to capture unique spectral characteristics. The frequency centroid represents the center of mass of the frequency distribution, offering insights into dominant frequency components. Spectral flatness, on the other hand, provides information about the flatness or peakiness of the spectrum, contributing to the identification of distinctive frequency patterns. These features are particularly pertinent for discerning spectral signatures crucial for device differentiation.

Given the power spectral density P_i at frequency bin f_i, these features are computed as:

Frequency Centroid:

$$\text{Freq\_Centroid} = \frac{\sum_{i=1}^{N} f_i \cdot P_i}{\sum_{i=1}^{N} P_i} \tag{7}$$

Spectral Flatness:

$$\text{Spectral\_Flatness} = \frac{N\sqrt{P_1 \cdot P_2 \cdot \ldots \cdot P_N}}{\frac{1}{N}\sum_{i=1}^{N} P_i} \tag{8}$$

These frequency-domain features are particularly effective at capturing RF front-end characteristics such as local oscillator imperfections and filter response variations, which create device-specific spectral signatures.

3.3. Statistical features

The statistical category involves a set of features, including kurtosis, skewness, and entropy. Kurtosis and skewness of both magnitude and phase distributions offer valuable statistical insights into the shape and asymmetry of the signal distribution. Entropy, measuring uncertainty and randomness, is crucial for identifying patterns or irregularities within the signal. This selection ensures a robust statistical toolkit for capturing unique statistical fingerprints.

For a signal with N samples, these statistical features are computed as follows:

Kurtosis of magnitude/angle:

$$\text{Kurtosis\_Mag} = \frac{\frac{1}{N}\sum_{i=1}^{N}(M_i - \bar{M})^4}{(\frac{1}{N}\sum_{i=1}^{N}(M_i - \bar{M})^2)^2} \tag{9}$$

$$\text{Kurtosis\_Ang} = \frac{\frac{1}{N}\sum_{i=1}^{N}(A_i - \bar{A})^4}{(\frac{1}{N}\sum_{i=1}^{N}(A_i - \bar{A})^2)^2} \tag{10}$$

Skewness of magnitude/angle:

$$\text{Skewness\_Mag} = \frac{\frac{1}{N}\sum_{i=1}^{N}(M_i - \bar{M})^3}{(\frac{1}{N}\sum_{i=1}^{N}(M_i - \bar{M})^2)^{3/2}} \tag{11}$$

$$\text{Skewness\_Ang} = \frac{\frac{1}{N}\sum_{i=1}^{N}(A_i - \bar{A})^3}{(\frac{1}{N}\sum_{i=1}^{N}(A_i - \bar{A})^2)^{3/2}} \tag{12}$$

Entropy:

$$\text{Entropy} = -\sum_{i=1}^{N} P_i \log_2(P_i) \tag{13}$$

where M_i represents magnitude samples, A_i represents angle samples, and P_i represents the probability of each sample. These statistical measures are particularly effective at capturing subtle variations in signal distributions that arise from hardware-specific characteristics.

3.4. Fundamental characteristics

Fundamental characteristics, including magnitude, phase, and power, constitute the cornerstone of the feature set. Magnitude represents the instantaneous amplitude or strength of the RF signal, phase captures the signal's instantaneous phase, and power characterizes the instantaneous power. These features provide a foundation for understanding the signal's fundamental attributes, offering key insights into amplitude, phase relationships, and power dynamics.

For each I/Q sample, these fundamental characteristics are computed as:

Magnitude:

$$\text{Magnitude} = \sqrt{I_i^2 + Q_i^2} \tag{14}$$

Table 1
Summary of RF fingerprinting features.

Features	Description	Formula[a]
Mean_I	Mean of in-phase components	$\frac{1}{N}\sum_{i=1}^{N} I_i$
Mean_Q	Mean of quadrature components	$\frac{1}{N}\sum_{i=1}^{N} Q_i$
Std_Q	Standard deviation of in-phase components	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i-\bar{I})^2}$
Std_Q	Standard deviation of quadrature components	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(Q_i-\bar{Q})^2}$
Freq_Centroid	Center of mass of the frequency distribution	$\frac{\sum_{i=1}^{N} f_i \cdot P_i}{\sum_{i=1}^{N} P_i}$
Spectral_Flatness	Ratio of geometric mean to arithmetic mean	$\frac{\sqrt[N]{P_1 \cdot P_2 \cdot \ldots \cdot P_N}}{\frac{1}{N}\sum_{i=1}^{N} P_i}$
Kurtosis_Mag	Tailedness of magnitude distribution	$\frac{\frac{1}{N}\sum_{i=1}^{N}(M_i-\bar{M})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(M_i-\bar{M})^2\right)^2}$
Skewness_Mag	Asymmetry of magnitude distribution	$\frac{\frac{1}{N}\sum_{i=1}^{N}(M_i-\bar{M})^3}{\left(\frac{1}{N}\sum_{i=1}^{N}(M_i-\bar{M})^2\right)^{3/2}}$
Kurtosis_Ang	Tailedness of angle distribution	$\frac{\frac{1}{N}\sum_{i=1}^{N}(A_i-\bar{A})^4}{\left(\frac{1}{N}\sum_{i=1}^{N}(A_i-\bar{A})^2\right)^2}$
Skewness_Ang	Asymmetry of angle distribution	$\frac{\frac{1}{N}\sum_{i=1}^{N}(A_i-\bar{A})^3}{\left(\frac{1}{N}\sum_{i=1}^{N}(A_i-\bar{A})^2\right)^{3/2}}$
Entropy	Entropy of the IQ samples	$-\sum_{i=1}^{N} P_i \log_2(P_i)$
Magnitude	Instantaneous amplitude of the RF signal	$\sqrt{I_i^2+Q_i^2}$
Phase	Instantaneous phase of the RF signal	$\arctan\left(\frac{Q_i}{I_i}\right)$
Power	Instantaneous power of the RF signal	$I_i^2+Q_i^2$

[a] Note: I_i and Q_i represent the in-phase and quadrature components, $\bar{I}$ and $\bar{Q}$ are the mean of in-phase and quadrature components, N is the total number of samples, f_i is the frequency at the ith bin, P_i is the power at the ith bin, M_i is the magnitude at the ith sample, A_i is the phase at the ith sample.

Phase:

$$\text{Phase} = \arctan\left(\frac{Q_i}{I_i}\right) \tag{15}$$

Power:

$$\text{Power} = I_i^2 + Q_i^2 \tag{16}$$

These fundamental features directly reflect the inherent characteristics of the RF hardware components, such as amplifier non-linearities and phase noise, which contribute to device-specific signatures.

It is important to note that the initial list of features serves as a potential candidate pool, subject to subsequent refinement using a recursive feature elimination (RFE) algorithm applied to real-world data, as we will discuss in the next section. This iterative process aims to identify a shortened list of features that significantly impact device selection while maintaining the necessary discriminatory power for effective RF fingerprinting. Table 1 provides a summary of these features along with their brief descriptions and formulas.

Building upon the initial selection of RF features grouped into distinct categories, we introduce an augmentation to our feature generation methodology, which incorporates summary statistic vectors for each feature slice over time. This approach aims to enhance the robustness of our RF fingerprinting model by tracking the temporal distribution characteristics of selected features. For each chosen feature slice F, we calculate a suite of summary statistics vectors, providing a comprehensive representation of its temporal evolution. The summary vector SV_F is constructed as shown in Eq. (4).

Here, $\mu(F)$ represents the mean value, $\sigma^2(F)$ is the variance, $\sigma(F)$ is the standard deviation, $\min(F)$ and $\max(F)$ are the minimum and maximum values, and $P_p(F)$ represents the vector of percentiles of the slice F.

In addition to these measures, including percentiles improves our understanding of the feature's distribution characteristics. The median, especially resistant to extreme values, encapsulates a robust measure of central tendency. This comprehensive suite of summary statistics vectors serves as a powerful tool for encapsulating the temporal nuances of our selected features, contributing to the adaptability and resilience of the model under diverse operational conditions, ultimately improving the reliability and accuracy of RF device classification.

4. Experimental evaluation

4.1. Features selection

4.1.1. Dataset description

The dataset used in our experimentation of this section originates from a testbed configuration described in [18], where the objective was to capture and analyze RF signals in the context of WiFi transmissions (IEEE 802.11 a/g) utilizing BPSK 1/2 modulation. We denote it as ACMWiSec21 Dataset in this paper. The testbed comprises one receiver and five transmitters, each consisting of a HackRF One device equipped with ANT500 antennas running GNU Radio. Their experimental setup established WiFi transmissions between the receiver and individual transmitters using open-source GNU Radio code.

The transmissions were sent at a center frequency of 2.45 GHz, with a bandwidth of 2 MHz and a sampling rate of 2 MHz, capturing the I/Q data. During the data collection period, the receiver stayed in one place while only one transmitter was active. The transmitters were placed about 3 ft from the receiver and remained stationary throughout the data collection process. The indoor experimental setting lasted for two days for data collection. Three transmissions were recorded from each transmitter on the first day, each lasting 30 s. During the time between the two transmissions, the transmitter stayed inactive for 15 s. The data transmission was consistent among all transmitters.

The data collection procedure was repeated on the following day. After equalizing the WiFi frame at the receiver side, the recorded I/Q data becomes the raw I/Q dataset. Each entry in this dataset contains 20 million I/Q samples, offering a comprehensive view of the RF signal properties across various transmitter setups.

4.1.2. Data exploration and features selection

Based on the raw I/Q data recorded for each device in the ACMWiSec21 dataset, we start by applying the first step of our methodology described in Section 3. Specifically, for raw I/Q data from each day, we divided the transmissions into $N = 150{,}000$ equal traces from all three transmissions per device. Each slice contains consecutive 200 I/Q samples, where the in-phase part and quadrature part are kept in their complex representation. After normalization of I/Q samples, we compute the n = 14 features described in Table 1 to form a matrix F of size n x N.

$$F_{\text{Device}_k} = \begin{bmatrix} F_{11} & \cdots & F_{1N} \\ \vdots & \ddots & \vdots \\ F_{n1} & \cdots & F_{nN} \end{bmatrix} \tag{17}$$

Subsequently, we apply the same data processing pipeline to four distinct devices within the dataset. This approach enables the extraction of essential features that encapsulate the temporal and spectral characteristics of the RF signals emitted by each device.

Feature relationships. To visually evaluate the interplay and potential correlations between the extracted features, we use Seaborn's pairplot visualization tool. The resulting pairplot, depicted in Fig. 2, presents a matrix of plots where each of selected feature is plotted against every other feature. The diagonal shows kernel density estimates of each feature's distribution, while the off-diagonal cells contain scatter plots showing relationships between feature pairs. Different colors in the scatter plots represent different device classes, allowing visual identification of feature combinations that provide good class separation. Notable clustering patterns are particularly visible in the relationships involving power, magnitude, and spectral flatness features.

The correlation matrix visualization in Fig. 3 uses a color-coded heatmap where darker red indicates positive correlation (1.0) and darker blue indicates negative correlation (−1.0), with gray representing near-zero correlation. The diagonal elements show perfect correlation (1.0) as expected for self-correlation. Several strong correlations are immediately visible through the intense coloring:

- A very strong positive correlation (0.96) between Kurtosis_Mag and Spectral_Flatness.
- A strong positive correlation (0.9) between Power and Spectral_Flatness.
- A significant negative correlation (−0.84) between Magnitude and Spectral_Flatness.
- Strong positive correlations (0.99) between Kurtosis_Mag and Skewness_Mag.
- Notable negative correlations (−0.74 to −0.76) between Entropy and standard deviation features (std_i and std_q).

These correlations highlight the potential discriminatory power of features related to signal strength, spectral characteristics, and distribution shape. The relationships between these features provide valuable insights for our RF fingerprinting approach, particularly in understanding which feature combinations offer the most discriminative power for device identification. As we progress to feature selection, these insights guide the identification of the most informative features for our RF fingerprinting model.

Feature selection with recursive feature elimination (RFE). In this section, we use Recursive Feature Elimination (RFE) to evaluate and select the most discriminative features for RF fingerprinting. A Linear Support Vector Machine (SVM) classifier is used for feature elimination. RFE iteratively prunes less informative features, allowing us to refine our feature set and enhance the model's ability to capture device-specific patterns. Following the RFE process, the optimal feature subset is identified as shown in Fig. 4, revealing that the five most crucial features for RF fingerprinting are magnitude, phase, power, entropy, and spectral flatness. To validate our findings, we applied PCA 3D visualization, as shown in Fig. 5, to the selected features, and the results show clear clusters, affirming the discriminative capabilities of the chosen feature set. We will use these five features in the next steps of our feature generation method.

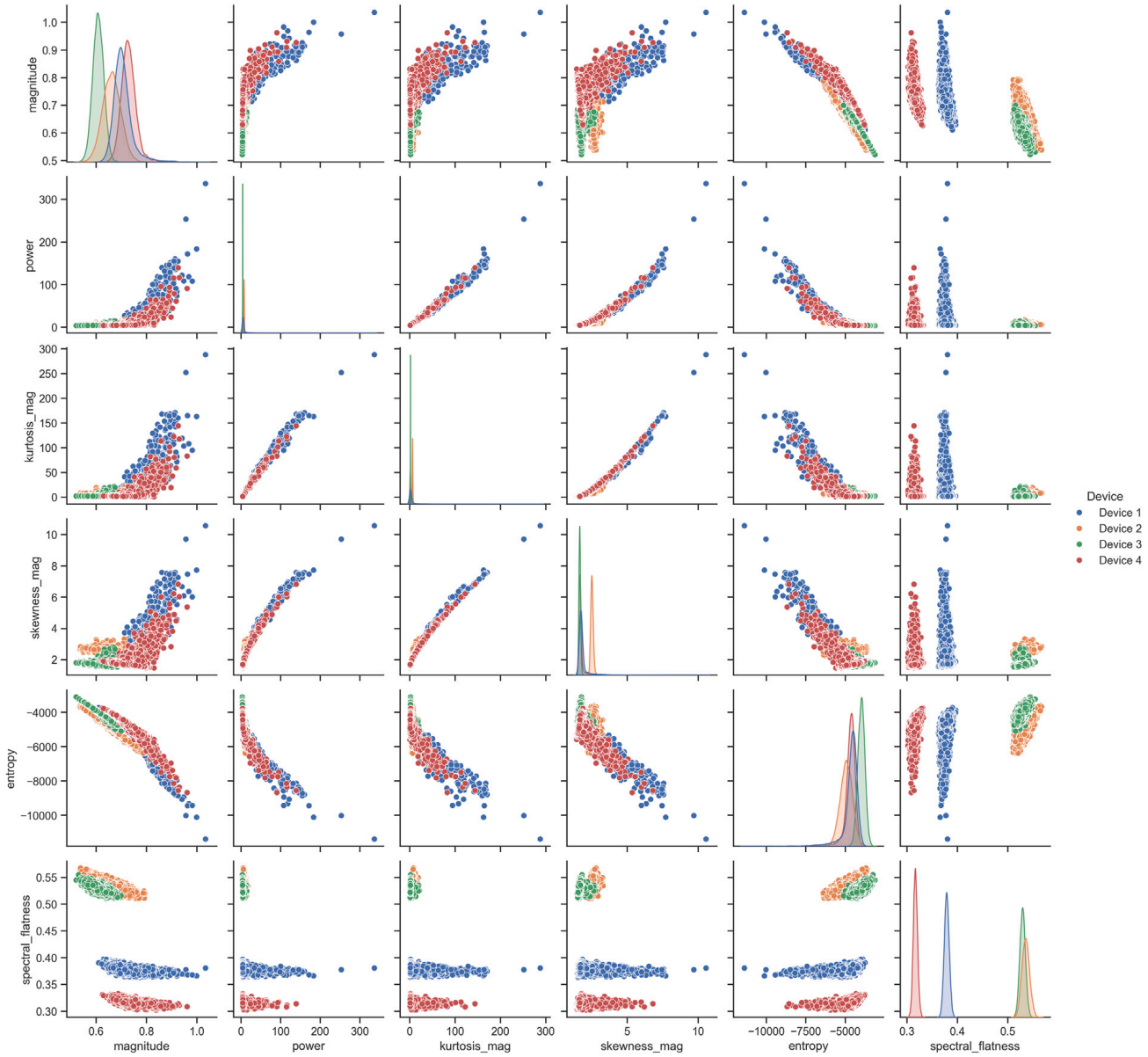


Fig. 2. Seaborn Pairplot showing the relationships between some features.

4.2. Classification with deep learning

4.2.1. DL architecture and settings

To evaluate the effectiveness of our proposed method, we leverage ResNet-50 neural networks. Our primary focus is to validate the enhanced classification performance rather than introducing novel DL architectures; thus, we adopt the ResNet-50-1D model inspired by the well-established ResNet architecture [33], as introduced in [34] for RF fingerprinting.

Our ResNet-50-1D architecture transforms the traditional 2D ResNet structure into a 1D variant optimized for RF fingerprinting. As shown in Fig. 6 and Table 2, the network begins with an input layer accepting feature vectors of length L with n features. This is followed by an initial convolution block (Conv1D with 64 filters, kernel size 7) combined with batch normalization and ReLU activation. After a MaxPool layer, the core of the architecture consists of four sequential stages of residual blocks, progressively increasing in complexity: The first stage contains 2 identity blocks with 256 filters, followed by a second stage with 3 identity blocks using 512 filters. The third stage deepens the network with 5 identity blocks operating with 1024 filters, and the final stage consists of 2 identity blocks with 2048 filters. Each stage is preceded by a convolutional layer (CVL) that handles the dimensional changes. These blocks implement residual connections through the mapping $X_{l+1} = F(X_l, W_l) + X_l$, where X_l represents the input to layer l, W_l the layer weights, and F the residual mapping. This architecture proves particularly effective for RF fingerprinting as it combines the gradient-stabilizing properties of residual connections with the ability to capture temporal patterns in our feature sequences.

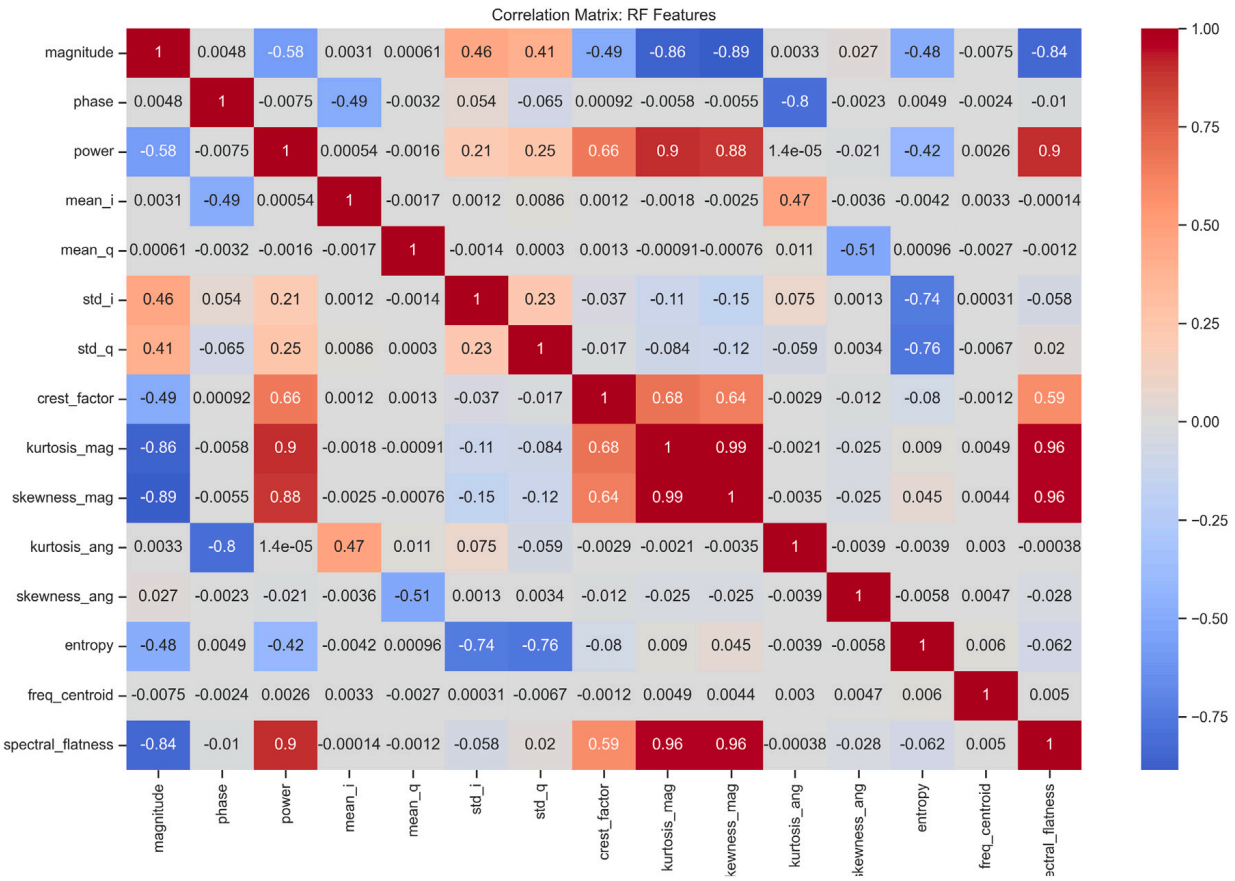


Fig. 3. Correlation matrix.

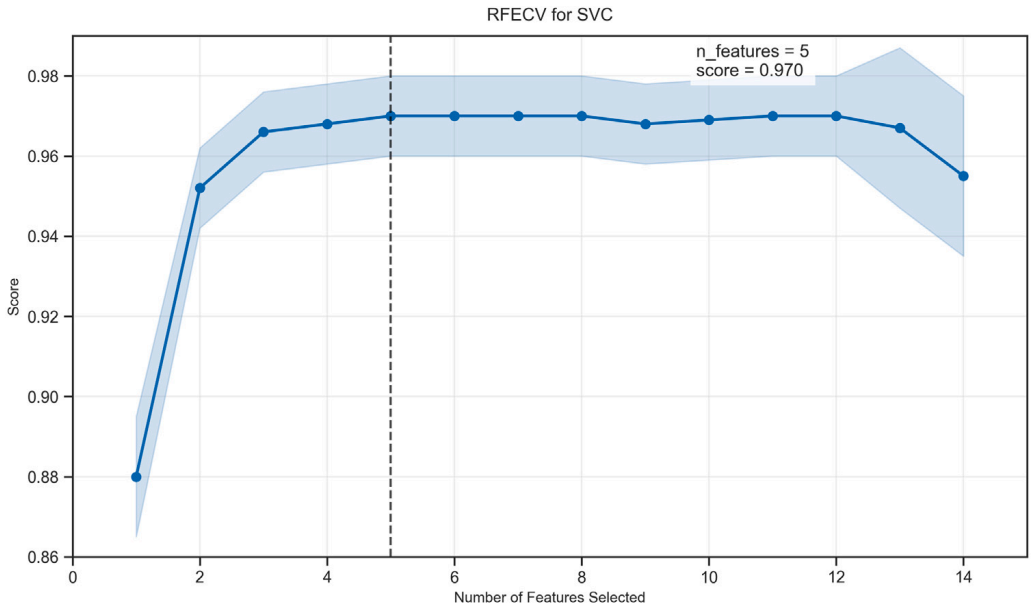


Fig. 4. Recursive feature elimination.

The network concludes with global average pooling and a dense layer with softmax activation, producing device identification probabilities across the number of classes in our dataset.

The implementation of neural networks is conducted in Python, utilizing Keras as the front end and Tensorflow as the back end. Experiments are executed on a Windows machine with Windows 10 Pro, a 5.1 GHz CPU, 32 GB RAM, and an NVIDIA GeForce RTX

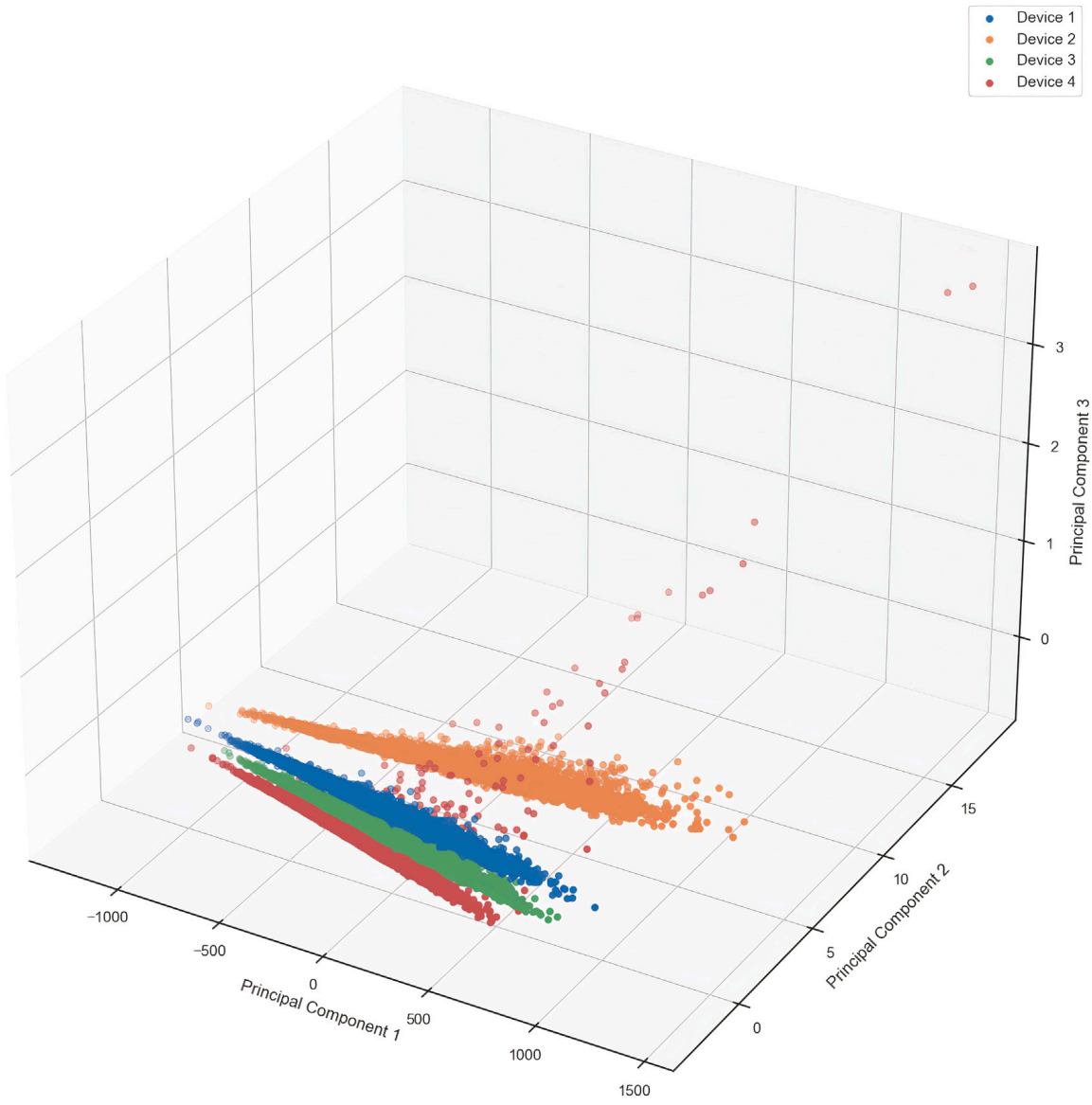


Fig. 5. Principal component analysis.

Table 2
ResNet-50-1D architecture details.

Layer	Output shape	Parameters
Input	(L, n)	–
Conv1D (64, 7) + BN + ReLU	$(L, 64)$	448
MaxPool (2)	$(L/2, 64)$	–
CVL + 2 × ID Block	$(L/2, 256)$	70,144
CVL + 3 × ID Block	$(L/4, 512)$	379,392
CVL + 5 × ID Block	$(L/8, 1024)$	1,226,752
CVL + 2 × ID Block	$(L/16, 2048)$	1,673,216
GlobalAvgPool	(2048)	–
Dense + Softmax	$(num\_classes)$	2048 × num_classes

2070 GPU. Algorithm 1 outlines the details of training and inference for deep attention networks. We allocate signal transmissions for dataset partitioning into training and test sets with a 6:4 ratio. The transmissions in both sets are then fed into the feature generator, as described in Section 3. In all test scenarios, MAX_EPOCH in Algorithm 1 is set to 100, or the training is stopped earlier if the training accuracy does not improve for 10 consecutive epochs.

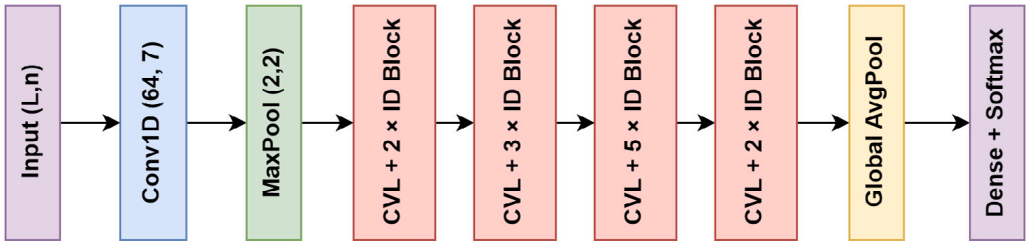


Fig. 6. ResNet-50 architectures.

Algorithm 1 Training and Inferring of DL Model

Training of deep CNN model:
Randomly Initialize network weights
for epoch = 1 to MAX_EPOCHS **do**
  for iteration = 1 to STEPS **do**
    Sample a batch of signal segments Generate statistics vectors batch Input batch into the network and Compute loss Compute gradients Update optimized weights
  Stop training once the model stops learning
Inferring using deep CNN model:
Freeze the networks with learned weights
Get a testing signal
Generate statistics vectors and input them into the network
Predict the fingerprint type of the testing signal using trained networks

4.2.2. Evaluation scenarios

Since our proposed method is tailored for RF fingerprinting models based on CNNs, we establish a benchmark using several State-of-the-Art (SOTA) networks and incorporate two publicly available datasets.

- The first dataset, ACMWiSec21 [18], is previously detailed in this section, wherein the authors evaluate three neural networks: Homegrown, Baseline, and ResNet-50, with a focus on fine-tuning for cross-day scenarios. They used slices of I/Q samples directly as input to the neural networks. We use five devices from this dataset for our tastings and compare results from their ResNet-50 model without fine-tuning.
- Additionally, we utilize a subset of a public dataset introduced in [17], referred to as the GLOBECOM22 Dataset in this paper. The authors used the SOTA attention mechanism to evaluate the generalization capability of RF fingerprinting under different environmental conditions. In this paper, we refer to their model as xDom. In this dataset, the experimental testbed captures WiFi and Bluetooth emissions from IoT devices in an indoor, in-the-wild laboratory setting. The capturing process involved one emitter at a time using a passive receiving USRP X300 radio with a UBX160 daughterboard and VERT2450 antenna. The receiver was tuned to 2.44 GHz with a sampling rate of 66.67 MS/s, deliberately avoiding standard WiFi or Bluetooth bandwidth to maintain generality. The emitters consist of 10 commercial IoT chipsets present in Raspberry Pis and Lenovo laptops, including 2 Lenovo laptops and 8 Raspberry Pi 4Bs. Notably, the dataset was collected over two different time frames, Day 1 and Day 2, with captures separated by several months. We use the 10 available devices for our evaluation.

Our evaluation uses the widely adopted classification accuracy to ensure a fair comparison with existing approaches. This metric signifies the proportion of correctly identified test samples out of the total test samples. It provides a comprehensive evaluation of the effectiveness of our RF fingerprinting methodology across diverse datasets and classification scenarios. We consider three sets of experimental evaluations to assess the generalization capability of our model for real-world deployment across various scenarios:

- **Train Test Same time frame Same Day (TTSS):** In this scenario, both the training and testing sets consist of samples drawn from data captured in the same time frame, location, and testbed setup.
- **Train Test Different time frame Same day (TTDS):** Here, the model is trained with samples collected from a specific time frame on a given day and tested on samples from a different time frame within the same day. This scenario allows us to evaluate the model's robustness to temporal variations on a single day.
- **Train Test Different time frame Different day (TTDD):** Representing the most challenging scenario, the model is trained with samples from a specific time frame on one day and tested on samples from a different time frame on a different day. This scenario introduces significant changes in environmental conditions, providing a tough test of the model's adaptability.

84

4.2.3. Evaluation results

Our method performance. The obtained results are shown in Tables 3 and 4 demonstrate the effectiveness of our proposed RF fingerprinting method across different scenarios and datasets. In the case of the GLOBECOM22 dataset, our method achieves a high

Table 3
Accuracy of our method vs. ResNet-50 (5 devices from GLOBECOM22 dataset).

Scenario	Our method (without Statistic Vectors)	Our method (with Statistic Vectors)	ResNet-50
TTSS	0.987	0.996	0.624
TTDS	0.661	0.721	0.600
TTDD	0.483	0.521	0.3752

Table 4
Accuracy of our method vs. xDom (10 devices from ACMWiSec21 dataset).

Scenario	Our method (without Statistic Vectors)	Our method (with Statistic Vectors)	xDom
TTSS	0.981	0.993	0.991
TTDS	0.826	0.949	0.991
TTDD	0.387	0.613	0.638

classification accuracy of 99.6% in the TTSS scenario, indicating excellent performance when training and testing on samples from the same time frame, location, and testbed setup. Notably, the introduction of the statistical vector further improves performance, particularly evident in the TTDS and TTDD scenarios with accuracy gains of 6% and 3.8%, respectively.

Our method demonstrates impressive accuracy across various scenarios with the ACMWiSec21 dataset, and the inclusion of statistical vectors consistently improves the results. Specifically, the TTDD scenario shows a significant enhancement, with accuracy rising from 38.7% to 54.3%. The results highlight the importance of including statistical summaries of specific RF features, which enhance the model's ability to adjust to different periods and environmental situations.

Our method vs. ResNet-50. As summarized in Table 3, a comparative analysis of our proposed RF fingerprinting method with the approach introduced in [18] reveals notable performance advantages. In the TTSS scenario, our method achieves an excellent identification accuracy of 99.6%, surpassing the ResNet-50 method's accuracy of 62.4%.

In addition, when dealing with the more complex scenarios of TTDS and TTDD, our method consistently surpasses the approach. Especially in the TTDS scenario, our approach has achieved an accuracy of 72.1% compared to the 60% reported in [18], demonstrating a notable enhancement. In the TTDD scenario, our approach achieves an accuracy of 52.1%, surpassing ResNet-50, which achieves an accuracy of 37.5%. These findings once again validate the efficiency of our suggested approach, particularly in situations with different time frames and environmental factors.

Our method vs. xDom. The comparative evaluation between our proposed method and the approach utilized by the authors of [17], reveals interesting insights. As shown in Table 4, in the TTSS scenario, both methods demonstrate high accuracy, with our method achieving 99.3% and xDom attaining 99.1%. This marginal difference suggests comparable performance in scenarios involving samples from the same time frame, location, and testbed setup. Moving to the TTDS scenario, where the model is trained on samples from one-time frame and tested on samples from a different time frame on the same day, our method, with an accuracy of 94.9%, exhibits competitive results against xDom, which achieves 99.1%. In the TTDD scenario, our method achieves an accuracy of 61.3%, while xDom attains a higher accuracy of 63.8%. It is worth noting that the xDom model incorporates an attention mechanism, contributing to its slightly superior performance in scenarios involving temporal variations. However, it is crucial to consider that the attention mechanism introduces higher computational complexity.

4.2.4. Impact of transmissions time

To address the generalization issue, we performed an investigation into the impact of training and testing from different transmission time. The results show some insightful findings, particularly when training/testing the model with transmissions from distinct days.

As shown in Table 5, when utilizing the first transmission (T1) from Day 1 as the training set, the accuracy on subsequent transmissions from the same day remains consistently high, indicating a robust model with acceptable accuracy. However, the scenario changes significantly when the model is tested with transmissions from Day 2. In this case, the accuracy drops considerably for transmissions from Day 2. Surprisingly, the accuracy remains relatively high for transmissions from Day 2 when the model is trained using transmission from the same day. This suggests that the model adapted well to the changing conditions within the same day. This observed discrepancy shows the sensitivity of the RF fingerprints to changes in environmental conditions, leading to a notable decline in accuracy when confronted with transmissions from different days. The results highlight the challenging nature of this problem, indicating that further research is required to enhance the model's ability to generalize across varying conditions. While the model demonstrates proficiency within a consistent timeframe, achieving robustness across different days remains a complex and open research question.

4.2.5. Impact of SNR

85

The Impact of SNR on the accuracy of the RF fingerprinting method is evident from the results obtained across different SNR levels as shown in Fig. 7. As the SNR decreases, there is a remarkable drop in accuracy, emphasizing the susceptibility to channel noise. At higher SNR levels, the accuracy reaches impressive levels, with a peak accuracy of 99.3% at 30 dBm SNR. However, as

Table 5
Accuracy of our method per transmission from GLOBECOM22 dataset.

Transmission	Training: T1 — Day1	Training: T1 — Day2
T1-Day1	0.996	0.256
T2-Day1	0.721	0.420
T3-Day1	0.701	0.313
T1-Day2	0.387	0.980
T2-Day2	0.521	0.872
T3-Day2	0.480	0.780

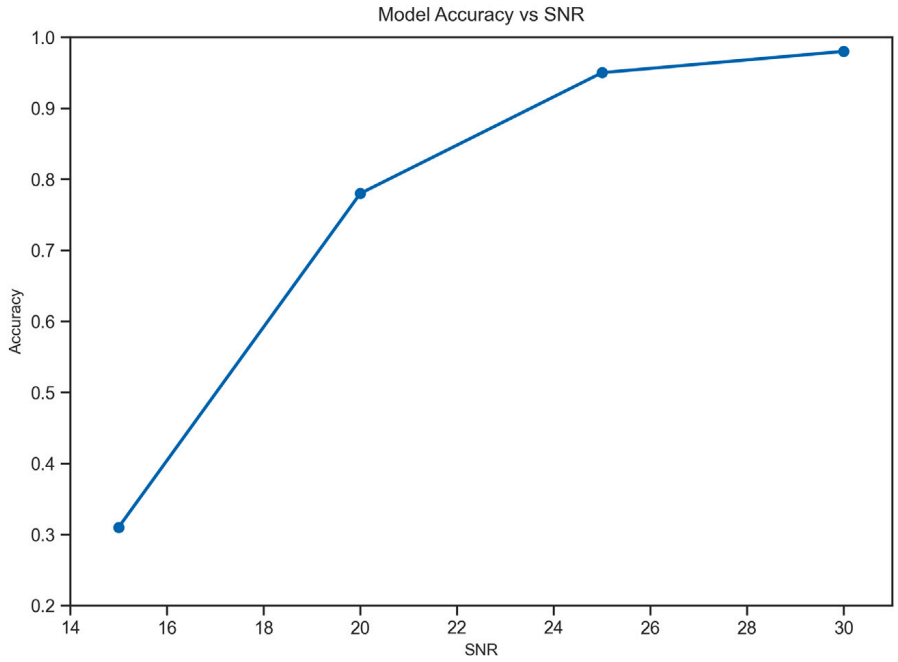


Fig. 7. Accuracy changes vs. SNR.

the SNR decreases to 25 dBm, 20 dBm, and 15 dBm, the accuracy gradually declines to 94.9%, 78.0%, and 31.1%, respectively. This decrease in accuracy underlines the challenge posed by low SNR conditions in real-world scenarios.

To address this, potential solutions involve exploring advanced noise reduction techniques or optimizing feature extraction processes to be less sensitive to noise. Additionally, further research into adaptive algorithms capable of adjusting to varying SNR conditions can contribute to enhancing the robustness of the RF fingerprinting method in diverse and noisy wireless environments.

4.3. Security analysis and robustness considerations

While our RF fingerprinting methodology demonstrates strong performance in device identification, it is essential to consider potential security vulnerabilities and robustness against adversarial attacks. RF fingerprinting systems, particularly those based on deep learning, can face various security challenges including signal replay attacks, spoofing attempts, and adversarial perturbations.

One significant concern in RF fingerprinting systems is the possibility of signal replay attacks, where an adversary captures and retransmits legitimate RF signals. Additionally, sophisticated attackers might attempt feature manipulation attacks targeting specific characteristics used in our feature extraction process, or develop adversarial perturbations designed to fool the deep learning classifier. These attack vectors could potentially compromise the system's ability to accurately identify and authenticate UAV devices.

Our methodology, however, incorporates several inherent defenses against such attacks. The multi-domain feature extraction approach, combining time-domain, frequency-domain, and statistical features, makes it more challenging for attackers to successfully manipulate all necessary signal characteristics. This complexity is further enhanced by our use of statistical summary vectors, which require potential adversaries to maintain consistency across multiple statistical measures. The combination of diverse feature domains and statistical summaries creates a robust framework that inherently resists simple spoofing or manipulation attempts. However, we acknowledge that comprehensive adversarial testing remains an important direction for future work. Further research should focus on systematic evaluation against various types of adversarial attacks, development of specific countermeasures for identified vulnerabilities, and integration of additional security mechanisms such as temporal consistency checks. These evaluations would strengthen the system's robustness and provide valuable insights for deploying RF fingerprinting in security-critical applications.

4.4. Practical applications and operational considerations

The proposed RF fingerprinting methodology has several potential applications across different operational scenarios. For instance, in airport security operations, this approach could enhance UAV detection and authentication systems by providing reliable identification of authorized drones while flagging potential security threats. The high accuracy achieved in same-day scenarios makes the system particularly suitable for continuous monitoring during critical operations, where consistent environmental conditions can be maintained.

In critical infrastructure protection, this methodology could be integrated into existing security frameworks to monitor UAV activities around sensitive facilities. The system's ability to distinguish between authorized and unauthorized devices becomes crucial in these scenarios, where early detection of potentially malicious drones is essential. The multi-domain feature extraction approach provides robust identification even in environments with various RF interference sources, which is common around industrial and infrastructure facilities.

Urban air mobility represents another significant application domain, particularly as cities begin to integrate UAV operations for various services. Our system's capability to handle multiple devices while maintaining high identification accuracy becomes valuable in these dense operational environments. The statistical feature vectors enable reliable device identification even in complex urban environments where signal propagation can be affected by buildings and other structures.

Emergency response scenarios also benefit from our approach, where rapid and accurate UAV identification is crucial. During disaster response or search and rescue operations, the ability to quickly authenticate authorized UAVs while detecting potential interference from unauthorized devices ensures secure and efficient emergency operations. The system's demonstrated performance in real-time processing makes it suitable for these time-critical applications.

5. Future work

This paper's proposed an RF fingerprinting method exhibits promising results, showcasing its potential for real-world applications, particularly in scenarios involving variations within the same day. The method's effectiveness is evident in its ability to achieve high accuracy when trained and tested on transmissions from identical time frames, locations, and testbed setups. However, challenges persist in achieving robust generalization across different days, suggesting several important directions for future research.

A primary focus for future work is improving the model's generalization capabilities. Our ongoing research explores continual learning approaches, incorporating temporal dependencies through hybrid architectures that combine CNNs with LSTM networks. This approach shows promise in addressing the challenge of cross-day performance variations by allowing the model to adapt to evolving signal characteristics while maintaining knowledge of previously learned patterns.

Additionally, we are investigating alternative learning paradigms that frame the authentication problem from different perspectives. One promising direction involves approaching UAV identification as an anomaly detection problem using autoencoder architectures. This approach could potentially offer more robust detection of unauthorized devices by learning the inherent patterns of authorized UAVs rather than relying solely on classification boundaries.

Real-world deployment considerations also present important research directions. Future work should focus on optimizing resource usage and computational efficiency, particularly for edge computing implementations. This includes developing lightweight architecture versions that maintain high accuracy while reducing computational overhead, making the system more suitable for resource-constrained environments.

Scalability remains another crucial aspect for future investigation. As the number of UAVs in operation continues to grow, developed systems must efficiently handle an increasing number of devices while maintaining its identification accuracy. This includes developing adaptive mechanisms for incorporating new device types and maintaining performance as the device population evolves.

The testing and validation of these approaches require more diverse datasets that capture a wider range of operational conditions and device types. Future work should include comprehensive evaluation across different environmental conditions, signal-to-noise ratios, and interference scenarios to ensure robust performance in real-world settings.

Our findings align with recent IoT research trends, particularly regarding the integration of AI/ML with IoT security frameworks. As IoT ecosystems continue to evolve, future research must focus on enhancing the adaptability of RF fingerprinting techniques to accommodate emerging IoT communication protocols and developing standardized frameworks for security evaluation across different UAV platforms and applications.

6. Conclusion

This work explores a new approach to RF fingerprinting for securing UAV communications through statistical features and deep learning techniques. Our investigation demonstrates that utilizing statistical characteristics significantly improves classification accuracy, particularly for same-day scenarios. While we achieve promising results in controlled environments, we identify key challenges in achieving robust cross-day generalization, presenting opportunities for further advancement in RF fingerprinting systems.

The methodology developed in this work provides a foundation for UAV authentication systems, with potential applications across various operational scenarios. As UAV applications continue to expand, our statistical feature-based approach contributes to the broader goal of ensuring secure and reliable UAV operations, while highlighting important directions for future research in this evolving field.

CRediT authorship contribution statement

Nordine Quadar: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Abdellah Chehri:** Writing – review & editing, Visualization, Validation, Supervision, Investigation, Funding acquisition, Formal analysis. **Benoit Debaque:** Writing – review & editing, Visualization, Validation, Resources, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors gratefully acknowledge the financial support provided by Thales Research and Technology, Quebec, Canada, and MITACS (Application Ref/Grant ID: IT40653), a Canadian non-profit organization dedicated to fostering research and innovation through academic, industry, and government collaborations.

Data availability

Data will be made available on request.

References

[1] A. Anitha, L. Arockiam, A review on intrusion detection systems to secure IoT networks, Int. J. Comput. Networks Appl. (2022).
[2] M.N. Adil, H. Song, S. Mastorakis, H. Abulkasim, A. Farouk, Z. Jin, UAV-assisted IoT applications, cybersecurity threats, AI-enabled solutions, open challenges with future research directions, IEEE Trans. Intell. Veh. (2023).
[3] S. Rangarajan, T. Al-Quraishi, Navigating the future of the internet of things: Emerging trends and transformative applications, Babylon. J. Internet Things 2023 (2023) 8–12.
[4] M.A.A. Kabir, W.M. Elmedany, M.S. Sharif, Securing IoT devices against emerging security threats: Challenges and mitigation techniques, J. Cyber Secur. Technol. 7 (2023) 199–223.
[5] A. Jagannath, J. Jagannath, P.S.P.V. Kumar, A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges, Comput. Netw. 219 (2022) 109455.
[6] A. Kumar, R. Saha, M. Conti, G. Kumar, W.J. Buchanan, T. hoon Kim, A comprehensive survey of authentication methods in internet-of-things and its conjunctions, J. Netw. Comput. Appl. 204 (2022) 103414.
[7] Z. Khan, M. Dhinakaran, P. Deepthi, S.S. Johar, R.J.M. Ventayen, R. Kalpana, The recent advancements of radio frequency machine learning (RFML) approaches in enhancing wireless security using multi regression analysis approach (MRAA), 2022 2nd Int. Conf. Adv. Comput. Innov. Technol. Eng. (ICACITE) (2022) 146–150.
[8] S. Riyaz, K. Sankhe, S. Ioannidis, K.R. Chowdhury, Deep learning convolutional neural networks for radio identification, IEEE Commun. Mag. 56 (2018) 146–152.
[9] L.J. Wong, I. WilliamH.Clark, B. Flowers, R.M. Buehrer, A.J. Michaels, W.C. Headley, The RFML ecosystem: A look at the unique challenges of applying deep learning to radio frequency applications, 2020, ArXiv: Signal Processing.
[10] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, T. Melodia, Neural networks for signal intelligence: Theory and practice, Mach. Learn. Futur. Wirel. Commun. (2020) 243–264.
[11] L. Xue, K. Obraczka, Z. Rezki, Cross-layer device fingerprinting and its applications to network security, ICC 2023 - IEEE Int. Conf. Commun. (2023) 5253–5258.
[12] V. Brik, S. Banerjee, M. Gruteser, S. Oh, Wireless device identification with radiometric signatures, in: ACM/IEEE International Conference on Mobile Computing and Networking, 2008.
[13] H. Givehchian, N. Bhaskar, E.R. Herrera, H.R.L. Soto, C. Dameff, D. Bharadia, A. Schulman, Evaluating physical-layer BLE location tracking attacks on mobile devices, 2022 IEEE Symp. Secur. Priv. (SP) (2022) 1690–1704.
[14] C. Xue, T. Li, Y. Li, Y. Ruan, R. Zhang, Radio frequency identification for drones using spectrogram and CNN, in: GLOBECOM 2022 - 2022 IEEE Global Communications Conference, 2022, pp. 4564–4569.
[15] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, K. Chowdhury, ORACLE: Optimized radio classification through convolutional neural networks, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, IEEE, 2019, pp. 370–378.
[16] A. Jagannath, J. Jagannath, Embedding-assisted attentional deep learning for real-world rf fingerprinting of bluetooth, IEEE Trans. Cogn. Commun. Netw. (2023).
[17] A. Jagannath, Z. Kane, J. Jagannath, RF fingerprinting needs attention: Multi-task approach for real-world WiFi and bluetooth, in: GLOBECOM 2022-2022 IEEE Global Communications Conference, IEEE, 2022, pp. 4607–4612.
[18] H. Li, C. Wang, N. Ghose, B. Wang, POSTER: Robust deep-learning-based radio fingerprinting with fine-tuning, in: Proc. ACM WiSec'21, 2021.
[19] O.O. Medaiyese, M. Ezuma, A.P. Lauf, I. Guvenc, Wavelet transform analytics for RF-based UAV detection and identification system using machine learning, Pervasive Mob. Comput. 82 (2022) 101569.
[20] T. Ding, L. Peng, Y. Qiu, Z. Wu, H. Fu, A research of i/q imbalance based RF fingerprint identification with LTE-RACH signals, in: 2021 IEEE 6th International Conference on Signal and Image Processing, ICSIP, IEEE, 2021, pp. 66–71.
[21] G. Baldini, Transient-based radio frequency fingerprinting with adaptive ensemble of transforms and convolutional neural network, Electron. Lett. 59 (22) (2023) e13032.
[22] A. Jagannath, J. Jagannath, P.S.P.V. Kumar, A comprehensive survey on radio frequency (rf) fingerprinting: Traditional approaches, deep learning, and open challenges, Comput. Netw. 219 (2022) 109455.
[23] H.J. Patel, Non-parametric feature generation for RF-fingerprinting on ZigBee devices, 2015 IEEE Symp. Comput. Intell. Secur. Déf. Appl. (CISDA) (2015) 1–5.

[24] L. Peng, A. Hu, J. Zhang, Y. Jiang, J. Yu, Y. Yan, Design of a hybrid RF fingerprint extraction and device classification scheme, IEEE Internet Things J. 6 (2019) 349–360.
[25] S. Deng, Z. Huang, X. Wang, G. Huang, Radio frequency fingerprint extraction based on multidimension permutation entropy, Int. J. Antennas Propag. 2017 (2017) 1–6.
[26] G. Baldini, R. Giuliani, G. Steri, R. Neisse, Physical layer authentication of internet of things wireless devices through permutation and dispersion entropy, 2017 Glob. Internet Things Summit (GIoTS) (2017) 1–6.
[27] Y. Yuan, X. Liu, Z. Liu, Z. Xu, MFMCF: A novel indoor location method combining multiple fingerprints and multiple classifiers, 2019 3rd Int. Symp. Auton. Syst. (ISAS) (2019) 216–221.
[28] M.W. Lukacs, P.J. Collins, M.A. Temple, Classification performance using 'rf-dna' fingerprinting of ultra-wideband noise waveforms, Electron. Lett. 51 (2015) 787–789.
[29] G. Shen, J. Zhang, A. Marshall, L. Peng, X. Wang, Radio frequency fingerprint identification for LoRa using deep learning, IEEE J. Sel. Areas Commun. 39 (2021) 2604–2616.
[30] G.R. Muns, D. Jaisinghani, K. Sankhe, K.R. Chowdhury, Trust in 5G open RANs through machine learning: RF fingerprinting on the POWDER PAWR platform, in: GLOBECOM 2020 - 2020 IEEE Global Communications Conference, 2020, pp. 1–6.
[31] T. Jian, B.C. Rendon, E. Ojuba, N.Y. Soltani, Z. Wang, K. Sankhe, A. Gritsenko, J.G. Dy, K.R. Chowdhury, S. Ioannidis, Deep learning for RF fingerprinting: A massive experimental study, IEEE Internet Things Mag. 3 (2020) 50–57.
[32] J. Yu, A. Hu, G. Li, L. Peng, A robust RF fingerprinting approach using multisampling convolutional neural network, IEEE Internet Things J. 6 (2019) 6786–6799.
[33] A. Al-Shawabka, F. Restuccia, S. D'oro, T. Jian, B.C. Rendon, N.Y. Soltani, J.G. Dy, K.R. Chowdhury, S. Ioannidis, T. Melodia, Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting, IEEE INFOCOM 2020 - IEEE Conf. Comput. Commun. (2020) 646–655.
[34] H. Gu, L. Su, W. Zhang, C. Ran, Attention is needed for RF fingerprinting, IEEE Access 11 (2023) 87316–87329.

4.3 Critical Analysis and Thesis Integration

The statistical feature enhancement methodology we presented in the preceding section represents our foundational approach for addressing RF fingerprinting generalization challenges. In this section, we provide critical analysis of what the approach contributes to the thesis objectives, examine its integration with the experimental framework from Chapter 3, assess computational efficiency for practical deployment, and identify specific limitations that motivate the advanced techniques in subsequent chapters.

4.3.1 Key Contributions to Thesis Objectives

The paper presented in Section 4.2 was published as "Advanced Security Frameworks for UAV and IoT: A Deep Learning Approach" in *Internet of Things* (Elsevier), vol. 32, Art. no. 101594, 2025 [30]. Within the thesis, this work directly addresses **Research Question 2** (RQ2): how can RF fingerprinting systems maintain identification accuracy under temporal evolution of device characteristics caused by aging and environmental drift? The paper's multi-domain statistical feature extraction methodology, cross-day evaluation protocol, and ResNet-50-1D baseline comparison were designed specifically to quantify and address temporal degradation as a standalone contribution. The critical analysis that follows examines how these published findings fit within the broader generalization framework of the thesis, identifies what the paper's evaluation does not cover, and establishes the bridge to the temporal modeling approach presented in Chapter 5.

Our statistical enhancement approach addresses the temporal generalization challenge identified in Chapter 2 through multi-domain feature extraction and temporal aggregation strategies. The experimental validation shows clear performance improvements across multiple evaluation scenarios using the GLOBECOM22 and ACMWiSec21 datasets from Chapter 3.

In same-day scenarios (TTSS), the approach achieves 99.6% accuracy with GLOBECOM22 and 99.3% with ACMWiSec21, confirming strong baseline performance under controlled operational conditions. The statistical vector augmentation provides consistent improvements: the GLOBECOM22 evaluation shows 98.7% accuracy without statistical vectors increasing to 99.6% with their inclusion.

Cross-day generalization performance demonstrates meaningful improvements over traditional approaches while revealing persistent challenges. In the most demanding TTDD scenario, our statistical enhancement achieves 52.1% accuracy compared to 37.5% for the ResNet-50 baseline on GLOBECOM22,

a 14.6 percentage point improvement. The ACMWiSec21 evaluation shows similar patterns with 61.3% accuracy. However, the gap between same-day (99.6%) and cross-day (52.1%) performance tells us that temporal generalization requires more sophisticated approaches than what statistical aggregation alone can provide.

The recursive feature elimination process successfully identified five key features (magnitude, phase, power, entropy, and spectral flatness) as the most discriminative across different operational scenarios. This selection achieves substantial dimensionality reduction from the initial 14-feature set while maintaining superior classification performance, suggesting that integrating domain knowledge can support effective feature engineering for RF fingerprinting.

We note that the multi-domain feature extraction principle establishes an important insight: hardware imperfections manifest differently across signal representation domains. Time-domain features capture power amplifier nonlinearities and I/Q imbalance effects directly, while frequency-domain characteristics reveal oscillator phase noise and spectral variations. Statistical measures, on the other hand, provide robust distribution characterization that maintains stability across varying conditions. This complementary characterization yields more informative fingerprint representations than single-domain approaches, a lesson that informs our subsequent work in later chapters.

4.3.2 Integration with Experimental Framework

The integration of statistical enhancement with the datasets from Chapter 3 provides evaluation capabilities showing both strengths and limitations of the approach. The GLOBECOM22 dataset's multi-protocol structure (Section 3.3.2) enables assessment across WiFi and Bluetooth signals sharing RF circuitry in combo chipsets, which validates the protocol-agnostic nature of our statistical feature extraction.

The combo chipset characteristics in GLOBECOM22 present particular challenges for RF fingerprinting, as shared antenna and RF front-end circuitry create subtle hardware variations that require highly discriminative feature extraction. Our experimental results show that multi-domain characterization successfully addresses these challenges by capturing diverse aspects of hardware imperfections across protocols. The months-long separation between Day 1 and Day 2 collections provides realistic temporal degradation assessment, and the 52.1% cross-day accuracy represents meaningful improvement over baseline approaches despite the persistent performance gap.

The ACMWiSec21 dataset's controlled cross-day evaluation framework (Section 3.3.2) provides assessment of temporal generalization capabilities.

Its two-day collection structure with 100,000 traces per device per day enables thorough statistical validation while maintaining controlled laboratory conditions that minimize environmental confounding factors. The fine-tuning assessment capabilities built into the dataset enable evaluation of adaptation strategies, though the published work focuses primarily on base model performance without fine-tuning.

Our evaluation framework also reveals scalability considerations through the different device populations. The GLOBECOM22 assessment uses 10 devices (8 Raspberry Pi 4B, 2 Lenovo laptops) while ACMWiSec21 employs 5 HackRF One transmitters. Consistent performance across both scales in our experiments suggests that statistical enhancement maintains effectiveness across device population variations, but the limited device diversity means we need additional validation with larger populations to properly assess scalability. The WiSIG dataset's 174 transmitters (Chapter 7) provides this larger-scale validation for subsequent approaches.

4.3.3 Limitations and Motivations for Advanced Techniques

While statistical feature enhancement shows meaningful improvements over traditional approaches, our critical analysis reveals specific limitations that motivate the advanced techniques we present in subsequent chapters. These limitations represent constraints of the statistical aggregation approach that require qualitatively different solutions.

Persistent Cross-Day Performance Degradation: The gap between same-day (99.6%) and cross-day (52.1%) performance indicates that temporal generalization requires more sophisticated modeling than statistical smoothing alone can provide. The summary statistics approach provides temporal robustness through aggregation but lacks the capability to capture long-term dependencies and aging patterns that evolve across extended time periods. This limitation motivates the temporal modeling approaches with CNN-LSTM-Attention architectures we present in Chapters 5 and 6, where recurrent architectures explicitly model temporal evolution and attention mechanisms allow adaptive focus on temporally stable characteristics.

Static Feature Extraction Constraints: The static nature of statistical feature extraction is a limitation when addressing dynamic environments with varying channel conditions and mobility effects. The predefined feature set and fixed aggregation strategy cannot adapt to changing operational conditions, which limits effectiveness in mobile scenarios where Doppler effects, multipath variations, and channel state changes require adaptive feature selection. In Chapter 5, we address this through LSTM-based temporal modeling

that can adaptively weight features based on their temporal stability and discriminative capability under varying channel conditions.

Cross-Transmission Generalization: While statistical enhancement improves cross-day performance, the approach still shows substantial degradation under varying transmission parameters. The static feature extraction lacks adaptive learning capability for efficient incorporation of new transmission conditions or device types without complete retraining. This becomes particularly evident in UAV controller scenarios where systematic transmission parameter variations create distribution shifts that challenge fixed feature extraction. We address this in Chapter 6 through progressive learning approaches that support continual adaptation to new transmission parameters while avoiding catastrophic forgetting.

Closed-Set Classification Assumptions: The classification-based approach assumes closed-set scenarios with known device identities during training, which limits applicability in open-set deployment environments where unknown devices may appear. The softmax classifier outputs confidence scores over known classes but cannot reliably detect device types not represented in the training set. In practical deployments where device populations evolve and security requirements mandate detection of unauthorized devices, this becomes a critical issue. We address this limitation in Chapter 7 through anomaly detection architectures based on reconstruction error rather than classification confidence.

Scalability Assessment: Our evaluation on relatively small device populations (5–10 devices) limits what we can conclude about scalability to larger deployments typical of operational UAV and IoT environments. The consistent performance across GLOBECOM22 (10 devices) and ACMWiSec21 (5 devices) suggests robustness to population variations, but the limited scale prevents assessment of per-device accuracy degradation, confusion patterns among similar devices, and computational scaling with population growth. WiSIG's 174 transmitters supports this larger-scale validation for the Transformer-based architecture in Chapter 7.

These limitations provide clear direction for the subsequent chapters. The temporal modeling in Chapter 5, the integration of statistical robustness with adaptive learning in Chapter 6, and the scalable anomaly detection architecture in Chapter 7 collectively build upon the statistical enhancement foundation while addressing the different aspects of the generalization challenge that this approach alone cannot solve.

4.4 Chapter Summary

In this chapter, we established statistical feature enhancement as our foundational approach for addressing temporal generalization challenges in RF fingerprinting. The multi-domain feature extraction methodology captures complementary signal characteristics across time, frequency, statistical, and fundamental domains, while sliding window temporal analysis and summary statistics aggregation provide robust characterization that maintains discriminative capability under varying operational conditions.

We validated the approach using GLOBECOM22 and ACMWiSec21 datasets across multiple evaluation scenarios. Same-day performance achieves 99.6% accuracy, establishing robust baseline capability, while cross-day scenarios show meaningful improvements over traditional methods (52.1% versus 37.5% for the ResNet-50 baseline). The recursive feature elimination process identifies magnitude, phase, power, entropy, and spectral flatness as the most discriminative features, achieving 64% dimensionality reduction while maintaining superior performance.

Our critical analysis reveals specific limitations that motivate techniques in subsequent chapters. The persistent gap between same-day and cross-day performance (99.6% versus 52.1%) indicates that temporal generalization requires more sophisticated modeling than what statistical aggregation can provide. Static feature extraction limits adaptability to dynamic operational conditions, which motivates the temporal modeling approaches in Chapter 5. Closed-set classification assumptions restrict applicability where unknown devices may appear, a problem we address with anomaly detection architectures in Chapter 7. Cross-transmission performance degradation motivates the continual learning methodologies in Chapter 6.

We emphasize that the statistical enhancement established here provides the essential foundations that enable the generalization framework we develop across Chapters 6 through 7. The multi-domain feature extraction principles inform our temporal modeling architectures, the summary statistics strategies guide continual learning methodologies, and the computational efficiency considerations provide the scalability foundation for large-scale deployments.

Chapter 5

5 Temporal Modeling for Mobile Scenarios

5.1 Introduction

The statistical feature enhancement methodology we presented in Chapter 4 demonstrated substantial improvements over traditional RF fingerprinting approaches, achieving 99.6% accuracy in same-day scenarios and meaningful cross-day performance gains (52.1% vs. 37.5% for conventional methods). However, our critical analysis revealed important limitations when addressing dynamic mobile environments where devices experience temporal variations, channel fluctuations, and mobility-induced effects that static feature extraction cannot adequately capture. We address these limitations in this chapter through the development of the CNN-LSTM-Attention architecture, a hybrid deep learning framework that combines spatial feature extraction with explicit temporal dependency modeling to handle the dynamics introduced by device mobility, Doppler effects, and temporal channel variations [52, 53].

Mobile RF fingerprinting scenarios introduce interconnected challenges that static approaches cannot address [54, 55]. Device mobility generates Doppler frequency shifts that modify signal spectral characteristics, creating time-varying distortions that can mask hardware-specific fingerprint features [56]. Temporal aging effects cause device characteristics to evolve gradually due to component aging, thermal variations, and environmental exposure [19]. Channel variations from multipath propagation, shadowing, and interference patterns that change with device location further complicate fingerprinting [57].

We validate the approach using the LoRa-60 dataset from Chapter 3. The experimental results demonstrate that temporal modeling achieves 99.6% accuracy in stationary conditions while maintaining 85.8% accuracy under high mobility conditions with 100 Hz Doppler shift, which amounts to improve-

ments over the statistical aggregation approach (52.1% cross-day performance from Chapter 4) for dynamic scenarios.

5.2 Theoretical Foundations

Developing effective temporal modeling approaches for mobile RF fingerprinting requires solid theoretical foundations in signal processing, deep learning architectures, and mobile channel modeling. We establish in this section the mathematical framework necessary for understanding how CNN-LSTM-Attention architectures address the temporal dynamics introduced by mobility while maintaining robust device identification.

5.2.1 Temporal Signal Processing for Mobile RF Environments

Mobile RF environments introduce temporal variations that manifest across multiple time scales and require sophisticated modeling to ensure robust fingerprinting performance [58, 59]. The core challenge lies in distinguishing between device-specific temporal patterns that contain fingerprint information and channel-induced variations that represent noise in the identification process [60].

We can model the temporal characteristics of RF signals in mobile scenarios as:

$$s_d(t) = s_{d,static}(t) + s_{d,mobile}(t) + s_{channel}(t) + n(t) \tag{5.1}$$

where $s_d(t)$ represents the received signal from device d, $s_{d,static}(t)$ contains time-invariant hardware fingerprint characteristics, $s_{d,mobile}(t)$ represents device-specific temporal patterns induced by mobility, $s_{channel}(t)$ captures channel-induced temporal variations, and $n(t)$ represents additive noise. The temporal modeling challenge involves learning representations that can separate device-specific components from channel-induced variations across varying mobility conditions.

Effective modeling requires consideration of multiple temporal scales [58]. Short-term variations (milliseconds to seconds) primarily result from Doppler effects and fast fading, while long-term patterns (minutes to hours) reflect device aging and environmental changes. Our temporal modeling architecture must capture dependencies across these diverse time scales to maintain robust performance.

5.2.2 CNN-LSTM-Attention Architecture Principles

The hybrid CNN-LSTM-Attention architecture combines spatial and temporal processing capabilities for RF fingerprinting [61]. Figure 5.1 illustrates the architecture pipeline from raw I/Q samples to device classification.

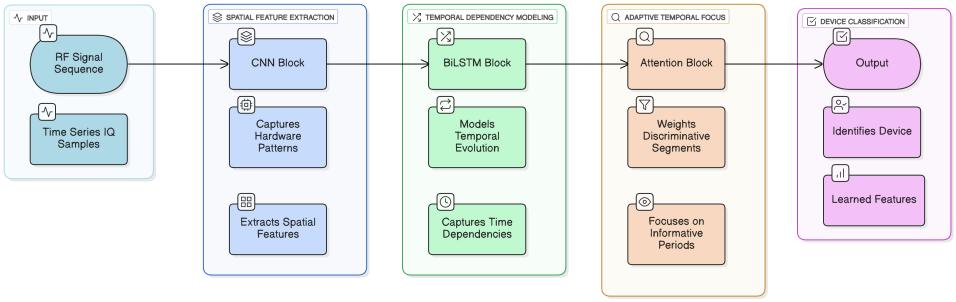


Figure 5.1: Conceptual architecture of the CNN-LSTM-Attention framework for temporal modeling in mobile RF fingerprinting. The pipeline processes raw I/Q samples through four main stages: (1) **Spatial Feature Extraction**: CNN layers capture hardware-specific patterns from RF signals while extracting discriminative spatial features; (2) **Temporal Dependency Modeling**: bidirectional LSTM layers model temporal evolution and capture time dependencies in both forward and backward directions; (3) **Adaptive Temporal Focus**: multi-head attention mechanism weights discriminative temporal segments and focuses on informative periods where hardware fingerprints remain stable despite channel variations; (4) **Device Classification**: fully connected layers identify devices based on the learned temporal-spatial feature representations.

The CNN component performs spatial feature extraction from RF signal representations through convolutional operations:

$$h_{i,j}^{(l)} = \sigma\left(\sum_{m}\sum_{n} w_{m,n}^{(l)} \cdot x_{i+m,j+n}^{(l-1)} + b^{(l)}\right) \tag{5.2}$$

where $h_{i,j}^{(l)}$ represents the activation at position (i, j) in layer l, $w_{m,n}^{(l)}$ are the learnable filter weights, $x^{(l-1)}$ is the input from the previous layer, and σ is the activation function. Hierarchical feature extraction through multiple convolutional layers allows the network to capture increasingly abstract representations of hardware-specific characteristics.

The LSTM component processes these spatial features to model temporal dependencies through bidirectional processing [62]:

$$\overrightarrow{h_t} = \text{LSTM}(\overrightarrow{h_{t-1}}, x_t, \theta_f) \tag{5.3}$$

$$\overleftarrow{h_t} = \text{LSTM}(\overleftarrow{h_{t+1}}, x_t, \theta_b) \tag{5.4}$$

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}] \tag{5.5}$$

where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ represent forward and backward hidden states, x_t is the input at time t, and θ_f and θ_b are the parameters. Bidirectional processing considers both past and future temporal context, which is particularly valuable for handling device-specific patterns that evolve gradually over time.

The attention mechanism dynamically focuses on discriminative temporal segments [63]:

$$e_{t,i} = f_{att}(h_t, h_i) \tag{5.6}$$

$$\alpha_{t,i} = \frac{\exp(e_{t,i})}{\sum_{j=1}^{T} \exp(e_{t,j})} \tag{5.7}$$

$$c_t = \sum_{i=1}^{T} \alpha_{t,i} h_i \tag{5.8}$$

where f_{att} is the attention function, $\alpha_{t,i}$ represents attention weights indicating the relative importance of each temporal position, and c_t is the context vector. The multi-head attention mechanism computes multiple attention patterns in parallel, allowing the model to focus on different aspects of the temporal sequence simultaneously.

5.2.3 Doppler Effect Modeling and Compensation

Doppler effects represent one of the most significant challenges in mobile RF fingerprinting, as they cause frequency shifts that can mask hardware-specific characteristics [56, 64]. The Doppler frequency shift for a mobile transmitter is:

$$f_d = \frac{v \cos(\theta)}{c} f_c \tag{5.9}$$

where v is the relative velocity, θ is the angle between the velocity vector and line of sight, c is the speed of light, and f_c is the carrier frequency. For

LoRa systems at 868 MHz, a 100 Hz Doppler shift corresponds to approximately 125 km/h velocity.

Doppler effects introduce both frequency shifting and spectral spreading that degrade classification performance. We can express the spectral spreading as:

$$B_d = 2f_d = \frac{2vf_c}{c} \tag{5.10}$$

where B_d represents the Doppler spread that causes signal energy to disperse across a wider frequency range.

Our CNN-LSTM-Attention architecture addresses Doppler effects through two complementary strategies: data augmentation during training that exposes the model to synthetic Doppler shifts, and adaptive temporal modeling where the temporal components learn device-specific patterns that remain consistent despite Doppler-induced variations. The attention mechanism focuses on temporal segments least affected by mobility effects [57].

5.2.4 Long-term Temporal Dependency Modeling

Long-term temporal dependencies in RF fingerprinting arise from device aging effects, environmental variations, and gradual operational changes [19, 65]. The LSTM gating mechanisms support modeling of these long-term dependencies:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{5.11}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5.12}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{5.13}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{5.14}$$

where f_t, i_t, and C_t represent the forget gate, input gate, and cell state respectively. The cell state C_t propagates information across many time steps, giving the model awareness of device-specific patterns that evolve gradually.

Bidirectional processing allows the model to consider both past and future temporal context, which is valuable for distinguishing between systematic temporal trends like device aging and transient variations such as temporary environmental effects [62].

5.3 CNN-LSTM-Attention Architecture Framework and Results

The theoretical foundations we established in Section 5.2 provide the basis for developing practical temporal modeling solutions for mobile RF fingerprinting. In this section, we present the complete CNN-LSTM-Attention architecture framework through published research that demonstrates the effectiveness of temporal modeling for LoRa device identification under challenging mobility conditions.

In the published work, we address the core challenge of maintaining robust RF fingerprinting performance when devices operate in dynamic environments with Doppler shifts, multipath propagation, and temporal variations. Our CNN-LSTM-Attention architecture combines the spatial feature extraction capabilities of convolutional neural networks with the temporal dependency modeling of bidirectional LSTM networks and the adaptive focus provided by multi-head attention [31, 66]. This hybrid approach enables the system to learn discriminative features that remain stable across diverse mobility scenarios while adapting to time-varying channel conditions.

We evaluated the architecture across multiple mobility scenarios using 30 commercial LoRa IoT devices from the LoRa-60 collection established in Chapter 3. The results show that temporal modeling provides marked performance improvements over static approaches: 99.6% accuracy in stationary conditions, 93.1% in line-of-sight mobile scenarios, 87.6% in non-line-of-sight mobile scenarios, and 85.8% under high mobility with 100 Hz Doppler shift. These results represent significant advances over conventional RF fingerprinting approaches that typically experience severe degradation under similar mobility conditions.

The following pages present the complete published work, including methodology, experimental setup, results, and detailed analysis of performance across diverse mobility scenarios.

Robust RF Fingerprinting for LoRa IoT Devices in Mobile Scenarios Using CNN-LSTM-Attention

Nordine Quadar
Dept. Math. and Computer Science
Royal Military College of Canada
Kingston, Canada
E-mail: quadar@rmc.ca

Abdelah Chehri
Dept. Math. and Computer Science
Royal Military College of Canada
Kingston, Canada
E-mail: chehri@rmc.ca

Benoit Debaque
Research and Technology Center n
Thales Digital Identity and Security
Quebec, Canada
E-mail: debaque@thalesgroup.com

2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring) | 979-8-3315-3147-8/25/$31.00 ©2025 IEEE | DOI: 10.1109/VTC2025-SPRING65109.2025.11174853

***Abstract*—This paper presents a study of Radio Frequency LoRa device classification performance under challenging channel's variation using a hybrid CNN-LSTM-Attention neural architecture. By addressing the temporal dynamics introduced by Doppler effects in mobile scenarios, combined with targeted data augmentation strategies, our method achieves 99.6% classification accuracy across 10 devices in stationary conditions and maintains robust performance of 85.8% even under high mobility conditions (100 Hz Doppler shift). The proposed hybrid architecture leverages convolutional layers for spatial feature extraction, LSTM layers for modeling temporal dependencies in RF emissions, and an attention mechanism to focus on the most discriminative temporal segments of the signal. Our experimental results, conducted using a dataset of 30 commercial LoRa IoT devices, demonstrate significant performance improvements over state-of-the-art approaches, particularly in challenging mobile environments where Doppler effects typically degrade classification reliability. The model maintains 93.1% accuracy in Line-of-Sight (LOS) mobile scenarios and 87.6% in Non-Line-of-Sight (NLOS) mobile environments, outperforming previous approaches by 3.1% and 2.6% respectively. This work contributes to the field of physical-layer security by demonstrating how temporal modeling techniques can enhance RF fingerprinting performance in realistic mobile deployment scenarios.**

Keywords—RF Fingerprinting; CNN; LSTM, Physical-Layer Security; Doppler.

I. INTRODUCTION

The exponential growth in wireless devices and IoT applications, with projections of approximately 30 billion connected devices by 2027, has intensified the need for reliable device identification and authentication methods in RF networks [1]. This need becomes particularly challenging in mobile environments, where device movement introduces Doppler effects and temporal variations in signal characteristics. Traditional approaches to Radio Frequency (RF) signals classification often focus on protocol-specific features or hardware imperfections extracted from stationary scenarios, but struggle to maintain performance when devices are in motion.

RF fingerprints originate from the inherent defects in the analog circuit components of a transmitter. These defects, similar to human fingerprints, exhibit characteristics such as universality, uniqueness, persistence, and robustness, making them suitable for RF signals identification [2]. The key challenge, however, is maintaining the reliability of these fingerprints when signal characteristics are altered by mobility-induced channel variations. While many existing approaches perform well in controlled, static environments, their accuracy degrades significantly in realistic deployment scenarios where transmitters or receivers are moving.

The challenge of modeling temporal dynamics in RF signals under mobility is particularly relevant to fingerprinting approaches. Raw IQ samples contain rich temporal information that becomes increasingly complex when subject to Doppler effects, requiring models capable of capturing both short and long-term dependencies in the signal. Deep learning techniques have emerged as a powerful tool for processing these complex temporal patterns, with recent studies demonstrating classification accuracies ranging from 80% to 90% using various neural architectures [3], [4], [5], [6]. However, these approaches often employ purely convolutional architectures that lack explicit temporal modeling, limiting their effectiveness in mobile scenarios where the sequential nature of the signal contains critical discriminative information.

The impact of Doppler effects on RF fingerprinting has been under-explored in existing literature. When transmitters are moving relative to receivers, the resulting frequency shifts create temporal dependencies in the signal that static feature extractors struggle to capture. These effects are particularly pronounced in outdoor IoT deployments or vehicular networks where devices may be moving at varying speeds. Recent work has shown that classification accuracy can drop by up to 40% when models trained on stationary data are applied to mobile scenarios [9], highlighting the need for specialized approaches.

This paper presents a methodology for RF signal classification that addresses the temporal dynamics introduced by device mobility. We evaluate the effectiveness of a hybrid CNN-LSTM-Attention architecture that combines the spatial feature extraction capabilities of Convolutional Neural Networks (CNNs) with the temporal modeling power of Long Short-Term Memory (LSTM) and selective focus of attention mechanisms. Our approach leverages data augmentation specifically designed to simulate varying degrees of Doppler shift, enabling robust performance across different mobility conditions. The proposed model demonstrates strong resilience to channel variations, maintaining high classification accuracy even at significant Doppler frequencies representative of real-world mobility scenarios.

The remainder of this paper is organized as follows: Section II provides background on RF fingerprinting techniques and their

challenges under mobility; Section III presents our methodology and neural architecture; Section IV describes the evaluation setup and results discussion; Section V provides conclusions and future work directions.

II. Related Work

A. Overview of RF Signal Classification and Mobile Environments

RF fingerprinting has emerged as a critical technique for device identification and authentication, operating at the physical layer to exploit unique characteristics in device transmissions. As detailed in [2], this technique leverages inherent defects in analog circuit components that create distinctive signal patterns. These hardware-level imperfections manifest as signal distortions that, while having minimal impact on data transmission, provide sufficiently distinctive fingerprints for identification. The approaches used to extract these features can be broadly categorized into traditional feature engineering and deep learning-based methods. Traditional approaches include modulation-based techniques [7], statistical methods [8], and transient-based analysis [9]. While effective in controlled environments, these approaches typically assume stationary conditions and struggle to maintain performance when devices are in motion. The resulting Doppler effects and time-varying channel characteristics introduce temporal dynamics that significantly complicate the fingerprinting process.

B. Deep Learning Approaches for RF Fingerprinting

Deep learning-based RFFI excels at extracting discriminative features from raw signals. CNNs have achieved over 90% accuracyin controlled settings [10] but struggle with modeling temporal dependencies critical for mobility. Recurrent models like LSTMs help capture sequential patterns, improving RF classification in varying channel conditions [11]. Attention mechanisms further enhance these models by focusing on key signal segments, making them effective for RF fingerprinting in dynamic environments.

C. Challenges of Channel Effects and Doppler Shift

The wireless channel significantly impacts RF fingerprinting performance, introducing distortions that can mask or alter the hardware-specific characteristics used for identification. In stationary scenarios, these effects are primarily characterized by multipath propagation and signal attenuation. Mobile scenarios introduce the additional complexity of Doppler shift, which causes time-varying frequency shifts proportional to the relative velocity between transmitter and receiver. Data augmentation has emerged as a key technique for improving robustness to channel variations. In [12] demonstrated that augmentation with synthetic channel models can improve classification accuracy in varying environments. However, most existing augmentation approaches focus on multipath effects and neglect the temporal dynamics introduced by Doppler shift. Recent research has highlighted the need for more sophisticated augmentation strategies that specifically address mobility-induced temporal variations. The evaluation of RF fingerprinting under controlled Doppler conditions remains an underexplored area. While studies like [11] have examined performance in real-world mobile scenarios, systematic analysis of how different levels of Doppler shift affect classification performance is limited. Additionally, comparative evaluation of different neural architectures across varying mobility conditions is needed to guide the development of more robust approaches. This paper addresses these gaps by proposing a hybrid CNN-LSTM-Attention architecture specifically designed to capture both spatial and temporal patterns in RF signals under mobility, combined with a Doppler-aware augmentation strategy to enhance model robustness across varying movement speeds.

III. Methodology

Our approach builds upon a hybrid deep learning framework for RF fingerprinting, specifically designed to address the challenges of device mobility and Doppler effects [13]. The proposed system leverages a CNN network, LSTM units, and attention mechanisms to effectively capture both spatial and temporal patterns in RF signals.

A. Signal Acquisition and Preprocessing

For our experiments, we utilize the LoRa dataset and preprocessing methodology established in [14]. In this approach, the baseband transmitted signal, x(t), undergoes signal modulation and up-conversion via hardware components such as oscillator and power amplifier, which introduce device-specific impairments denoted as f(·). The received signal model is given as:

$$y(t) = h(\tau, t) * f\big(x(t)\big) + n(t) \tag{1}$$

where $h(\tau, t)$ is the time-varying channel impulse response, n(t) is the additive white Gaussian noise, and $*$ denotes convolution operation. The digital I/Q samples $y[n]$ are obtained after analog-to-digital conversion.

Following the established preprocessing pipeline, the raw signals undergo synchronization to locate packet boundaries, carrier frequency offset (CFO) compensation to address frequency drift, and normalization to eliminate power-level dependencies. These preprocessing steps ensure that the classification model focuses on device-specific characteristics rather than channel or power variations.

B. Channel Effects under Mobility

We adopt the channel model framework from [14] which characterizes mobile RF environments through two primary components:

- **Multipath Effect:** Described by the Power Delay Profile (PDP), the multipath component introduces time dispersion. The exponential PDP is selected and the discrete model is given as:

$$\mathrm{P(p)} = \left(\frac{1}{\tau_\mathrm{d}}\right) \mathrm{e}^{-\frac{\mathrm{pTs}}{\tau_\mathrm{d}}},\ \mathrm{p} = 0, 1, \ldots, \mathrm{p}_{max}, \tag{2}$$

where τ_d is the RMS delay spread and p_{max} is the index of the last path.

- **Doppler Effect:** When transmitters or receivers are in motion, the Doppler effect causes frequency shifts proportional to the relative velocity. The Jakes Doppler spectrum model is adopted:

$$S(f) = \frac{1}{\pi f_d \sqrt{1-\left(\frac{f}{f_d}\right)^2}} \quad (3)$$

where f_d is the maximum Doppler shift, which relates directly to mobility speed. For LoRa systems operating at 868 MHz, a Doppler frequency of 10 Hz corresponds to approximately 12.5 km/h, while 100 Hz represents about 125 km/h.

C. Doppler-Aware Data Augmentation

Following the approach in [14], we use the data augmentation provided dataset that simulates realistic channel conditions with particular attention to Doppler effects. The dataset includes samples collected in controlled Line-of-Sight (LOS) stationary conditions, which are then augmented with synthetic channel effects using parameters randomly selected from ranges shown in Table I.

TABLE I. PARAMETERS OF THE CHANNEL SIMULATOR.

Parameter	Range
RMS delay spread τd (ns)	[5,300]
Maximum Doppler frequency fd (Hz)	[0,100]
Rician K-factor	[0,10]
SNR (dB)	[20,80]

The dataset includes variations covering stationary conditions (fd = 0 Hz) to high-mobility scenarios (fd = 100 Hz), enabling systematic evaluation of model performance across different movement speeds.

D. CNN-LSTM-Attention Architecture

The proposed neural architecture, shown in Figure 1 with details in Table II, consists of three main components designed to process the complex temporal dynamics of mobile RF signals:

1. *CNN Backbone:* The convolutional layers extract spatial features from the raw I/Q samples. The network employs convolutional layers with residual connections to enable deeper feature extraction while maintaining gradient flow. The first layer uses 64 7×7 filters with stride 2, followed by deeper layers with 128 3×3 filters.
2. *Bi-directional LSTM Layers:* The temporal features extracted by the CNN are processed through bi-directional LSTM layers with 128 hidden units. These recurrent layers explicitly model the sequential dependencies in the signal, capturing both forward and backward temporal relationships that are critical for identifying device-specific patterns in mobile scenarios.
3. *Multi-head Attention Mechanism:* To focus on the most discriminative temporal segments of the signal, we implement a multi-head attention mechanism with 4 attention heads. This component learns to dynamically weight different time steps based on their relevance for classification, allowing the model to focus on signal portions that remain distinctive despite channel variations.

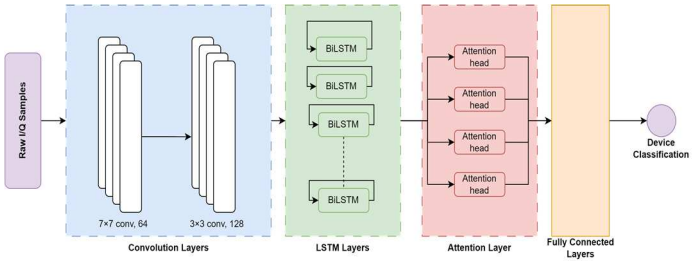


Figure 1. Hybrid CNN-LSTM-Attention architecture overview

The entire model is trained end-to-end using the cross-entropy loss function. To improve generalization, we apply dropout with a rate of 0.5 and batch normalization in all convolutional and dense layers. The output of the network is a softmax layer that produces the probability distribution over the possible device classes.

TABLE II. MODEL CONFIGURATION.

Input channels	Raw I/Q samples (batch_size × sequence_length × 2)
CNN – Layer 1	7×7 conv, 64
CNN – Layer 2	3×3 conv, 128
BiLSTM	128 hidden units
Attention heads	4
Dropout rate	0.5
Normalization	Batch normalization in all convolutional and dense layers

The proposed architecture addresses the challenges of mobile RF fingerprinting through several complementary mechanisms. The CNN component efficiently captures spatial patterns in the signal while the LSTM layers explicitly model the temporal dependencies introduced by mobility. This temporal modeling is crucial for tracking how device-specific impairments manifest across time under Doppler effects. The attention mechanism enhances performance by dynamically focusing on the most discriminative temporal segments, effectively filtering out periods where channel effects may mask the hardware fingerprint. Residual connections ensure effective gradient flow through the deeper layers during training, while the bi-directional processing captures dependencies in both time directions, creating a comprehensive representation of the signal's temporal characteristics.

E. Dataset Description

For our experiments, we utilize the LoRa dataset established in [14], which contains transmissions from 30 commercial off-the-shelf LoRa devices of four different models: 45 LoPy4 devices with SX1276 chipsets, 5 mbed SX1261 shields, 5 FiPy devices with SX1272 chipsets, and 5 Dragino SX1276 shields. The signals were captured using a USRP N210 software-defined radio platform with a sampling rate of 1 MHz at a carrier frequency of 868.1 MHz. The dataset includes multiple test scenarios: stationary with LOS and NLOS conditions, scenarios with objects moving around stationary devices, and fully mobile scenarios where the transmitting devices were carried by a person walking at approximately 2 m/s. For our evaluation of Doppler effects, we focus on the controlled mobility scenarios

with Doppler shifts ranging from 0 Hz (stationary) to 100 Hz (high mobility, equivalent to approximately 125 km/h at 868 MHz).

IV. Evalaution Results

This section presents the experimental evaluation of our proposed CNN-LSTM-Attention architecture across various mobility scenarios. We analyze the model's performance with respect to different Doppler frequencies, comparing against state-of-the-art approaches and evaluating the effectiveness of different architectural components.

A. Experimental Setup

For our evaluation, we conduct experiments using 30 LoRa devices from the dataset described in Section III.E. We train our models using data collected in controlled environments and evaluate performance across multiple test scenarios with varying mobility conditions. All models are implemented using Keras and trained on an NVIDIA GeForce GTX 1660 GPU. To evaluate performance under different mobility conditions, we organize our experiments into three main categories:

1) Stationary scenarios with no Doppler effect (fd = 0 Hz).

2) Low to moderate mobility (fd = 10 Hz, 30 Hz).

3) High mobility scenarios (fd = 50 Hz, 100 Hz).

For each scenario, we measure classification accuracy, precision, recall, and F1-score, with particular attention to how performance degrades as mobility increases. Additionally, we use t-SNE visualizations to examine how feature separability changes under different mobility conditions.

B. Performance Under Varying Doppler Conditions

Table III presents the classification performance of our CNN-LSTM-Attention model across different Doppler frequencies, both with and without data augmentation. As shown in the results, our model maintains good performance in stationary conditions (99.6% accuracy) and demonstrates strong resilience to increasing mobility when trained with Doppler-aware augmentation.

The most significant finding is the model's ability to maintain 85.8% accuracy even under extreme mobility conditions (fd = 100 Hz), representing a substantial improvement over previous approach. Without augmentation, performance drops more rapidly as Doppler frequency increases, highlighting the critical importance of Doppler-aware training for mobile scenarios.

As shown in Table III, classification accuracy remains high even as the number of devices increases, with only a small decrease from 99.6% with 10 devices to 98.3% with 30 devices under stationary conditions. Under mobility (fd = 30 Hz), the performance gap widens slightly, with accuracy dropping from 95.4% (10 devices) to 93.5% (30 devices).

The t-SNE visualizations reveal that while cluster separation remains distinct with 10 devices even under mobility, some overlap begins to appear with 30 devices in high-mobility scenarios. This suggests that as the number of devices increases, the challenge of maintaining discriminative features under mobility becomes more pronounced.

TABLE III. PERFORMANCE OF OUR APPROACH ACROSS DIFFERENT SCENARIOS.

# of Devices	Doppler	Aug[(1)]	Acc[(2)]	Pre[(3)]	Recall	F1
10	No	No	96.80%	96.50%	96.80%	96.60%
10	Yes fd = 10 Hz	No	94.20%	94.00%	93.90%	93.90%
10	Yes fd = 30 Hz	No	90.60%	90.30%	90.20%	90.20%
10	Yes fd = 50 Hz	No	85.30%	85.00%	84.80%	84.90%
10	Yes fd = 100 Hz	No	73.20%	72.80%	72.50%	72.60%
10	No	Yes	99.60%	99.50%	99.60%	99.50%
10	Yes fd = 10 Hz	Yes	97.80%	97.60%	97.80%	97.70%
10	Yes fd = 30 Hz	Yes	95.40%	95.10%	95.20%	95.10%
10	Yes fd = 50 Hz	Yes	92.10%	91.80%	91.70%	91.70%
10	Yes fd = 100 Hz	Yes	85.80%	85.30%	85.50%	85.40%
20	No	Yes	99.10%	99.00%	98.90%	98.90%
30	No	Yes	98.30%	98.20%	98.10%	98.10%
20	Yes fd = 30 Hz	Yes	94.70%	94.30%	94.20%	94.20%
30	Yes fd = 30 Hz	Yes	93.50%	93.20%	93.10%	93.10%

Note on Metrics:
[(1)]: Data augmentation used., [(2)]: Accuracy. [(3)]: Precision.

C. Comparative Performance Analysis

Table IV compares our CNN-LSTM-Attention approach with the state-of-the-art method from [14] across different test scenarios.

Our approach consistently outperforms the reference method across all test scenarios, with the most significant improvements observed in high-mobility conditions. Specifically, our model achieves a 5.8% accuracy improvement at fd = 100 Hz and a 7.1% improvement at fd = 50 Hz. These results highlight the advantage of explicit temporal modeling through the LSTM layers and attention mechanism when dealing with the time-varying characteristics introduced by Doppler effects. The cross-scenario performance, which measures how well models trained in one environment perform in different environments, also shows notable improvements.

Our approach achieves 42.6% cross-scenario accuracy compared to 37.9% for the reference method, indicating better generalization to unseen channel conditions. Beyond controlled Doppler experiments, we evaluated our model on realistic mobile scenarios including LOS and NLOS conditions.

The model maintains strong performance in these challenging real-world conditions, with 93.1% accuracy in LOS mobile scenarios and 87.6% in NLOS mobile scenarios. The performance degradation in NLOS conditions reflects the additional challenges introduced by signal obstructions combined with mobility effects.

TABLE IV. COMPARATIVE PERFORMANCE BETWEEN OUR APPROACH AND [14]

		Our approach		[14]	
Scenario	**Doppler**	**Acc**	**Cross-Scen**	**Acc**	**Cross-Scen**
10 Deviecs	No	99.60%	42.60%	98.50%	37.90%
10 Deviecs	Yes (fd = 10 Hz)	97.80%	41.20%	95.00%	36.30%
10 Deviecs	Yes (fd = 30 Hz)	95.40%	40.50%	91.00%	35.80%
10 Deviecs	Yes (fd = 50 Hz)	92.10%	39.20%	85.00%	34.10%
10 Deviecs	Yes (fd = 100 Hz)	85.80%	38.30%	80.00%	32.70%
20 Deviecs	No	99.10%	41.30%	97.00%	36.50%
30 Deviecs	No	98.30%	40.10%	96.00%	35.20%
Cross-manufacturer	No	91.30%	39.70%	88.70%	34.60%
LOS	Mixed	93.10%	40.80%	90.00%	39.00%
NLOS	Mixed	87.60%	39.40%	85.00%	38.00%

Note on Metrics:
- Accuracy: Percentage of correctly classified samples across all devices
- Cross-Scenario Performance: Accuracy when testing on data collected in environments different from training conditions.
- LOS Mobile: Line-of-Sight scenario where devices are moving but maintain direct signal path to receiver.
- NLOS Mobile: Non-Line-of-Sight scenario where obstacles block direct path while devices are moving.
- Doppler frequency (fd): Measure of frequency shift due to relative motion, with higher values representing faster movement.
- Cross-manufacturer: Performance when testing on devices from manufacturers different from those used in training, evaluating the model's ability to generalize across hardware variations.

D. Cross-Scenario Challenges and Limitations

While our CNN-LSTM-Attention model shows good improvements over existing approaches, cross-scenario performance remains a fundamental challenge in RF fingerprinting. Cross-scenario performance refers to a model's ability to maintain accuracy when tested in environments or conditions different from those encountered during training. As shown in Table IV, even our enhanced architecture achieves only 42.6% accuracy in cross-scenario evaluation, despite reaching 99.6% in matched conditions.

This substantial performance gap highlights a persistent challenge in the field: RF fingerprints that appear distinctive under one set of channel conditions may become less discriminative under different conditions. Although our temporal modeling approach mitigates this issue compared to purely spatial models (improving cross-scenario accuracy by 4.7% over the reference method), the problem is not fully solved. The degradation is particularly pronounced when moving from controlled indoor environments to outdoor scenarios with different multipath characteristics, or when device orientations change significantly.

V. CONCLUSION

In this paper, we investigated the challenges of RF fingerprinting under mobility and proposed a hybrid CNN-LSTM-Attention architecture specifically designed to address the temporal dynamics introduced by Doppler effects. Our evaluation using 30 commercial LoRa devices demonstrated that explicit temporal modeling significantly enhances classification robustness in mobile scenarios. Future work could explore adaptive techniques to further address cross-scenario challenges, more comprehensive channel modeling approaches, and optimization for deployment on resource-constrained devices.

REFERENCES

[1] F. Bruegge et al., “State of IoT—Spring 2023,” 2023.

[2] A. Jagannath, J. Jagannath, and P. S. P. V. Kumar, “A Comprehensive Survey on Radio Frequency (RF) Fingerprinting: Traditional Approaches, Deep Learning, and Open Challenges,” Comput. Networks, vol. 219, p. 109455, 2022.

[3] T. Jian et al., “Deep learning for RF fingerprinting: A massive experimental study,” IEEE Internet of Things Magazine, vol. 3, no. 1, pp. 50–57, 2020.

[4] G. Shen, J. Zhang, A. Marshall, L. Peng, and X. Wang, “Radio frequency fingerprint identification for LoRa using deep learning,” IEEE Journal on Selected Areas in Communications, vol. 39, no. 8, pp. 2604–2616, 2021.

[5] F. Zhuo, Y. Huang, and J. Chen, “Radio Frequency Fingerprint Extraction of Radio Emitter Based on I/Q Imbalance,” Procedia Comput Sci, vol. 107, pp. 472–477, 2017.

[6] H. Gu, L. Su, W. Zhang, and C. Ran, “Attention is needed for RF fingerprinting,” IEEE Access, 2023.

[7] V. Brik, S. Banerjee, M. Gruteser, and S. Oh, “PARADIS : Physical 802. 11 Device Identification with Radiometric Signatures,” 2008.

[8] “Anova-Based RF DNA Analysis - Identifying Significant Parameters for Device Classification,” in Proceedings of the International Conference on Wireless Information Networks and Systems, SciTePress - Science and and Technology Publications, 2010, pp. 47–52. doi: 10.5220/0002994100470052.

[9] Y. Yuan, Z. Huang, H. Wu, and X. Wang, “Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features,” IET Commun., vol. 8, pp. 2404–2412, 2014.

[10] A. Jagannath and J. Jagannath, “Embedding-assisted attentional deep learning for real-world RF fingerprinting of Bluetooth,” IEEE Trans Cogn Commun Netw, 2023.

[11] X. Zeng, D. Liu, C. Yu and D. Lin, "RF Fingerprint Recognition Method Based on Lightweight Deep Learning for Doppler Resilience," 2023 20th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2023, pp. 1-6, doi: 10.1109/ICCWAMTIP60502.2023.10387115.

[12] N. Quadar, A. Chehri, and B. Debaque, “Wireless Security and IoT Device Identification using RF Fingerprinting and Deep Learning,” in 2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall), IEEE, Oct. 2024, pp. 1–5. doi: 10.1109/VTC2024-Fall63153.2024.10757486.

[13] H. Park, S. Kim, S. Min Ko, and T. Kim, “CNN-Based RF Fingerprinting Method for Securing Passive Keyless Entry and Start System,” Computers, Materials & Continua, vol. 76, no. 2, pp. 1891–1909, 2023, doi: 10.32604/cmc.2023.039464.

[14] G. Shen, J. Zhang, A. Marshall and J. R. Cavallaro, "Towards Scalable and Channel-Robust Radio Frequency Fingerprint Identification for LoRa," in IEEE Transactions on Information Forensics and Security, vol. 17, pp. 774-787, 2022.

5.4 Critical Analysis and Integration

The CNN-LSTM-Attention architecture we presented in the preceding section represents an important advancement in addressing temporal and mobility challenges in RF fingerprinting. We provide here a critical analysis of the approach's contributions, evaluate its integration with the experimental framework from Chapter 3, and identify specific limitations that motivate the cyclostationary feature engineering and progressive learning approaches in Chapter 6.

5.4.1 Temporal Modeling Contributions and Dataset Integration

The paper presented in Section 5.3 was published as "Robust RF Fingerprinting for LoRa IoT Devices in Mobile Scenarios Using CNN-LSTM-Attention" in *Proc. IEEE 101st Veh. Technol. Conf. (VTC-Spring)*, 2025 [31]. Within the thesis, this work directly addresses **Research Question 3** (RQ3): how can RF fingerprinting handle mobile scenarios with Doppler effects and time-varying channel conditions? The paper's CNN-LSTM-Attention architecture and Doppler augmentation methodology were validated on the LoRa-60 dataset specifically to characterize performance under mobility as a standalone contribution. The critical analysis below examines how these findings integrate with the thesis experimental framework, what limitations they reveal regarding transmission parameter robustness, and how they motivate the cyclostationary approach in Chapter 6.

The CNN-LSTM-Attention architecture shows that explicit temporal modeling provides marked performance improvements over the statistical aggregation approaches from Chapter 4. Our hybrid architecture achieves 99.6% accuracy in stationary conditions while maintaining 85.8% accuracy under high mobility (100 Hz Doppler shift), a 33.7 percentage point improvement over statistical enhancement's cross-day performance (52.1% from Chapter 4) when addressing temporal dynamics introduced by mobility and channel variations.

Our evaluation using the LoRa-60 dataset from Chapter 3 validates performance across diverse mobility scenarios. We observe 93.1% accuracy in LOS mobile scenarios showing effective Doppler compensation, and 87.6% accuracy in NLOS conditions confirming multipath robustness. The progressive degradation with increasing Doppler shift (99.6% $\rightarrow$ 93.1% $\rightarrow$ 85.8%) suggests that temporal modeling can maintain acceptable performance even as mobility-induced challenges intensify. The key architectural components, CNN spatial

feature extraction, bidirectional LSTM temporal modeling (128 hidden units), and multi-head attention (4 heads), work together to enable robust identification under dynamic conditions where statistical approaches experience severe degradation.

5.4.2 Limitations Motivating Cyclostationary Feature Engineering

While temporal modeling effectively addresses mobility-induced challenges, our analysis reveals specific limitations that motivate the advanced feature engineering in Chapter 6.

Perhaps the most critical limitation concerns robustness to transmission parameter variations. Although the CNN-LSTM-Attention architecture handles devices operating at consistent transmission parameters across different mobility scenarios, we have not validated it under the systematic cross-transmission variations characteristic of UAV controller datasets. Practical deployments may encounter both mobility effects and transmission parameter variations simultaneously, creating compound generalization challenges that temporal modeling alone cannot fully address.

Another limitation is that the CNN spatial features, while effective for capturing hardware imperfections under mobility, do not explicitly capture higher-order statistical periodic patterns that cyclostationary analysis can reveal. Standard convolutional operations extract second-order statistical features through learned filter weights, but hardware-specific periodic patterns that manifest as fourth-order and sixth-order statistical relationships remain unexploited. In Chapter 6, we address this by enhancing the CNN-LSTM architecture with cyclostationary feature extraction layers that embed squaring, cubic transformations, and spectral correlation operations directly within the CNN component. These layers capture hardware-specific periodic patterns that remain discriminative across both mobility effects and transmission parameter variations.

We also note that the cross-scenario performance degradation remains significant despite temporal modeling improvements, with accuracy dropping from 99.6% (same-scenario) to 42.6% (cross-scenario), a 57 percentage point drop that suggests that temporal modeling alone does not fully address generalization across diverse operational conditions. This motivates the progressive learning methodology in Chapter 6, which enables efficient incorporation of new operational scenarios through three-stage adaptation, achieving 38% memory reduction compared to complete retraining while maintaining competitive generalization.

A practical concern is that the static architecture requires complete retraining when new devices are added, creating computational overhead and operational disruption. Chapter 6's progressive learning framework addresses this through continual adaptation using selective parameter updating guided by Fisher information importance weighting. This maintains temporal modeling capabilities while enabling device population evolution without requiring complete system redeployment.

5.4.3 Foundation for Enhanced Temporal-Feature Integration

What we demonstrated here with temporal modeling establishes essential foundations that Chapter 6 builds upon by integrating cyclostationary feature extraction with the CNN-LSTM-Attention architecture. The demonstrated effectiveness of bidirectional LSTM layers for capturing device-specific temporal patterns gives us confidence that adding richer feature representations through cyclostationary operations will enhance cross-transmission robustness while maintaining mobility adaptation capabilities.

Our LoRa dataset evaluation reveals that feature extraction quality is what fundamentally limits generalization performance. By embedding cyclostationary operations within the CNN layers, Chapter 6 enables the LSTM to model temporal dependencies of hardware-specific periodic patterns rather than just learned convolutional features. This can support robust performance across both mobility variations (confirmed in this chapter) and transmission parameter changes (confirmed in Chapter 6). The attention mechanism observed ability to focus on discriminative temporal segments informs the device-specific attention heads we develop in Chapter 6, extending temporal attention to enable adaptive feature weighting tailored to individual devices.

5.5 Chapter Summary

In this chapter, we established temporal modeling as an important approach for addressing mobility-induced challenges in RF fingerprinting through the CNN-LSTM-Attention architecture. Building upon the statistical enhancement foundation from Chapter 4, we demonstrated that explicit modeling of temporal dependencies provides marked performance improvements over static feature aggregation when devices operate under dynamic conditions. Our experimental validation using the LoRa-60 dataset shows robust performance across diverse mobility scenarios: 99.6% accuracy in stationary conditions, 93.1% in LOS mobile scenarios, 87.6% in NLOS conditions, and 85.8% under high mobility with 100 Hz Doppler shift. The CNN-LSTM-Attention

architecture successfully combines spatial feature extraction with temporal dependency modeling and adaptive focus through attention mechanisms.

Our critical analysis revealed specific limitations motivating Chapter 6. While temporal modeling effectively addresses mobility-induced challenges, the architecture shows limited robustness to transmission parameter variations and does not explicitly capture higher-order statistical periodic patterns. The cross-scenario performance degradation remains significant (99.6% $\rightarrow$ 42.6%), and adding new devices requires complete retraining. In Chapter 6, we address these limitations by enhancing the CNN-LSTM-Attention architecture with cyclostationary feature extraction layers and progressive learning methodology, achieving 85.9% cross-transmission generalization on UAV controller datasets while maintaining the mobility adaptation we established here.

The integration of temporal modeling (this chapter) with cyclostationary features and progressive learning (Chapter 6) creates a framework that addresses cross-day temporal variations (Chapter 4), mobility effects (this chapter), and cross-transmission robustness (Chapter 6). This progressive development establishes the foundations for the scalable architectures and unknown device detection capabilities in Chapter 7, where we use Transformer-based approaches to address large-scale deployment requirements while building upon the temporal modeling and feature engineering principles validated in this and previous chapters.

Chapter 6

6 Advanced Feature Engineering with Continual Learning

6.1 Introduction

The temporal modeling approaches we presented in Chapter 5 demonstrated substantial improvements in addressing mobility-induced challenges through CNN-LSTM-Attention architectures, achieving 99.6% accuracy in stationary conditions and 85.8% accuracy under high mobility with 100 Hz Doppler shift. However, our critical analysis revealed that while temporal modeling effectively handles mobility effects, it exhibits limited robustness to transmission parameter variations, a limitation that motivates more sophisticated feature engineering strategies. The persistent cross-scenario performance degradation (99.6% same-scenario to 42.6% cross-scenario) and the lack of explicit higher-order statistical feature extraction indicate that temporal modeling alone cannot fully address the compound generalization challenges encountered when transmission parameters vary significantly from training conditions.

In this chapter, we introduce advanced feature engineering approaches that exploit cyclostationary signal processing principles to achieve robust cross-transmission generalization while maintaining computational efficiency suitable for practical deployment. Communication signals exhibit cyclostationarity through periodic statistical properties arising from modulation processes, symbol timing, and carrier frequencies [67, 68]. Unlike conventional spectral analysis that assumes signal stationarity, cyclostationary analysis captures higher-order statistical relationships that remain relatively invariant to channel effects and transmission parameter variations [69, 70], providing enhanced distinguishing power for RF fingerprinting.

The core innovation lies in developing custom neural network layers that

perform operations analogous to cyclostationary analysis, supporting adaptive learning of distinguishing signal characteristics without the computationally intensive preprocessing or accurate parameter estimation that traditional cyclostationary approaches demand. Our cyclostationary feature extraction framework captures higher-order statistical relationships through specialized neural network layers implementing squaring, cubic, and spectral correlation operations, revealing quadratic and cubic relationships invisible in conventional second-order statistics [71, 72].

Progressive learning represents the second major contribution, addressing the critical challenge of adapting RF fingerprinting systems to new devices without catastrophic forgetting of previously learned capabilities [73, 74]. Traditional approaches require complete retraining when new devices join the network, incurring substantial computational costs and potentially degrading performance on existing device classifications. Our progressive learning methodology enables efficient adaptation through a three-stage process that achieves 38% memory reduction (1471MB versus 2390MB) compared to bulk training while maintaining competitive generalization performance.

We validate the approach using the UAV controller datasets from Chapter 3 (Section 3.3.3), specifically the 8-device and 17-device collections that provide controlled cross-transmission evaluation scenarios. In our evaluation, the approach reaches 97.2% validation accuracy and 85.9% cross-transmission generalization on the 8-device dataset, substantially exceeding the cross-scenario performance from Chapter 5 (42.6%). On the larger 17-device dataset, performance maintains 70–80% generalization accuracy, demonstrating scalability despite increased device diversity.

The strategic significance of this chapter lies in complementing the temporal modeling foundation from Chapter 5 by adding robust cross-transmission generalization capabilities that temporal architectures alone cannot provide. While Chapter 5 addresses mobility-induced temporal dynamics through CNN-LSTM-Attention architectures, we enhance those architectures here with cyclostationary feature extraction to address transmission parameter variations. The progressive learning strategies also inform the anomaly detection methodologies in Chapter 7, positioning advanced feature engineering as essential for RF fingerprinting frameworks that address the full spectrum of practical deployment challenges.

Section 6.2 presents the theoretical foundations for cyclostationary feature extraction and progressive learning, providing the mathematical framework that extends beyond what we could include in our submitted work due to space constraints. Section 6.3 presents the complete paper with empirical validation and practical implementation. Section 6.4 provides critical analysis examining

contributions, integration with our experimental framework, and limitations motivating the advanced techniques in Chapter 7. Section 6.5 summarizes key contributions and transitions to scalable architecture approaches.

6.2 Theoretical Foundations and Mathematical Framework

We present in this section the theoretical foundations for cyclostationary feature extraction and progressive learning, which represent core contributions of this thesis. The theoretical framework demonstrates how cyclostationary signal processing principles, traditionally requiring explicit parameter estimation and computationally intensive preprocessing, can be embedded within differentiable neural network architectures to enable adaptive learning without that preprocessing overhead. This contribution addresses a longstanding limitation in RF fingerprinting where robust feature extraction methods prove computationally prohibitive for resource-constrained deployment environments.

The expanded theoretical treatment we provide here goes substantially beyond the mathematical exposition feasible within the space constraints of the submitted work in Section 6.3. While our paper demonstrates empirical validation and provides the implementation details necessary for reproducibility, its condensed theoretical presentation focuses on practical application rather than full mathematical derivations. We include here the complete theoretical framework with detailed derivations of higher-order statistical relationships, rigorous analysis of cyclic cumulant properties, progressive learning formulations with Fisher information-based importance weighting, and theoretical justification for design decisions that are essential for understanding the contributions.

We argue that the theoretical significance lies in three principal innovations. First, we establish mathematically that higher-order statistical transformations, specifically fourth-order (through squaring operations) and sixth-order (through cubic transformations), capture hardware-specific periodic patterns that remain discriminative across transmission parameter variations where conventional second-order statistics fail [75, 71]. Second, embedding cyclostationary operations in neural networks eliminates the preprocessing overhead and parameter estimation requirements of traditional cyclostationary analysis, supporting end-to-end learning through differentiable operations suitable for real-time processing in resource-constrained UAV environments. Third, the progressive learning methodology achieves marked memory efficiency improvements (38% reduction: 1471MB versus 2390MB for the 8-device

evaluation) through selective parameter updating and knowledge preservation based on Fisher information importance weighting.

6.2.1 Cyclostationary Signal Processing Theory

Communication signals inherently exhibit cyclostationarity due to periodic statistical properties arising from modulation processes, symbol timing, and carrier frequency effects [67]. Unlike traditional stationary signal analysis that assumes time-invariant statistical properties, cyclostationary analysis captures temporal variations that occur periodically, providing enhanced discriminative power for RF fingerprinting where hardware-specific characteristics manifest through subtle periodic patterns in transmitted signals.

The theoretical foundation of cyclostationary analysis rests on the Spectral Correlation Function (SCF) and cyclic cumulants that characterize periodic statistical relationships [68]. For a complex-valued signal $x(t)$, we define the spectral correlation function at cycle frequency α and spectral frequency f as:

$$S_x^\alpha(f) = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[X_T\left(f+\frac{\alpha}{2}\right)X_T^*\left(f-\frac{\alpha}{2}\right)\right] \tag{6.1}$$

where $X_T(f)$ represents the finite-time Fourier transform of $x(t)$ over interval T, and α denotes the cycle frequency at which periodic correlations occur. The cycle frequencies correspond to periodicities in the signal generation process, including symbol rates, carrier frequencies, and their harmonics [70].

Higher-order cyclic cumulants extend second-order spectral correlation to capture more complex statistical relationships through moment-based characterizations. The n-th order cyclic cumulant at cycle frequency α provides information about periodic statistical dependencies that are invisible to second-order analysis:

$$C_{n,x}^\alpha(\tau_1,\tau_2,\ldots,\tau_{n-1}) = \text{cum}\{x(t), x(t+\tau_1),\ldots,x(t+\tau_{n-1})\}e^{-j2\pi\alpha t} \tag{6.2}$$

The cumulant representation can reveal hardware-specific periodic patterns that persist across varying transmission parameters, supporting robust device identification even when channel conditions or protocol configurations differ from training scenarios [71, 72].

For RF fingerprinting, the most informative cycle frequencies typically occur at harmonics of the symbol rate $1/T_s$, multiples of the carrier frequency f_c, and combinations thereof. We can express the cycle frequency set for typical PSK/QAM communication signals as:

$$\alpha = (n - 2m)f_c \pm \frac{k}{T_s} \tag{6.3}$$

where n represents the statistical order, m denotes the number of conjugations, and k encompasses non-negative integers constrained by practical estimation requirements [69].

What makes cyclostationary features theoretically advantageous for RF fingerprinting is their robustness to transmission parameter variations and channel effects that typically degrade conventional time-domain and frequency-domain features. While amplitude and phase characteristics may be distorted by channel conditions, the underlying periodic statistical relationships captured by cyclic cumulants tend to preserve hardware-specific signatures across diverse operational scenarios [75].

Traditional cyclostationary analysis requires accurate estimation of signal parameters including symbol rate, carrier frequency offset, and cycle frequency patterns prior to feature extraction. This preprocessing requirement creates computational overhead and introduces estimation errors that can degrade fingerprinting performance, particularly in dynamic environments where parameters may vary rapidly [76]. The neural network-based approach we developed addresses these limitations by embedding cyclostationary operations directly within trainable architectures that can adaptively learn optimal feature extraction without explicit parameter estimation.

6.2.2 Higher-Order Feature Transformations

Higher-order feature transformations rest on the principle that nonlinear operations on complex-valued signals can reveal statistical relationships invisible to linear processing. For RF fingerprinting, the most significant transformations involve squaring, cubic, and higher-order power operations that expose quadratic and higher-order statistical dependencies in hardware-generated signals.

The squaring transformation of a complex signal $z(t) = I(t) + jQ(t)$ produces:

$$z^{(2)}(t) = z(t)^2 = [I(t)^2 - Q(t)^2] + j[2I(t)Q(t)] \tag{6.4}$$

This operation reveals second-order periodic patterns and emphasizes hardware nonlinearities that manifest as quadratic distortion in transmitted signals. The cubic transformation extends this to third-order relationships:

$$z^{(3)}(t) = z(t)^3 = [I(t)^3 - 3I(t)Q(t)^2] + j[3I(t)^2Q(t) - Q(t)^3] \tag{6.5}$$

These higher-order transformations capture power amplifier nonlinearities, oscillator instabilities, and I/Q imbalance effects that create device-specific signatures in the higher-order statistical domain [77, 78]. Transforming to the frequency domain through Fourier analysis reveals cycle frequency patterns corresponding to hardware characteristics.

We note that the theoretical significance of sixth-order features lies in their ability to capture even higher-order nonlinearities while maintaining computational tractability. The power-of-three operation followed by spectral analysis provides access to sixth-order cyclic cumulants that offer enhanced discriminative power for devices with subtle hardware differences:

$$|FFT(z^{(3)}(t))|^2 \propto |C_{6,z}^{\alpha}(f)|^2 \tag{6.6}$$

where the spectral magnitude reveals periodic patterns corresponding to sixth-order statistical dependencies in device hardware [70].

Integrating multiple statistical orders through feature fusion preserves discriminative information across different aspects of hardware imperfections. The concatenation of second-order, fourth-order, and sixth-order features creates representations that capture both subtle and pronounced hardware variations while remaining computationally efficient for real-time processing.

6.2.3 Progressive Learning and Continual Adaptation

Progressive learning addresses the challenge of incorporating new devices into RF fingerprinting systems without degrading performance on previously learned devices, a problem known as catastrophic forgetting in the continual learning literature [73, 79]. The theoretical framework rests on selective parameter updating strategies that preserve important knowledge while enabling adaptation to new conditions.

The Elastic Weight Consolidation (EWC) approach provides the theoretical foundation for preventing catastrophic forgetting through Fisher information-based regularization. For a neural network with parameters θ, the EWC objective function incorporates both new task learning and knowledge preservation:

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{new}}(\theta) + \frac{\lambda}{2}\sum_i F_i(\theta_i - \theta_i^*)^2 \tag{6.7}$$

where $\mathcal{L}_{\text{new}}(\theta)$ represents the loss for new device learning, F_i denotes the Fisher information matrix elements indicating parameter importance, θ_i^* represents optimal parameters for previous devices, and λ controls the trade-off between adaptation and preservation [73].

Progressive neural network architectures provide an alternative approach through lateral connections that enable knowledge transfer while preserving previously learned representations. Each new device introduces additional network capacity with connections to previous device-specific modules, allowing feature sharing without interference [80]. For RF fingerprinting, progressive learning must address unique challenges including device-specific attention mechanisms, hardware characteristic preservation, and efficient memory utilization in resource-constrained environments. We developed a three-stage progressive learning methodology that addresses these requirements:

Stage 1: New Device Feature Learning, we adapt feature extraction layers to capture device-specific characteristics while preserving shared hardware analysis capabilities through selective parameter updates guided by importance weighting.

Stage 2: Attention Mechanism Refinement, we optimize device-specific attention heads to focus on discriminative features for the new device while maintaining attention patterns for existing devices through regularized attention weight updates.

Stage 3: Classification Integration, we integrate new device classification capabilities through expanded output layers and knowledge distillation techniques that preserve classification boundaries for existing devices while accommodating new device signatures.

Memory efficiency of progressive learning becomes critical for practical deployment where computational resources are constrained [81]. Our analysis of memory requirements demonstrates that progressive approaches can achieve substantial reductions compared to bulk retraining while maintaining competitive performance through strategic parameter sharing and selective updating [82, 83].

6.3 DeepRFFinger: Empirical Validation and Implementation

This section presents the complete submitted work that establishes the empirical validation of cyclostationary feature extraction and progressive learning for cross-transmission generalization in RF fingerprinting. Building upon the theoretical foundations in Section 6.2, our paper demonstrates practical im-

plementation through custom neural network layers, experimental evaluation across UAV controller datasets, and systematic performance analysis confirming the theoretical predictions.

We highlight the contributions that translate our theoretical principles into practical implementation. First, the custom feature extraction layers embed cyclostationary operations, including squaring, cubic transformations, and spectral analysis, directly within differentiable neural network architectures, eliminating separate preprocessing while preserving the discriminative advantages of cyclostationary analysis. Second, the three-stage progressive learning methodology enables efficient incorporation of new devices through selective parameter updating guided by Fisher information importance weighting, achieving significant memory efficiency while preventing catastrophic forgetting. Third, device-specific attention mechanisms enable adaptive focus on discriminative features for each device, enhancing cross-transmission robustness through learned importance weighting.

We performed the experimental validation on both 8-device and 17-device UAV controller datasets, providing controlled assessment of cross-transmission generalization while maintaining the experimental rigor necessary for comparative analysis. Our evaluation protocol follows the standardized framework from Chapter 3, with training on Tx1–Tx2 transmission parameters and evaluation across Tx3–Tx19 variations.

The key experimental results demonstrate the effectiveness of our integrated approach. On the 8-device dataset, we achieve 97.2% validation accuracy and 85.9% cross-transmission generalization, which represents clear improvements over existing approaches. Progressive learning maintains 83.8% generalization while achieving 38% memory reduction (1471MB versus 2390MB). Scalability to the 17-device dataset with 70–80% generalization accuracy confirms effectiveness despite increased device diversity. Ablation studies, confusion matrix analyses, and comparative evaluations establish the contribution of individual components while confirming the integrated framework.

The complete experimental protocols, architectural specifications, training procedures, and performance analyses are presented in our paper below.

Integrating Sensing, Communication, and Computing for Robust RF Fingerprinting in Consumer Electronics

Nourdine Quadar, *Member, IEEE,* Abdellah Chehri, *Senior Member, IEEE,* and Benoit Debaque, *Member, IEEE*

***Abstract*—Securing connected consumer electronics is essential as smart homes, wearables, health monitoring systems, and entertainment platforms increasingly rely on seamless integration of sensing, communication, computing, and control. These devices operate in dynamic wireless environments where traditional cryptographic methods alone cannot fully safeguard against cyber and physical threats. To address this challenge, we propose a deep learning–driven radio frequency fingerprinting (RFF) framework that strengthens physical-layer security while remaining lightweight for resource-constrained consumer electronics. Integrating sensing, communication, and computing for robust RF fingerprinting in consumer electronics, the framework combines cyclostationary feature extraction with continual learning and device-aware attention mechanisms to enable resilient authentication across heterogeneous transmission settings. By adapting to invariant signal patterns, the system achieves real-time scalability and robustness, aligning with the demands of consumer electronics that require secure, personalized, and context-aware services. Experimental evaluation demonstrates 97.2% validation accuracy and 85.9% generalization accuracy across transmission variations, outperforming state-of-the-art approaches by 25%. On a larger dataset of 17 devices, the method sustains 70–80% accuracy while reducing memory requirements by 38%, underscoring its suitability for smart wearables, connected health devices, and home automation systems. This work advances the integration of sensing, communication, computing, and control by embedding secure, adaptive device identification into wireless deep learning architectures, paving the way for resilient consumer electronics in real-world environments.**

***Index Terms*—Radio frequency fingerprinting (RFF); Physical-layer security; Wireless device authentication; Deep learning; Cross-transmission generalization; Cyclostationary feature extraction; Continual learning framework**

I. Introduction

The unprecedented proliferation of Internet of Things devices and unmanned aerial vehicles has reshaped the wireless communication landscape, driving pervasive connectivity across industrial automation, smart cities, defense, and critical infrastructure. With billions of heterogeneous IoT nodes and UAV platforms expected to operate concurrently, the expansion of the wireless attack surface has become a pressing concern [1], [5]. These systems are frequently deployed in resource-constrained environments, often relying on lightweight protocols and ad hoc networking, which makes them particularly vulnerable to spoofing, jamming, eavesdropping, and advanced adversarial intrusions. Conventional cryptographic and software-based security mechanisms, while indispensable, face scalability challenges and can be circumvented by sophisticated attacks. As a result, there is a growing consensus in the research community that complementary physical-layer security solutions are required to provide resilient, device-specific authentication and safeguard IoT ecosystems and UAV communications against evolving threats.

Current deep learning approaches for RF fingerprinting typically follow one of two paths. The first directly processes raw in-phase and quadrature (I/Q) samples through convolutional neural networks (CNNs) or recurrent architectures [6]. Although these methods perform well when the test conditions closely match the training conditions, they exhibit weak generalization when transmission parameters change [7]. The second approach employs extensive manual feature engineering prior to classification, which improves generalization but requires domain expertise and computationally intensive pre-processing steps [8], [9].

Our work addresses this fundamental challenge through two complementary innovations. First, we introduce a novel neural network architecture that incorporate custom feature extraction layers based on cyclostationary signal processing, which is a powerful technique in signal processing that captures periodic statistical properties inherent to communication signals [10]. Rather than manually extracting these features beforehand, we design differentiable neural network layers that perform operations analogous to cyclostationary analysis, allowing the network to adaptively learn the most discriminative signal characteristics. This approach maintains the computational efficiency of end-to-end deep learning while incorporating the robust generalization properties of signal processing theory.

Second, we develop a progressive learning methodology based on continual learning principles [11], allowing the system to efficiently adapt to new devices while maintaining performance on those previously encountered ones. Unlike conventional training approaches that require complete retraining when new devices are introduced [12], our progressive framework sequentially adapts to new transmitters through targeted parameter updates and knowledge preservation mechanisms. This approach significantly reduces computational requirements while improving overall system adaptability.

By integrating these approaches, we achieve significant improvements in cross-transmission generalization. Baseline methods, such as Raw I/Q CNN-LSTM, which directly processes I/Q samples using convolutional and recurrent layers,

Nourdine Quadar and Abdellah Chehri are with the Department of Mathematics and Computer Science at the Royal Military College of Canada (RMC), Kingston, Ontario, Canada. E-mail: quadar@rmc.ca; chehri@rmc.ca.

Besnoit Debaque is with Research and Technology Center of Thales Digital Identity and Security, Quebec, Canada. E-mail: debaque@thalesgroup.com

achieve high training accuracy (up to 91.1%) but suffer from poor generalization (67%) due to overfitting to specific transmission parameters [13]. Similarly, features engineering extraction methods improve generalization but require extensive preprocessing, limiting scalability, while traditional retraining approaches demand complete retraining for new devices [14]-[15].

The proposed framework addresses these limitations by embedding cyclostationary feature extraction directly within the neural architecture to enable robust cross-transmission generalization, while employing a continual learning strategy that reduces memory consumption by 38% (1471 MB vs. 2390 MB). Comprehensive evaluations conducted on UAV controller transmission datasets demonstrate a validation accuracy of 97.2% and a cross-transmission generalization accuracy of 85.9% on a set of eight UAV controllers, surpassing state-of-the-art baselines by 25% [16]. When extended to a larger dataset comprising 17 devices, the framework sustains generalization accuracy in the range of 70-80%, thereby confirming scalability in the presence of increased device diversity and hardware similarity. In addition, the progressive learning paradigm significantly reduces memory overhead compared to conventional bulk training, enhancing suitability for resource-constrained environments such as IoT and UAV networks. Collectively, these results establish the framework as a robust and scalable solution for secure device authentication in dynamic wireless environments. Future work will focus on real-world deployment, strengthening resilience against adversarial attacks, and optimizing scalability for very large device population.

The remainder of this paper is organized as follows. Section II reviews background concepts and related work on RF fingerprinting, cyclostationary analysis, and continual learning. Section III details our custom feature extraction layers and their mathematical foundations. Section IV presents our progressive learning methodology for efficient classifier adaptation. Section V describes our experimental setup and evaluation methodology. Section VI presents comprehensive results and comparative analysis. Finally, Sections VII and VIII provide a discussion and concluding remarks, including directions for future research.

II. Background and Related Work

A. RF Fingerprinting Techniques

Radio frequency fingerprinting, previously known as specific-emitter identification, leverages the unique electromagnetic characteristics imparted by hardware imperfections during signal transmission. These imperfections include oscillator instabilities, power amplifier nonlinearities, and in-phase/quadrature imbalances, which create distinctive "fingerprints" that persist across transmissions [17]. RFF has evolved significantly, from traditional methods that depend on handmade features to state-of-art deep learning approaches.

Early techniques used transient signal analysis, extracting features from the on/off behaviors of transmitters. These methods, while effective in controlled environments, often struggled with real-world channel effects and required precisely capture of transient events. Spectral analysis techniques subsequently emerged, focusing on frequency domain representations to identify distinctive transmitter characteristics. However, these approaches typically required extensive preprocessing and feature selection, limiting their adaptability to various signal conditions.

Deep learning has revolutionized RF fingerprinting by allowing end-to-end feature learning directly from I/Q samples [18]. CNNs have demonstrated remarkable classification accuracy, particularly when training and testing conditions align closely [19]. Recurrent architectures, including Long Short-Term Memory (LSTM) networks, have proven effective in capturing temporal dependencies in RF signals. Despite these advances, maintaining performance across varying transmission parameters remains challenging, as neural networks often overfit to specific conditions present in training data.

B. Cyclostationary Signal Analysis

Communication signals inherently exhibit cyclostationarity, statistical properties that vary periodically, arising from modulation, coding, and other transmission processes. Cyclostationary analysis provides powerful tools for characterizing these periodicities, offering several advantages over conventional signal processing techniques. Unlike spectral analysis, which assumes signal stationarity, cyclostationary methods capture higher-order statistical relationships that remain relatively invariant to channel effects.

The theoretical foundation of cyclostationary analysis lies in the spectral correlation function (SCF) and cyclic cumulants (CCs). The SCF measures correlation between frequency components separated by cycle frequencies that are some specific frequencies at which the signal exhibits periodicity. Cyclic cumulants extend this concept to higher statistical orders, providing greater discriminative power for complex signals. These features have proven particularly valuable for modulation classification and signal intelligence applications.

For RF fingerprinting, cyclostationary features offer enhanced robustness against channel variations and changes in transmission parameters. However, traditional implementations require accurate parameter estimation and computationally intensive preprocessing, constraining their application in resource-limited environments. Our work addresses these limitations by embedding cyclostationary analysis directly within the neural network architecture, creating an adaptive feature extraction process that maintains generalization benefits while eliminating separate preprocessing requirements.

C. Continual Learning for RF Classification

The dynamic nature of wireless environments requires adaptive classification systems capable of incorporating new devices without compromising performance on existing ones. Progressive learning, which is a specialized form of continual learning, offers a promising framework to address this challenge, enabling systems to learn sequentially from new data without catastrophic forgetting of previously acquired knowledge.

Conventional approaches to the incorporation of new devices typically require complete retraining, incurring substantial computational costs, and potentially degrading performance on previously encountered classes [20]. Several techniques have been proposed to mitigate these issues, including replay-based methods that retain samples from previous tasks, regularization approaches that restrict weight updates to preserve important parameters, and architectural strategies that allocate new capacity for new tasks [21].

In the RF domain, progressive learning faces unique challenges due to the complex and high-dimensional nature of signal data and the potential for substantial distribution shifts between different devices or transmission parameters. Recent work has explored knowledge distillation and feature alignment techniques to facilitate knowledge transfer between models trained on different devices. However, these approaches often still require substantial computational resources and may not fully preserve performance across the entire device history.

D. Attention Mechanisms for RF Signal Processing

Traditional CNNs and RNNs process all parts of an input sequence with equal importance, potentially diluting discriminative features in RF signals that may appear only in specific time or frequency regions. Attention mechanisms address this limitation by selectively focusing on the most relevant parts of the input. They achieve this by computing importance weights for different positions or features in an input sequence. These weights determine how much influence each part of the input has on the final representation. In RF fingerprinting, this capability is particularly valuable as distinctive device characteristics may manifest only in specific signal segments or frequency bands.

Attention mechanisms typically compute a context vector as a weighted sum of input features:

$$c = \sum_i \alpha_i h_i \tag{1}$$

where h_i is the i-th input feature and α_i is its attention weight, determined by a trainable scoring function.

Recent work has shown that attention mechanisms significantly improve RF fingerprinting performance by helping models focus on the most discriminative signal characteristics. Multi-head attention extends this concept by allowing multiple "views" of the input, each focusing on different aspects of the signal. In the context of RF fingerprinting, device-specific attention heads can learn to identify the unique characteristics of each transmitter, significantly enhancing discrimination capability.

Our approach also leverages device-specific attention mechanisms to identify the most relevant features for each device, enabling more effective cross-transmission generalization while maintaining robust device discrimination.

E. Gaps in Existing Approaches

Despite advances in RF fingerprinting, significant gaps remain in achieving robust cross-transmission generalization and computational efficiency. Deep learning methods that process raw I/Q samples, such as those in [13] and [8], achieve high accuracy (up to 93.2% training accuracy) but struggle with generalization, with accuracies dropping to 64% on unseen transmission parameters. These approaches often overfit to specific transmission conditions, limiting their practical deployment in dynamic environments. Additionally, methods relying on manual feature engineering, such as [9], could improve generalization but require extensive preprocessing, increasing computational costs and reducing adaptability. In the context of continual learning, existing techniques [12], [20] often demand complete retraining for new devices, incurring high memory and time costs. DeepRFFinger addresses these gaps by integrating cyclostationary feature extraction within the neural network, achieving a 25% improvement in generalization accuracy (85.9%) over the best state-of-the-art method (67%), while reducing memory usage by 38% through progressive learning.

III. Custom Feature-Extraction Layers for RF Fingerprinting

RF fingerprinting performance fundamentally depends on the ability to extract distinctive features that remain consistent across different transmission parameters while varying sufficiently between devices. Conventional approaches typically employ either generic neural network architectures that learn feature representations directly from I/Q samples, or manually engineered features extracted through domain-specific signal processing. Our approach bridges these methodologies by integrating cyclostationary signal processing directly into the neural network architecture through custom-designed feature extraction layers.

A. Motivation and Theoretical Foundation

Communication signals exhibit inherent cyclostationary statistical properties that vary periodically as a consequence of modulation schemes, channel coding, and other transmission processes. These cyclostationary characteristics generate distinctive signal fingerprints that remain relatively invariant across diverse channel conditions and transmission parameters, thereby providing a robust foundation for device identification in dynamic wireless environments.

The theoretical underpinning of this work is that hardware-induced imperfections in RF transmitters—such as oscillator drift, power amplifier nonlinearities, and in-phase/quadrature (I/Q) imbalances—manifest as unique cyclostationary features. Conventional methods typically extract these features through cyclic cumulant computations, which demand extensive preprocessing and precise parameter estimation. In contrast, the proposed framework embeds the mathematical operations of cyclostationary analysis directly into the neural network architecture as differentiable layers. This design enables the network to autonomously learn and optimize discriminative cyclostationary representations, eliminating the need for explicit parameter estimation and enhancing adaptability to varying transmission conditions.

This approach offers several advantages:

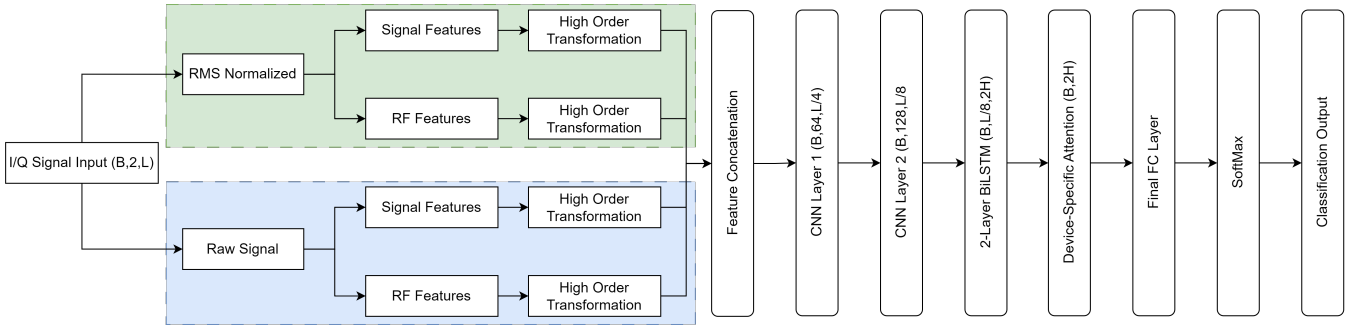


Fig. 1. Overall architecture of the proposed RF fingerprinting system with dual-path approach. The network processes I/Q signal input (B,2,L) where B is batch size, 2 represents I and Q components, and L is the sequence length. The signal flows through parallel RMS Normalized (green) and Raw Signal (blue) paths, extracting both signal and RF hardware features at multiple statistical orders. The feature concatenation layer combines all features into a (B,68,L/2) tensor, followed by two CNN layers reducing spatial dimensions (B,64,L/4 and B,128,L/8 respectively). The 2-Layer BiLSTM processes temporal information producing (B,L/8,2H) where H is the hidden dimension. Finally, device-specific attention heads (B,2H) focuses on discriminative features before classification through FC layer, SoftMax, and final classification output.

- It eliminates the need for separate, computationally intensive pre-processing.
- It allows the network to learn optimal cyclostationary representations directly from the data.
- It preserves the generalization benefits of cyclostationary analysis within an end-to-end learning framework.
- It enables the network to adapt its feature extraction to specific hardware characteristics across devices.

B. Architecture Overview

The proposed architecture, illustrated in Fig. 1, applies cyclostationary analysis principles to both signal properties and RF hardware imperfections for enhanced fingerprinting. While cyclostationary analysis has been successfully applied to modulation classification, our innovation lies in extending this approach to extract unique device fingerprints from subtle hardware variations.

Our architecture processes I/Q signals through two parallel paths: an RMS Normalized path for training stability and a Raw Signal path to preserve relative power differences between devices. Unlike previous approaches that focus primarily on protocol-level characteristics, our framework targets hardware-specific imperfections across multiple RF components:

- **Frontend impairments**: I/Q imbalance, DC offset, and gain mismatch characteristics.
- **Oscillator characteristics**: Phase noise, frequency stability, and jitter patterns.
- **Amplifier signatures**: Nonlinear distortion patterns unique to specific amplifier implementations.
- **Filter characteristics**: Frequency response anomalies and group delay variations.

Each component's features are extracted at multiple statistical orders (2nd, 4th, and 6th) to capture increasingly subtle nonlinear relationships. These features are then processed through specialized CNN branches, fused, and analyzed by a CNN-LSTM backbone with device-specific attention mechanisms to maximize discriminative capability.

C. Custom Layer Design

Our custom layers implement cyclostationary-inspired analysis principles specifically tailored for RF hardware fingerprinting. These differentiable layers extract statistical features at multiple orders that capture both signal characteristics and hardware-specific impairments.

1) Normalization Layers: We implement two parallel normalization approaches to start our feature extraction process:

- **RMS Normalization Layer**: Transforms I/Q signals to have unit power, focusing on shape characteristics while removing amplitude variations:

$$\text{signal} = I + jQ \tag{2}$$

$$\text{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} |\text{signal}_n|^2} \tag{3}$$

$$\text{signal}_{\text{norm}} = \frac{\text{signal}}{\text{RMS}} \tag{4}$$

$$I_{\text{output}} = \text{Re}(\text{signal}_{\text{norm}}), \quad Q_{\text{output}} = \text{Im}(\text{signal}_{\text{norm}}) \tag{5}$$

  where N is the number of samples in the input signal.
- **Raw Signal Path**: Preserves absolute magnitude characteristics that might be distinctive for device identification but could be lost during normalization.

This dual-path approach enables our model to simultaneously capture both shape-based and amplitude-dependent device characteristics.

2) Signal Feature Extractors: For both normalized and raw signal paths, we extract time and frequency domain features to capture diverse signal characteristics:

- **Time Domain Features**: Signal warping extracts temporal patterns while preserving relative magnitude relationships:

$$\text{signal}_{\text{warped}} = \frac{\text{signal}}{\max(|\text{signal}|)^{1/k}} \tag{6}$$

  where k is a scaling factor determined by the feature order (2, 4, or 6).

- **Frequency Domain Features**: FFT and spectral warping reveal spectral patterns unique to each transmitter:

$$X(f) = \text{FFT}(\text{signal}) \tag{7}$$

$$X_{\text{warped}}(f) = \frac{X(f)}{\max(|X(f)|)^{1/2}} \tag{8}$$

3) RF Hardware Feature Extractors: Our specialized extractors target specific RF hardware imperfections that create unique device signatures:

- **Frontend Impairment Extractor**: Captures I/Q imbalance through gain and phase metrics:

$$G_{imb} = \log\left(\frac{\sum I^2}{\sum Q^2}\right) \tag{9}$$

This captures manufacturing variations in I/Q modulators that create consistent device-specific patterns.

- **Oscillator Characteristics Extractor**: Analyzes phase variations that reveal oscillator stability differences:

$$\phi(t) = \arctan\left(\frac{Q(t)}{I(t)}\right) \tag{10}$$

$$\Delta\phi(t) = \phi(t) - \phi(t-1) \tag{11}$$

These phase variations are particularly effective for distinguishing devices with different clock generation circuitry.

- **Amplifier Signature Extractor**: Identifies nonlinear characteristics through envelope analysis:

$$e(t) = \sqrt{I(t)^2 + Q(t)^2} \tag{12}$$

Power amplifier nonlinearities vary significantly between devices and remain consistent across transmission parameters.

- **Filter Characteristics Extractor**: Captures frequency response variations:

$$M(f) = |\text{FFT}(I(t) + jQ(t))| \tag{13}$$

$$M_{\text{norm}}(f) = \frac{M(f)}{max(M(f))} \tag{14}$$

Mnorm(f) normalizes the magnitude spectrum by its peak value, preserving spectral shape while removing absolute amplitude differences between devices, enabling comparison of filter response characteristics across varying transmission power levels.

4) Higher-Order Feature Transformations: A key innovation in our approach is applying cyclostationary-inspired transformations to create higher-order features for both signal and RF hardware characteristics, as illustrated in Fig. 2:

- **Second-Order Features**: Direct features from signal processing and RF hardware extractors serve as our baseline.
- **Fourth-Order Features**: Squaring transformations reveal quadratic relationships invisible in base features:

$$I^{(4)} = (I^{(2)})^2 - (Q^{(2)})^2 \tag{15}$$

$$Q^{(4)} = 2I^{(2)}Q^{(2)} \tag{16}$$

- **Sixth-Order Features**: Power-of-three transformations capture higher-order nonlinearities:

$$I^{(6)} = (I^{(2)})^3 - 3I^{(2)}(Q^{(2)})^2 \tag{17}$$

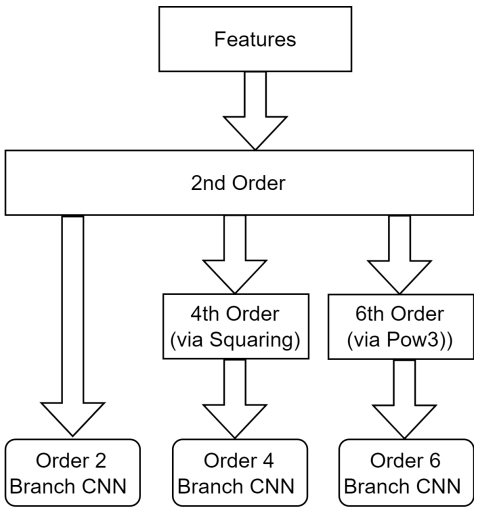


Fig. 2. Feature transformation hierarchy. The 2nd order features are processed into 4th order (via squaring) and 6th order (via pow3) representations before being processed by dedicated branch CNNs.

Algorithm 1 Combined Signal and RF Hardware Feature Extraction with Higher-Order Transformations

Require: I/Q signal x, sequence length L
Extract signal features:
$f_{sig}^{(2)} \leftarrow \text{SignalFeatureExtractor}(x)$ ▷ 2nd order
$f_{sig}^{(4)} \leftarrow \text{SquaringLayer}(f_{sig}^{(2)})$ ▷ 4th order
$f_{sig}^{(6)} \leftarrow \text{Pow3Layer}(f_{sig}^{(2)})$ ▷ 6th order
Add $[f_{sig}^{(2)}, f_{sig}^{(4)}, f_{sig}^{(6)}]$ to f_{sig}
Extract RF hardware features:
for each base feature type t in [Frontend, Oscillator, Amplifier, Filter] **do**
$f_t^{(2)} \leftarrow \text{BaseFeatureExtractor}_t(x)$ ▷ 2nd order
$f_t^{(4)} \leftarrow \text{SquaringLayer}(f_t^{(2)})$ ▷ 4th order
$f_t^{(6)} \leftarrow \text{Pow3Layer}(f_t^{(2)})$ ▷ 6th order
Add $[f_t^{(2)}, f_t^{(4)}, f_t^{(6)}]$ to f_{rf}
end for
Feature concatenation:
$features \leftarrow \text{Concatenate}([f_{sig}, f_{rf}])$
return $features$

$$Q^{(6)} = 3(I^{(2)})^2Q^{(2)} - (Q^{(2)})^3 \tag{18}$$

D. Feature Fusion and Attention-Based Architecture

After extracting multi-order cyclostationary features from RF hardware components, our approach employs feature fusion followed by a CNN-LSTM architecture with device-specific attention mechanisms for classification, as detailed in Algorithm 1.

1) Feature Fusion Strategy: After extracting multi-order cyclostationary features from both signal processing and RF hardware domains, we employ a feature fusion strategy that preserves the discriminative information from each extracted feature while creating a unified representation. The fusion process combines features from different statistical orders across all feature extractors through channel-wise concatenation.

TABLE I
CNN-LSTM-ATTENTION ARCHITECTURE

Component	Layer Type	Output Shape
Input	I/Q Data	$(B, 2, L)$
Feature Extraction	Combined Features	$(B, C_{\text{feat}}, L/2)$
CNN Backbone	Conv1D + BatchNorm + ReLU	$(B, 64, L/4)$
	Conv1D + BatchNorm + ReLU	$(B, 128, L/8)$
Sequence Processing	BiLSTM (2 layers)	$(B, L/8, 2H)$
Attention	Device-Specific Attention heads	$(B, 2H)$
Classification	Fully Connected + ReLU	$(B, 1024)$
	Fully Connected + ReLU	$(B, 256)$
	Device-Specific Classifier	(B, N_{classes})

B: batch size, L: sequence length, C_{feat}: feature channels, H: hidden size, N_{classes}: number of classes

2) CNN-LSTM Processing: The fused features are processed through a hierarchical deep learning architecture that extracts both spatial and temporal patterns, as detailed in Table I:

- **CNN layers:** Extract local patterns and reduce dimensionality while preserving key signal characteristics.
- **LSTM layers:** Model temporal dependencies and long-range patterns that may span across the signal.

This combination is particularly effective for RF signals, where both spectral (frequency domain) and temporal patterns contain complementary discriminative information.

3) Device-Specific Attention Mechanisms: A key innovation in our architecture is the use of device-specific attention heads that learn to focus on the most discriminative features for each device. Each attention head computes a context vector as a weighted sum of the LSTM outputs:

$$c_i = \sum_t \alpha_{i,t} \cdot h_t \tag{19}$$

where c_i is the context vector for device i, h_t is the LSTM output at time step t, and $\alpha_{i,t}$ are the attention weights computed by a trainable scoring function.

This approach is particularly valuable for RF fingerprinting, where different devices may have distinctive characteristics in different signal regions or frequency bands. The device-specific attention mechanisms learn which features are most discriminative for each device, enhancing classification performance especially when devices share similar characteristics in some dimensions.

4) Device-Specific Classification: The device-specific context vectors are passed through a shared fully connected layer followed by device-specific classification layers. During training, only the current device's classifier is updated, while during inference, all device classifiers are evaluated.

This architecture combines the strengths of convolutional, recurrent, and attention-based approaches to create a robust RF fingerprinting framework that can effectively generalize across different transmission parameters while maintaining device-specific discrimination capability.

IV. PROGRESSIVE LEARNING METHODOLOGY FOR RF FINGERPRINTING

RF fingerprinting systems deployed in practical environments must continually adapt to new devices without forgetting previously learned fingerprints [22]- [28]. This requirement presents a fundamental challenge in deep learning: integrating new knowledge while preserving existing capabilities, a challenge known as catastrophic forgetting in continual learning literature [29]. Our progressive learning methodology addresses this challenge by enabling efficient adaptation to new devices while maintaining classification performance across previously encountered transmitters.

A. Motivation and Progressive Learning Framework

Traditional approaches to incorporating new devices typically require complete retraining of the network using all available data. This approach is computationally expensive, time-consuming, and potentially impractical for resource-constrained environments. Our progressive learning framework draws inspiration from continual learning paradigms but is specifically tailored to the unique challenges of RF fingerprinting:

- Signal distributions vary significantly between devices, creating a domain shift problem even among devices of the same type.
- Limited data may be available for new devices compared to existing ones.
- Hardware-specific characteristics that create useful fingerprints may be masked by protocol-specific features or channel conditions.
- Previously learned features remain highly relevant for new device identification.

Our approach differs from standard continual learning in that we don't just add new classification capabilities but actively leverage transfer learning to enhance cross-transmission generalization. Using our progressive learning methodology, we achieve a validation accuracy of 97.2% and a cross-transmission generalization accuracy of 85.9%, a 25% improvement over the best baseline (67%) [13], while reducing memory usage by 38% (1471 MB).

B. Three-Stage Progressive Learning Process

Our progressive learning methodology follows a three-stage process that gradually integrates new devices while preserving performance on existing ones:

1) *Stage 1: New Device Training*: In this initial stage, we train only the device-specific components (attention heads and classification layers) for the new device, while keeping the shared feature extraction layers frozen:

$$\min_{\theta_{\text{new}}} \mathcal{L}_{\text{CE}}(f_{\theta_{\text{new}}}(x_{\text{new}}), y_{\text{new}})$$

where θ_{new} represents the parameters of the device-specific components, and $\mathcal{L}_{\text{CE}}$ is the cross-entropy loss function.

To avoid catastrophic forgetting, we employ a data-balancing strategy where the training batch contains a majority (70%) of samples from the new device and a minority (30%) from previously encountered devices. This approach allows the device-specific components to learn distinctive characteristics of the new device while maintaining awareness of existing devices.

2) *Stage 2: Device Heads Tuning*: Once the new device components have been trained, we fine-tune all attention heads while keeping feature extraction layers frozen. During this stage, we use balanced batches with equal representation from all devices, enabling the attention mechanisms to optimize feature selection across the entire device population. A consistency regularization term helps prevent drastic changes to previously optimized attention patterns.

3) *Stage 3: Final Adaptation*: In the final stage, we perform limited fine-tuning of the entire network, including the feature extraction layers, CNN-LSTM backbone, and attention mechanisms. This step allows the model to adapt to subtle cross-device interactions that may not be captured by device-specific components alone. We incorporate Elastic Weight Consolidation (EWC) to prevent catastrophic forgetting during this stage. EWC estimates the importance of parameters for previously learned tasks using the Fisher information matrix and penalizes significant changes to these important parameters, as shown in Algorithm 2. This fine-tuning improves cross-transmission generalization by refining the shared feature extractors to better capture hardware-specific patterns across all devices, while EWC ensures that performance on previously learned devices remains stable.

Algorithm 2 Progressive Learning for RF Fingerprinting

Require: Initial model M trained on device set D_{init}, new device data x_{new}
Stage 1: New Device Training
Freeze all feature extraction layers in M
Expand model architecture for new device by adding new attention heads
Train only device-specific components:
$\theta_{\text{new}} \leftarrow \arg\min_\theta \mathcal{L}_{\text{CE}}(f_\theta(x_{\text{new}}), y_{\text{new}})$ using
batches with 70% new device data, 30% existing devices
Stage 2: Device Heads Tuning
Keep feature extraction layers frozen
Fine-tune all attention heads:
$\theta_{\text{heads}} \leftarrow \arg\min_\theta \mathcal{L}_{\text{CE}}(f_\theta(x), y) + \lambda_1 \mathcal{L}_{\text{consistency}}(\theta, \theta^{\text{old}})$
using balanced batches across all devices
Stage 3: Final Adaptation
Unfreeze all layers
Apply EWC regularization
$\theta_{\text{all}} \leftarrow \arg\min_\theta \mathcal{L}_{\text{CE}}(f_\theta(x), y) + \lambda_2 \mathcal{L}_{\text{EWC}}(\theta, \theta^*)$
where $\mathcal{L}_{\text{EWC}}(\theta, \theta^*) = \sum_i \frac{1}{2} F_i(\theta_i - \theta_i^*)^2$, F_i is the Fisher information for parameter i, and θ^* are the parameters after Stage 2
return Updated model M' with preserved knowledge from D_{init} and new device capability

C. Progressive Learning Implementation

The implementation of our progressive learning methodology introduces minimal computational overhead compared to standard training approaches, while offering significant practical benefits. The three-stage process is implemented sequentially, with early stopping applied at each stage to prevent overfitting.

For Stage 1, we use a higher learning rate (1×10^{-3}) to quickly adapt the new device components. In Stage 2, we reduce the learning rate to 5×10^{-4} for more controlled adaptation of attention mechanisms. Finally, in Stage 3, we use an even lower learning rate (1×10^{-4}) to fine-tune the entire network without disrupting previously learned features.

Progressive learning significantly lessens the computational burden by avoiding the need for complete retraining when new devices are introduced. For example, Table III shows that our progressive approach reduces memory usage by 38% (1471 MB vs. 2390 MB for bulk training). This efficiency makes DeepRFFinger suitable for environments with limited processing power, such as IoT devices or edge computing systems, where the cost of retraining can be prohibitive. The methodology is inherently scalable, as it can incrementally accommodate new devices without extensive recalibration of the entire model, as demonstrated by its performance on the 17-device dataset (70-80% generalization accuracy, Section VI-G). Progressive learning also allows the system to adapt to changing environmental conditions, such as variations in signal conditions or transmission parameters, without losing previously refined capabilities. New features or classes can be introduced rapidly since the model can learn them without being taken offline for complete retraining, enabling faster deployment in dynamic settings. Additionally, continual learning leads to more effective use of computational and memory resources by focusing on selective retraining rather than holding large datasets or complex models in memory at all times, as evidenced by the reduced memory footprint compared to bulk training.

V. Experimental Setup

To evaluate the effectiveness of our proposed approach, we conducted extensive experiments on datasets of UAV controller signals. This section details the datasets, evaluation metrics, baseline comparisons, and implementation details.

A. Datasets

We evaluate DeepRFFinger using two datasets of UAVs controller transmissions to assess both performance and scalability.

The first dataset [16] includes transmissions from 8 unique devices (matrice, Q205, inspire2, mini2, NineEAGLES, DX4e, Frysky, witoys) captured across different transmission configurations, varying in parameters such as carrier frequency, bandwidth, and symbol rate. Each transmission configuration contains 10M I/Q samples per device. We selected 10 transmissions to use in our model, first two Tx1 and Tx2 for training, while 8 randomly selected from Tx3 to Tx19 are reserved for testing cross-transmission generalization. The dataset is preprocessed to remove channel effects and noise, ensuring that the model focuses on hardware-specific characteristics.

To evaluate scalability and stability with a larger number of devices, we introduce a second dataset [30] comprising 17 unique UAV controllers. This dataset also contains different transmissions where we selected same number of transmissions as dataset 1 and called them the same way to make

the comparison easy (Tx1 to Tx19). Moreover, to make the evaluation consistent with the first dataset, transmissions are limited to 10M sample and Tx1 and Tx2 are used for training, and Tx3 to Tx19 are reserved for testing cross-transmission generalization. Preprocessing steps mirror those of the first dataset to ensure focus on hardware-specific features. This larger dataset allows us to analyze DeepRFFinger's performance under increased device diversity, providing insights into its scalability and robustness.

B. Evaluation Metrics

We evaluated the performance of our model using several metrics:

- **Classification accuracy:** The percentage of correctly classified samples.
- **Cross-transmission generalization accuracy:** Accuracy when testing on transmission instances not seen during training.
- **F1-score:** The harmonic mean of precision and recall, calculated for each device class.
- **Confusion matrix:** Visualization of classification errors between different device classes.
- **Progressive learning efficiency:** Performance improvement after each stage of progressive learning.

C. Baseline Comparisons

We compared our approach with several baseline methods:

- **Raw I/Q CNN-LSTM:** A conventional deep learning approach that directly processes raw I/Q samples using CNN and LSTM layers without custom feature extraction.
- **Statistical feature extraction:** An approach based on extracting statistical features from I/Q samples before classification, a method we introduced in [14].
- **Traditional retraining:** A single-stage training approach that completely retrains the network whenever a new device is added.

Our evaluation focuses particularly on cross-transmission generalization performance, assessing how well each method maintains classification accuracy when faced with previously unseen transmission parameters.

D. Implementation Details

Our model was implemented using PyTorch and trained on NVIDIA V100 GPUs. The training process used the following hyperparameters:

- **Raw IQ:** A conventional deep learning approach that directly processes raw I/Q samples using CNN and LSTM layers without custom feature extraction.
- **RF Features:** A domain-specific approach that extracts hardware-level cyclostationary features (frontend impairments, oscillator characteristics, amplifier signatures, and filter characteristics) prior to classification.
- **Signal Features:** An approach that extracts signal-level cyclostationary features, without considering hardware-specific characteristics.
- **Combined Features:** Our full feature extraction approach with bulk training that integrates both signal and RF hardware features.
- **Combined with CL:** Our full feature extraction approach with progressive continual learning (CL) methodology.

For the cyclostationary feature extraction layers, we used a sequence length of 4096 samples with 20% overlap between consecutive windows. The CNN-LSTM backbone consisted of 2 CNN layers and 2 LSTM layers, with attention mechanisms implemented after the LSTM layers.

For the progressive learning methodology, we implemented the three-stage approach described in Section IV, with stage-specific learning rates and early stopping criteria. This approach enabled efficient adaptation to new devices while maintaining performance on existing ones.

To ensure reproducibility, we fixed the random seed across all experiments and maintained consistent data splitting, preprocessing, and evaluation criteria.

VI. Results and Analysis

In this section, we present a comprehensive evaluation of DeepRFFinger, comparing it with baseline approaches and analyzing the impact of each component on the system's overall effectiveness.

A. Cross-Transmission Generalization Performance

Table II compares our DeepRFFinger approach with recent state-of-the-art RF fingerprinting methods. Our progressive DeepRFFinger with cyclostationary feature extraction achieves 85.9% cross-transmission generalization accuracy, substantially outperforming existing methods where generalization accuracy typically ranges from 27% to 67% [13]. Within our 8-device dataset, DeepRFFinger (bulk training) achieves a validation accuracy of 97.2% and a generalization accuracy of 85.9%, compared to 87.0% validation and 73.4% generalization for the Raw I/Q CNN-LSTM baseline. The progressive variant maintains the same 83.8% generalization accuracy with a validation accuracy of 93.9%.

Fig. 3 illustrates the generalization performance across different test transmissions (Tx3-Tx19) on the 8-device dataset, demonstrating consistently higher accuracy with our approach compared to baseline methods. Most notably, the performance degradation typically observed when encountering new transmission parameters is significantly reduced with our cyclostationary feature-based approach, maintaining an average F1-score of 0.75 across unseen transmissions.

B. Ablation Study Results

To quantify the contribution of each component in our system, we conducted a systematic ablation study, with results presented in Table III. This analysis revealed several key insights:

- **Impact of Cyclostationary Feature Extraction:** Using our custom cyclostationary feature extraction layers consistently improves generalization accuracy by 17 percentage points compared to using raw I/Q data, regardless of

TABLE II
COMPARISON OF RF FINGERPRINTING METHODS

Ref	Year	Method	Features Type	Train acc	Val acc	Gen acc
			Prior Work			
[13]	2024	Parallel CNN-LSTM	Raw IQ and RF Features	0.911	0.877	0.670
[6]	2024	YOLO based	Spectrogram	0.909	0.816	0.488
[8]	2022	CNN+attention	Raw IQ	0.932	0.840	0.640
[9]	2021	Mbed-ATN	magnitude and phase, PSD	0.910	0.880	0.270
			Our Work			
-	2025	DeepRFFinger	Raw IQ	0.839	0.870	0.734
-	2025	DeepRFFinger	Cyclostationary	0.977	0.972	0.859
-	2025	Progressive DeepRFFinger	Cyclostationary	0.939	0.931	0.838

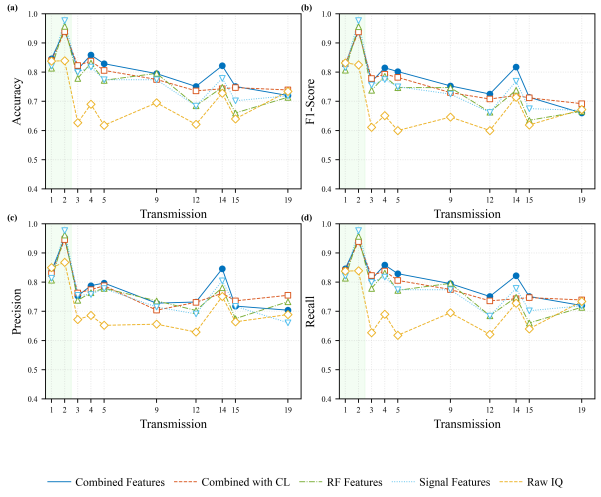


Fig. 3. Cross-transmission generalization performance across different test transmissions. DeepRFFinger with cyclostationary features maintains consistently higher accuracy across varying transmission parameters compared to baseline approaches.

training methodology. This confirms that cyclostationary features effectively capture hardware-specific characteristics that remain consistent across transmission parameters. The F1-score improves from 0.714 (Raw I/Q, bulk) to 0.815 (Cyclostationary, bulk).

- **Value of Progressive Learning:** While bulk training with cyclostationary features achieves slightly higher validation accuracy (97.2% vs. 96.6%), our progressive learning approach delivers competitive generalization performance (83.8%) while reducing memory requirements by approximately 38% (1471MB vs. 2390MB). This trade-off demonstrates the efficiency advantages of our approach for practical deployment scenarios.
- **Necessity of Device-Specific Attention:** Removing the device-specific attention mechanism causes substantial performance degradation across all metrics, with generalization accuracy decreasing by 14.8 percentage points (from 85.9% to 69%) and F1-score dropping to 0.681128 This confirms that attention mechanisms are essential for identifying the most relevant features for each device, particularly for cross-transmission generalization.

C. Confusion Matrix Analysis

The confusion matrices in Fig. 4 provide insight into the classification performance between devices on the 8-device dataset. The raw I/Q baseline shows significant off-diagonal elements, indicating frequent misclassifications between similar devices. In contrast, DeepRFFinger demonstrates much stronger diagonal dominance, with minimal confusion even between devices from the same manufacturer.

This improved discrimination capability is particularly notable for devices sharing similar hardware architectures or from the same manufacturer, where subtle hardware imperfections rather than protocol-level characteristics become the primary differentiating factors. DeepRFFinger effectively isolates these hardware-specific fingerprints, enabling robust device identification even across varying transmission parameters.

D. Progressive Learning Dynamics

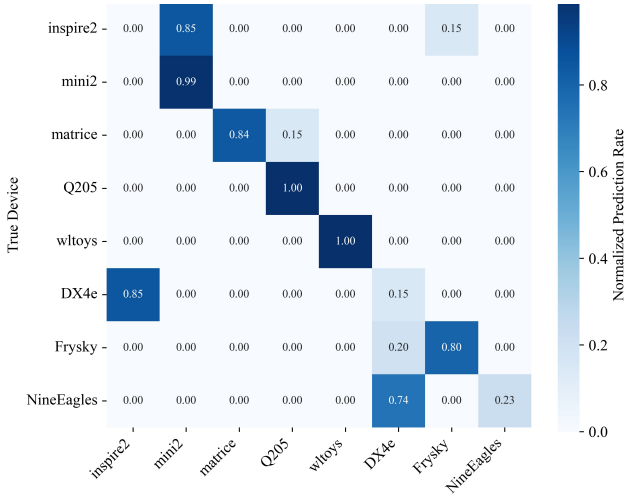
Fig. 5 visualizes the validation accuracy throughout our three-stage progressive learning process on the 8-device dataset. The graph reveals distinct performance patterns at each stage:

- **Stage 1 (New Device Training):** Rapid initial improvement as the model learns device-specific characteristics, with validation accuracy quickly reaching approximately 80%. The plateau in this stage indicate the successful convergence of device-specific components without modifying shared feature extractors.
- **Stage 2 (Device Heads Tuning):** Secondary improvement phase as attention mechanisms are refined across all devices, allowing better selection of features and reaching approximately 90% validation accuracy. The smoother progression reflects the balanced training batches with equal representation from all devices.
- **Stage 3 (Final Adaptation):** Incremental refinement through limited fine-tuning of the entire network, with validation accuracy exceeding 93%. The stable performance throughout this stage confirms that our EWC regularization effectively prevents catastrophic forgetting while allowing beneficial adaptations.

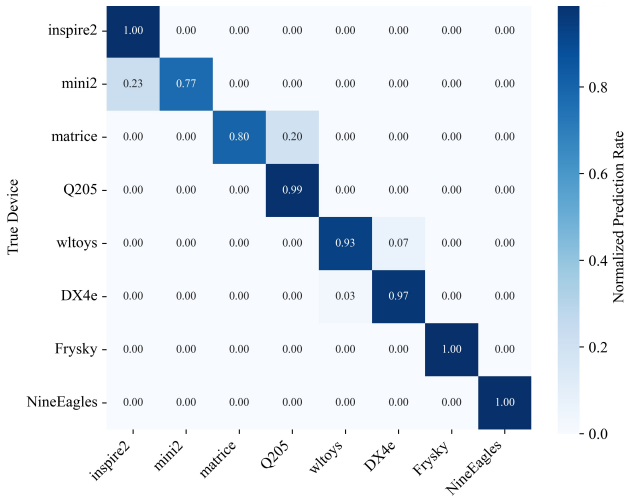
This stage-wise progression validates our progressive learning methodology, confirming that each stage contributes meaningfully to overall performance while maintaining efficiency.

TABLE III
ABLATION STUDY RESULTS

DeepRFFinger Configuration	Feature Type	Val. Acc.	Gen. Acc.	F1-Score	Prec.	Memory (MB)
With CL	Cyclostationary	0.96.6	0.838	0.793	0.763	1471
	Raw IQ	0.839	0.728	0.672	0.689	**1464**
Bulk Training	Cyclostationary	**0.972**	**0.859**	**0.815**	**0.788**	2390
	Raw IQ	0.870	0.734	0.714	0.751	2404
Without Device-Specific Attention	Cyclostationary	0.821	0.690	0.681	0.689	2030
	Raw IQ	0.782	0.618	0.623	0.614	2150



(a) Raw I/Q CNN-LSTM



(b) DeepRFFinger

Fig. 4. Confusion matrices for (a) Raw I/Q CNN-LSTM baseline and (b) DeepRFFinger with cyclostationary features. The baseline shows significant confusion between devices, while DeepRFFinger demonstrates clear diagonal dominance, indicating superior device discrimination capability.

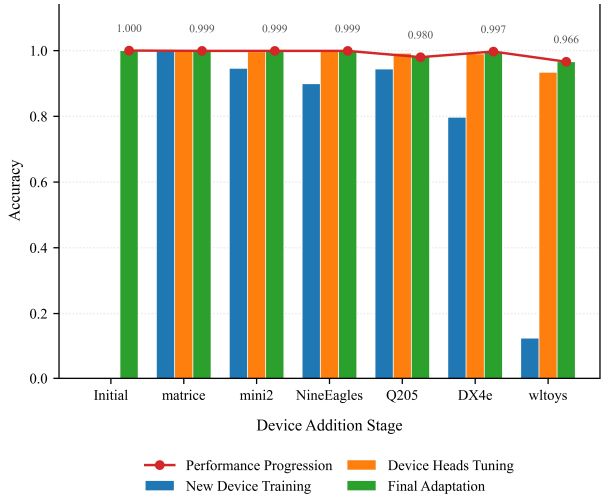
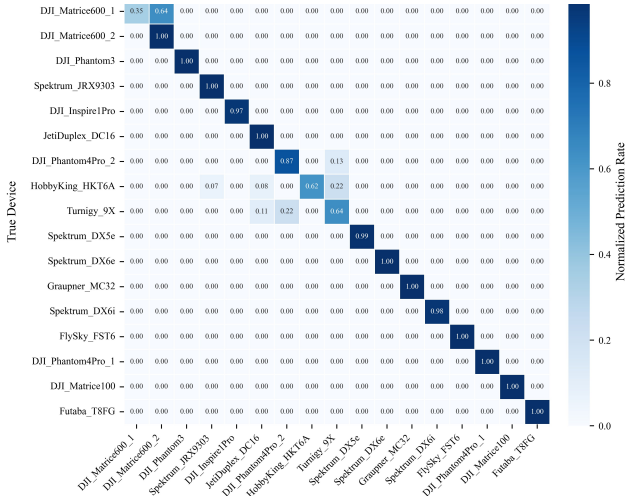


Fig. 5. Progressive learning validation accuracy across three stages: Stage 1 (New Device Training), Stage 2 (Device Heads Tuning), and Stage 3 (Final Adaptation). Each stage contributes distinct improvements to overall performance.

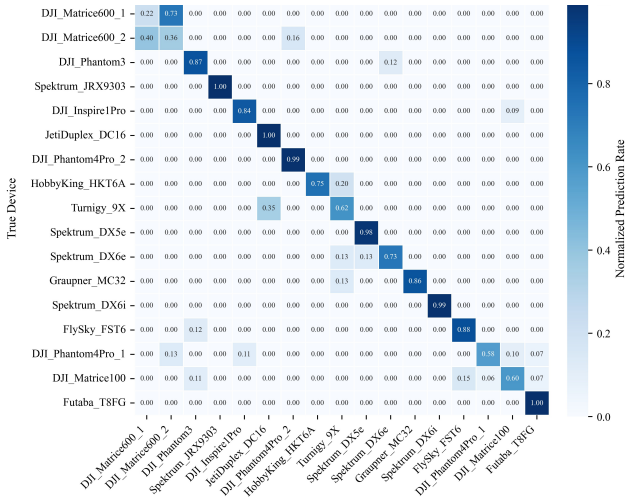
E. Scalability on Second Dataset

To evaluate DeepRFFinger's scalability, we tested it on a larger dataset of 17 UAV controllers [30]. Fig. 6 presents confusion matrices for the 17-device dataset, where devices are grouped based on similarity in hardware characteristics using hierarchical clustering. This ordering highlights patterns of misclassification among similar devices. On seen transmissions (Tx1, Tx2), the model achieves high classification performance, with accuracies of 90.8% (F1-score 0.901) and 86.3% (F1-score 0.848), respectively, compared to 84.6% (F1-score 0.828) and 94.3% (F1-score 0.943) for the 8-device dataset. On unseen transmissions (Tx3-Tx19), performance remains robust, with an average generalization accuracy of 73.9% (F1-score 0.723). Specific examples include Tx3, with an accuracy of 80% (F1-score 0.775), and Tx15, with a lower accuracy of 69.7% (F1-score of 0.686), reflecting the challenge of greater device diversity. The confusion matrix for unseen transmissions (Fig. 6b) reveals increased misclassifications, such as a value of 0.73 between DJI-Matrice600-1 and DJI-Matrice600-2, indicating that 73% of samples from DJI-Matrice600-1 are misclassified as DJI-Matrice600-2. This misclassification arises due to overlapping hardware characteristics (e.g., similar oscillator designs) in the larger device set, as both devices are from manufacturers with same hardware profiles. Despite this, the overall diagonal dominance in the matrix demonstrates that DeepRFFinger maintains robust discrimination, supporting its scalability to larger device sets while highlighting areas for improvement in handling hardware similarity.

Table IV summarizes the cross-transmission generalization performance for selected transmissions (Tx1-Tx19), comparing the 17-device dataset with the 8-device dataset. The 17-device dataset shows a slight performance drop compared to the 8-device dataset's average of 79.2% (F1-score 0.76)



(a) Seen Transmissions



(b) Unseen Transmissions

Fig. 6. Confusion matrices for the 17-device dataset. (a) Performance on seen transmissions, showing near-perfect classification. (b) Performance on unseen transmissions, with an average generalization accuracy of 70-80% and some confusion (e.g., 0.73 between DJI-Matrice600-1 and DJI-Matrice600-2).

on unseen transmissions, highlighting the increased difficulty of scaling to more devices with greater hardware similarity. However, the model maintains robust performance across varying transmission parameters, confirming its scalability.

VII. Discussion

The results presented in the previous section demonstrate the effectiveness of DeepRFFinger in addressing the challenges of cross-transmission generalization and device-specific RF fingerprinting. Several key insights emerge from our evaluation:

A. Hardware vs. Protocol-Level Characteristics

Hardware-specific characteristics, such as oscillator imperfections and front-end impairments, consistently outperform protocol-level features in terms of discriminative power. This finding aligns with RF hardware theory, as manufacturing variations in hardware components create unique fingerprints

TABLE IV
Cross-Transmission Generalization Performance Comparison

Transmission ID	8-Device Dataset		17-Device Dataset	
	Acc.	F1-Score	Acc.	F1-Score
Tx1 (Seen)	0.8457	0.8281	0.9076	0.9013
Tx2 (Seen)	0.9436	0.9431	0.8631	0.8480
Tx3	0.8098	0.7672	0.8004	0.7750
Tx4	0.8590	0.8150	0.7404	0.7342
Tx5	0.8288	0.8016	0.7538	0.7330
Tx9	0.7954	0.7529	0.7409	0.7151
Tx12	0.7509	0.7245	0.7357	0.7158
Tx14	0.8216	0.8182	0.7280	0.7061
Tx15	0.7510	0.7140	0.6971	0.6858
Tx19	0.7197	0.6603	0.7186	0.7191
Avg. (Tx3-Tx19)	0.7921	0.7568	0.7393	0.7230

that are more stable across different transmission parameters than protocol-level characteristics, which can vary significantly with transmission settings. For example, in the 17-device dataset, the confusion matrix for unseen transmissions (Fig. 6b) shows minimal confusion (e.g., 0.73 between DJI-Matrice600-1 and DJI-Matrice600-2) despite these devices sharing similar protocol-level characteristics, suggesting that hardware-specific features drive the discrimination. This insight suggests that future RF fingerprinting systems should prioritize hardware-level feature extraction to achieve robust generalization across diverse transmission scenarios.

B. Advantages of Progressive Learning

The progressive learning approach implemented in DeepRFFinger offers several practical advantages. By incrementally adapting to new devices while preserving performance on previously learned devices, our system avoids the catastrophic forgetting typically associated with sequential learning tasks. The three-stage learning process ensures scalability, as demonstrated by the consistent performance across training stages (Fig. 5), with validation accuracy reaching 96.6% on the 8-device dataset. Moreover, the 38% reduction in memory usage (1471MB vs. 2390MB) compared to bulk training makes DeepRFFinger particularly suitable for resource-constrained environments, such as edge devices in IoT networks. This efficiency, combined with robust generalization (83.8% accuracy on unseen transmissions for the 8-device dataset), positions DeepRFFinger as a viable solution for practical deployment in dynamic wireless environments.

C. Limitations and Future Work

Despite its strong performance, DeepRFFinger has several limitations that warrant further investigation:

- *Scalability to Very Large Device Sets*: Although the 17-device dataset evaluation demonstrates scalability (70-80% generalization accuracy), scaling to datasets with hundreds of devices may introduce additional challenges, such as increased feature overlap and computational complexity. Future work could explore hierarchical classification, where devices are first grouped into clusters based on shared hardware characteristics (e.g., manufacturer or oscillator type) before fine-grained classification, reducing feature overlap. Additionally, feature selection

techniques, such as mutual information-based filtering, could prioritize the most discriminative features to maintain performance on larger device sets.

- *Real-World Deployment Challenges*: Our experiments were conducted in a controlled lab environment. Real-world scenarios may introduce additional noise, interference, multipath fading, and Doppler effects due to mobility, which could degrade performance. Future research should evaluate DeepRFFinger in real-world settings, such as outdoor UAV networks, using standardized testing protocols like those in [28]. This could involve deploying the system on software-defined radios to capture live UAV transmissions and assessing its robustness under varying environmental conditions (e.g., urban vs. rural settings).
- *Adversarial Robustness*: The current system does not account for adversarial attacks, such as feature perturbation where attacker alter signal characteristics to evade detection. To enhance security, future work could incorporate adversarial training strategies, such as training with adversarial examples generated via gradient-based attacks (e.g., Fast Gradient Sign Method) to improve robustness against spoofing. Additionally, robust feature selection methods, such as focusing on immutable hardware features (e.g., oscillator characteristics), could mitigate the impact of feature perturbation attacks.

D. Implications for RF Fingerprinting

The success of DeepRFFinger has significant implications for the field of RF fingerprinting. By demonstrating the feasibility of cross-transmission generalization using cyclostationary features and progressive learning, our work paves the way for more robust and scalable RF fingerprinting systems. The system's ability to maintain high generalization accuracy (85.9% for 8 devices, 70-80% for 17 devices) across varying transmission parameters enhances security in wireless networks by enabling reliable device identification in dynamic environments. Moreover, the reduced memory footprint of our progressive learning approach (1471MB vs. 2390MB) makes it feasible to deploy such systems on resource-constrained devices, broadening their applicability to IoT and edge computing scenarios, such as securing UAV networks or smart city infrastructures.

VIII. Conclusion

This paper presented a deep learning–based RF fingerprinting framework that leverages cyclostationary feature extraction and progressive learning to achieve robust cross-transmission generalization. Evaluation on an 8-device dataset demonstrated a generalization accuracy of 85.9% (F1-score average of 0.76) on unseen transmissions, outperforming baseline methods by 25%, while reducing memory usage by 38% (1471MB vs. 2390MB) through progressive learning. When scaled to a 17-device dataset, the framework maintained strong generalization accuracy in the range of 70–80% (F1-score average of 0.72), confirming its scalability to larger device populations. These findings highlight the ability of the proposed approach to effectively capture hardware-specific fingerprints, enabling reliable device identification in dynamic wireless environments. Future work will focus on real-world deployment, enhancing robustness against adversarial attacks, and further optimizing scalability for very large device sets, paving the way for secure and efficient RF fingerprinting in IoT and edge computing applications.

Acknowledgment

This work was supported by the Mitacs Accelerate Fellowship program, which provided essential funding for the research. The authors would also like to thank Thales Canada for providing access to the hardware/software used in our evaluations, as well as Thales-Ottawa and Quebec RF engineering team for their valuable technical insights and feedback during the development of this work.

References

[1] Y. Xie, Y. Hong, S. Qiao, J. Yao, G. Liu and S. Pang, "A Time-Aware Generative Network for Enhancing Transaction Security in Consumer Electronics," in IEEE Transactions on Consumer Electronics, vol. 71, no. 2, pp. 6818-6828, May 2025, doi: 10.1109/TCE.2024.3511260.

[2] S. Qin et al., "A Partially Labeled Anomaly Data Detection Approach Based on Prioritized Deep Reinforcement Learning for Consumer Electronics Security," in IEEE Transactions on Consumer Electronics, vol. 70, no. 4, pp. 6452-6462, Nov. 2024, doi: 10.1109/TCE.2024.3445629.

[3] H. Byeon et al., "Lightweight AI and Blockchain Optimization for Enhancing Consumer Electronics Decision-Making," in IEEE Transactions on Consumer Electronics, vol. 71, no. 2, pp. 6007-6015, May 2025, doi: 10.1109/TCE.2025.3563412.

[4] A. Rehman, K. Cengiz, S. Ali and K. Ahmad Awan, "H-SecNet: Lightweight and Adaptable Security Framework for IoT-Integrated Consumer Electronics," in IEEE Transactions on Consumer Electronics, vol. 71, no. 3, pp. 8708-8715, Aug. 2025, doi: 10.1109/TCE.2025.3595664.

[5] S. Selvarajan, H. Manoharan, A. O. Khadidos, A. O. Khadidos, A. M. Alshareef and A. Y. Alsobhi, "Secured 6G Communication for Consumer Electronics With Advanced Artificial Intelligence Algorithms," in IEEE Transactions on Consumer Electronics, vol. 70, no. 3, pp. 5711-5718, Aug. 2024, doi: 10.1109/TCE.2024.3382779.

[6] A. Jagannath, Z. Kane, and J. Jagannath, "RF fingerprinting needs attention: Multi-task approach for real-world WiFi and Bluetooth," in Proc. IEEE Global Communications Conference (GLOBECOM), IEEE, 2022.

[7] H. Kulhandjian et al., "AI-based RF-Fingerprinting Framework and Implementation using Software-Defined Radios," in Proc. International Conference on Computing, Networking and Communications (ICNC), IEEE, 2023.

[8] A. Jagannath and J. Jagannath, "Embedding-assisted attentional deep learning for real-world RF fingerprinting of Bluetooth," IEEE Transactions on Cognitive Communications and Networking, vol. 9, no. 4, pp. 940-949, 2023.

[9] J. A. Gutierrez del Arroyo, B. J. Borghetti, and M. A. Temple, "Fingerprint Extraction Through Distortion Reconstruction (FEDR): A CNN-Based Approach to RF Fingerprinting," IEEE Transactions on Information Forensics and Security, 2024.

[10] J. A. Snoap, D. C. Popescu, and C. M. Spooner, "Deep-learning-based classifier with custom feature-extraction layers for digitally modulated signals," IEEE Transactions on Broadcasting, 2024.

[11] L. Wang et al., "A comprehensive survey of continual learning: Theory, method and application," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.

[12] Z. Wang et al., "Federated continual learning for edge-ai: A comprehensive survey," arXiv preprint arXiv:2411.13740, 2024.

[13] P. Patil et al., "Classification of RF Transmitters in the Presence of Multipath Effects using CNN-LSTM," in Proc. IEEE International Conference on Communications Workshops (ICC Workshops), IEEE, 2024.

[14] N. Quadar, A. Chehri, and B. Debaque, "Wireless Security and IoT Device Identification using RF Fingerprinting and Deep Learning," in Proc. IEEE 100th Vehicular Technology Conference (VTC2024-Fall), IEEE, 2024.

[15] J. Sun et al., "Research Progress on the Application of RF Fingerprint Technology in UAV Signal Recognition," in Proc. 7th International Conference on Communication and Information Systems (ICCIS), IEEE, 2023.
[16] S. Basak et al., "Combined RF-based drone detection and classification," IEEE Transactions on Cognitive Communications and Networking, vol. 8, no. 1, pp. 111-120, 2021.
[17] A. Jagannath, J. Jagannath, and P. S. P. Vasanth Kumar, "A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges," Computer Networks, vol. 219, p. 109455, 2022.
[18] H. Gu, L. Su, W. Zhang, and C. Ran, "Attention is needed for RF fingerprinting," IEEE Access, vol. 11, pp. 87316-87329, 2023.
[19] Z. Wen et al., "A hybrid CNN-RF classifier with multi-dimensional early-exit strategy for radio frequency fingerprinting," in Proc. IEEE International Conference on Communications (ICC), IEEE, 2023.
[20] J. Xiao et al., "Progressive Unsupervised Domain Adaptation for Radio Frequency Signal Attribute Recognition across Communication Scenarios," Remote Sensing, vol. 16, no. 19, p. 3696, 2024.
[21] N. Quadar, A. Chehri, B. Debaque, Advanced security frameworks for UAV and IoT: A deep learning approach, Internet of Things, Volume 32, 2025, 101594, ISSN 2542-6605.
[22] J. He, S. Huang, Z. Yang, K. Yu, H. Huan and Z. Feng, "Channel-Agnostic Radio Frequency Fingerprint Identification Using Spectral Quotient Constellation Errors," in IEEE Transactions on Wireless Communications, vol. 23, no. 1, pp. 158-170, Jan. 2024, doi: 10.1109/TWC.2023.3276519.
[23] Y. Zeng et al., "Multi-Channel Attentive Feature Fusion for Radio Frequency Fingerprinting," in IEEE Transactions on Wireless Communications, vol. 23, no. 5, pp. 4243-4254, May 2024, doi: 10.1109/TWC.2023.3316286.
[24] Q. Jiang and J. Sha, "RF Fingerprinting Identification in Low SNR Scenarios for Automatic Identification System," in IEEE Transactions on Wireless Communications, vol. 23, no. 3, pp. 2070-2081, March 2024, doi: 10.1109/TWC.2023.3294988
[25] W. Wu et al., "Reliable resource allocation with RF fingerprinting authentication in secure IoT networks," Science China Information Sciences, vol. 65, no. 7, p. 170304, 2022.
[26] T. Zhao et al., "Drone RF Signal Detection and Fingerprinting: UAVSig Dataset and Deep Learning Approach," in Proc. IEEE Military Communications Conference (MILCOM), IEEE, 2024.
[27] A. Ahmed et al., "A comprehensive survey on deep learning-based LoRa radio frequency fingerprinting identification," Sensors, vol. 24, no. 13, p. 4411, 2024.
[28] M. Ezuma, F. Erden, C. Kumar Anjinappa, O. Ozdemir and I. Guvenc, "Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference," in *IEEE Open Journal of the Communications Society*, vol. 1, pp. 60-76, 2020, doi: 10.1109/OJCOMS.2019.2955889.
[29] E. L. Aleixo, J. G. Colonna, M. Cristo, and E. Fernandes, "Catastrophic Forgetting in Deep Learning: A Comprehensive Taxonomy," JBCS, vol. 30, no. 1, pp. 175–211, Aug. 2024.
[30] Martins Ezuma, Fatih Erden, Chethan K. Anjinappa, Ozgur Ozdemir, Ismail Guvenc, November 25, 2020, "Drone Remote Controller RF Signal Dataset," IEEE Dataport, doi: https://dx.doi.org/10.21227/ss99-8d56.

6.4 Critical Analysis and Thesis Integration

The cyclostationary-based feature extraction and progressive learning methodology we presented in the preceding section marks a clear step forward in addressing RF fingerprinting cross-transmission generalization challenges. We provide here a critical analysis of the approach's contributions, assess scalability characteristics, and identify limitations that motivate the scalable architecture in Chapter 7.

6.4.1 Theoretical Contributions and Performance Assessment

The paper presented in Section 6.3 is submitted as "Integrating Sensing, Communication, and Computing for Robust RF Fingerprinting in Consumer Electronics" to *IEEE Transactions On Consumer Electronics* [32]. Within the thesis, this work directly addresses **Research Question 1** (RQ1): how can RF fingerprinting achieve robustness to transmission parameter variations while maintaining hardware-specific discriminability? The paper's cyclostationary feature extraction layers and three-stage progressive learning methodology were designed and validated on UAV controller datasets specifically to address cross-transmission generalization and memory-efficient continual learning as complementary contributions. The critical analysis below examines integration with the broader thesis framework, identifies scalability limitations that motivate the embedding-based architecture in Chapter 7, and positions progressive learning within the efficiency objective of RQ5.

Our cyclostationary feature extraction framework addresses limitations of both statistical enhancement (Chapter 4) and temporal modeling (Chapter 5) through principled integration of signal processing domain knowledge with adaptive deep learning architectures. The theoretical contribution lies in demonstrating that higher-order statistical transformations can be embedded within differentiable neural network operations, enabling end-to-end learning without preprocessing overhead while preserving the discriminative advantages of cyclostationary analysis.

Empirical validation confirms substantial performance improvements across multiple evaluation dimensions. The 85.9% cross-transmission generalization accuracy on the 8-device UAV controller dataset represents a 43.3 percentage point improvement over the cross-scenario performance from Chapter 5 (42.6%) and a 33.8 percentage point improvement over statistical enhancement (52.1% from Chapter 4). This result shows that cyclostationary features capture hardware-specific periodic patterns that remain discriminative across transmission parameter variations where conventional statistical fea-

tures and temporal modeling exhibit substantial degradation. Progressive learning reaches competitive generalization performance (83.8%) while providing marked memory efficiency improvements (38% reduction: 1471MB versus 2390MB for bulk training). This efficiency gain addresses practical deployment constraints where computational resources are limited, enabling incorporation of new devices without the complete retraining that traditional approaches require. The three-stage process systematically handles feature adaptation, attention mechanism refinement, and classification integration while preserving performance on existing devices through Fisher information-based importance weighting.

Ablation study results establish what each individual component contributes within the integrated framework. Cyclostationary feature extraction layers consistently improve generalization accuracy by 17 percentage points compared to raw I/Q processing, suggesting that principled domain knowledge integration enhances robustness beyond purely data-driven approaches. Device-specific attention mechanisms prove essential for cross-transmission performance, with their removal causing 14.8 percentage point degradation, which shows how important adaptive focus on discriminative features is for this task.

6.4.2 Limitations and Motivations for Scalable Architectures

While integrating cyclostationary feature extraction with temporal modeling (from Chapter 5) yields notable improvements in addressing both cross-transmission and mobility challenges, our critical analysis reveals specific limitations that motivate Chapter 7. Our combined CNN-LSTM-Attention architecture with cyclostationary features shows effectiveness across 8-device and 17-device populations, but the performance degradation pattern (85.9% on 8 devices to 70–80% on 17 devices) suggests that architectural enhancements are necessary for large-scale deployments. While the device-specific attention mechanisms provide adaptive feature weighting, it still introduce computational overhead that scales linearly with device population, which becomes prohibitive for deployments involving hundreds of devices typical of the WiSIG dataset's 174 transmitters we evaluate in Chapter 7.

The closed-set classification approach demonstrates excellent performance for known devices but lacks open-set recognition capabilities necessary for practical deployment. The softmax classifier provides confidence scores over known classes but cannot reliably detect device types not represented in training. This motivates the anomaly detection architectures in Chapter 7, where

reconstruction-based approaches enable unknown device detection suitable for open-set scenarios.

Another limitation is that the architecture, despite progressive learning capabilities for adding new devices, cannot adapt to evolving operational conditions without periodic retraining. While progressive learning addresses device population evolution efficiently (38% memory reduction), it does not handle gradual distribution shifts caused by environmental changes or aging effects across the entire device population. In Chapter 7, we address this through pre-training and fine-tuning paradigms that enable more flexible adaptation. We also note that the evaluation on UAV datasets with controlled transmission parameter variations shows effectiveness for cross-transmission scenarios but does not address the scalability requirements of large-scale IoT deployments where hundreds of devices must be simultaneously managed. WiSIG's 174 transmitters enables evaluation of scalable architectures that build upon our cyclostationary feature extraction principles and address computational efficiency and unknown device detection through transformer-based attention and autoencoder reconstruction approaches.

6.5 Chapter Summary

In this chapter, we established advanced feature engineering approaches that address RF fingerprinting generalization limitations through principled integration of cyclostationary signal processing theory with progressive learning methodologies. Building upon the temporal modeling foundation from Chapter 5, our cyclostationary feature extraction framework shows that higher-order statistical transformations, embedded within differentiable neural network operations, capture hardware-specific periodic patterns that remain discriminative across transmission parameter variations where conventional second-order statistics and temporal modeling alone prove insufficient.

Our critical analysis revealed that while integrating cyclostationary features with temporal modeling effectively addresses both cross-transmission and mobility challenges for moderate-scale deployments, specific limitations motivate Chapter 7. The performance degradation across device populations (85.9% on 8 devices to 70–80% on 17 devices) indicates scalability constraints for large-scale IoT deployments. Device-specific attention mechanisms introduce computational overhead that scales linearly with population size, while the closed-set classification approach lacks the unknown device detection capabilities necessary for open-set deployment. In Chapter 7, we address these limitations through MADE architectures that combine reconstruction-based

anomaly detection with efficient self-attention mechanisms, enabling scalable unknown device detection across 174 transmitters while building upon the cyclostationary feature extraction and temporal modeling principles we validated here and in Chapter 5.

Chapter 7

7 Anomaly Detection and Scalable Architecture

7.1 Introduction

The RF fingerprinting approaches we developed across Chapters 4 through 6 demonstrate clear advances in addressing generalization challenges: statistical enhancement achieves 99.6% same-day accuracy with meaningful cross-day improvements (52.1%), temporal modeling maintains 85.8% accuracy under high-mobility conditions with 100 Hz Doppler shift, and cyclostationary feature extraction achieves 85.9% cross-transmission generalization. However, our critical analysis across these chapters reveals scalability limitations that constrain practical deployment in large-scale wireless environments.

A key issue is that the classification-based frameworks we employed in previous chapters assume closed-set scenarios with predetermined device identities, which limits applicability when unknown devices are encountered during deployment. This closed-set assumption prevents effective handling of rogue devices or sophisticated spoofing attacks that represent critical security threats [23]. Computational and memory requirements also scale unfavorably with device population size, as traditional classification requires $O(N)$ parameter growth as populations expand [54]. Our experimental validation across Chapters 4 through 6 consistently shows performance degradation when scaling beyond moderate device populations, the 17-device UAV evaluation drops to 70–80% accuracy compared to 85.9% on 8 devices, indicating architectural constraints that become prohibitive for contemporary IoT deployments involving hundreds of devices.

This chapter builds upon the methodological foundations from previous chapters while addressing their scalability limitations through architectural innovation. The temporal modeling framework from Chapter 5 demonstrated that CNN-LSTM-Attention architectures effectively capture temporal depen-

dencies under mobility conditions, establishing the importance of attention mechanisms for adaptive focus on discriminative signal characteristics. The cyclostationary feature extraction from Chapter 6 proved that higher-order statistical transformations capture hardware-specific periodic patterns robust to transmission parameter variations.

We propose a Transformer-based architecture that integrates these proven principles within a scalable framework. The transformer-based attention mechanism extends the temporal attention concepts from Chapter 5 through self-attention that scales efficiently to large device populations, replacing device-specific attention heads with shared attention parameters that remain constant regardless of device count. The architecture integrates cyclostationary features from Chapter 6 through flexible input processing, combining the robust generalization properties of higher-order statistical features with efficient embedding-based similarity matching. We also introduce a dual training paradigm that combines reconstruction learning with classification fine-tuning. This addresses the closed-set limitation by enabling unknown device detection through reconstruction error analysis, extending the progressive learning concepts from Chapter 6 to open-set scenarios.

We perform the experimental validation across four evaluation dimensions: WiSIG dataset ablation studies establishing optimal architectural configurations, multi-dataset evaluation demonstrating architectural robustness across diverse RF environments, progressive scalability analysis validating performance across device populations from 5 to 80 devices, and reconstruction performance assessment that provides proof of concept for anomaly detection. The evaluation demonstrates that the MADE architecture addresses the scalability limitations identified in previous chapters while maintaining the robust generalization properties established through statistical enhancement, temporal modeling, and cyclostationary feature extraction. On WiSIG RX77, the architecture reaches 99.9% accuracy under optimal conditions, 92.9% cross-transmission generalization on UAV controllers, and 5.0% performance degradation across 16-fold device scaling (5 to 80 devices).

The work presented in this chapter has been published as "Scalable Deep Learning for RF Fingerprinting: The MADE Architecture for Robust Physical-Layer Device Identification" in *IEEE Open Journal of the Communications Society* [33]. Section 7.2 presents the complete published work. Section 7.3 provides our critical analysis of the approach's contributions, integration with the experimental framework from Chapter 3, scalability assessment, and limitations defining future research directions. Section 7.4 summarizes the key contributions and positions the MADE architecture within the generalization framework developed across the thesis.

7.2 MADE Architecture for Scalable RF Fingerprinting

This section presents the complete published work introducing the MAsked Denoising autoEncoder (MADE) architecture as a solution to scalability and unknown device detection challenges in RF fingerprinting. The architecture represents a different path from the classification-based frameworks of previous chapters, employing transformer-based self-attention with dual-objective training that combines reconstruction learning with supervised classification. This unified framework enables both accurate device identification for known transmitters and principled anomaly detection for unknown devices through reconstruction error analysis.

The architecture integrates three key innovations building upon foundations from previous chapters: transformer-based self-attention mechanisms that extend the temporal modeling concepts from Chapter 5 with near-constant parameter complexity regardless of device population size; flexible integration of cyclostationary features from Chapter 6 that maintain robust cross-transmission generalization properties; and masked denoising autoencoder pre-training adapted from natural language processing and computer vision [84, 85, 86] to RF signal characteristics. The embedding-based similarity matching framework addresses the $O(N)$ parameter growth and $O(N^2)$ training complexity constraints of traditional classification approaches [54], enabling practical deployment at IoT network scales involving hundreds or thousands of devices. The experimental validation across WiFi (WiSIG) and UAV controller datasets demonstrates MADE's effectiveness: 99.9% accuracy under optimal conditions (WiSIG RX77), 92.9% cross-transmission generalization on UAV controllers, and controlled 5.0% performance degradation across 16-fold device scaling from 5 to 80 devices. The complete architectural specifications, mathematical foundations, experimental protocols, and performance analyses are presented in the published work below.

Received 27 December 2025; revised 26 January 2026; accepted 15 February 2026. Date of publication 24 February 2026; date of current version 4 March 2026.

Digital Object Identifier 10.1109/OJCOMS.2026.3667450

Scalable Deep Learning for RF Fingerprinting: The MADE Architecture for Robust Physical-Layer Device Identification

NORDINE QUADAR[1], ABDELLAH CHEHRI[1] (Senior Member, IEEE), AND BENOIT DEBAQUE[2] (Member, IEEE)

[1]Department of Computer Science, Royal Military College of Canada, Kingston, ON K7K 7B4, Canada
[2]Thales Defense and Security, Quebec, QC G1P 4P5, Canada

CORRESPONDING AUTHOR: A. CHEHRI (chehri@rmc.ca)

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN-2022-3256 and in part by Thales Digital Identity and Security and the Mitacs Accelerate Fellowship Program.

ABSTRACT Radio Frequency (RF) fingerprinting stands out as a powerful technique for device identification, exploiting intrinsic hardware-induced signal variations that are difficult to mimic and thus provide a robust layer of security in wireless systems. In recent years, deep learning techniques have been increasingly integrated with RF fingerprinting to enhance device identification and strengthen security against spoofing and unauthorized access. Nevertheless, existing deep learning approaches face severe scalability challenges in large-scale wireless environments. Conventional classification-based frameworks are constrained by closed-set assumptions that require predefined device identities, thereby limiting adaptability to unknown devices and causing exponential performance degradation as device populations expand. These limitations undermine the practicality of current solutions in dynamic, heterogeneous deployments. This paper introduces the **MAsked Denoising autoEncoder (MADE)**, a novel architecture designed to overcome these fundamental constraints through embedding-based similarity matching and dual-objective training. Beyond conventional approaches, **MADE** advances the field by enabling scalable and robust physical-layer device identification which is an essential prerequisite for authentication systems that can ensure that each device's unique RF signature serves as a reliable basis for recognition even in large-scale and dynamic environments. The architecture integrates masked denoising autoencoder pre-training with transformer-based feature extraction, combining reconstruction learning with classification fine-tuning to support both supervised device identification and open-set recognition. Extensive experimental validation across WiFi (WiSIG) and UAV controller datasets demonstrates **MADE**'s effectiveness: 99.9% accuracy under optimal conditions, 92.9% cross-transmission generalization on UAV controllers, and controlled 5.0% performance degradation across 16-fold device scaling (5 to 80 devices). In addition, ablation analysis confirms that cyclostationary-based features, when combined with denoising, yield optimal performance. Finally, **MADE** achieves practical deployment feasibility with 12–15ms inference latency and sub-linear computational scaling, establishing a scalable foundation for next-generation wireless security applications.

INDEX TERMS RF fingerprinting, device identification, transformer, autoencoder, deep learning, wireless security, IoT security, UAV detection, scalability.

I. INTRODUCTION

RADIO Frequency (RF) fingerprinting leverages inherent hardware-induced imperfections in wireless transmitters to achieve fine-grained device identification, thereby establishing a resilient physical-layer security framework that complements and reinforces conventional cryptographic mechanisms. Subtle variations introduced by analog front-end components, such as nonlinearities in power amplifiers, phase noise from local oscillators, in-phase (I) and quadrature (Q), or IQ, imbalance in mixers, and distortions in digital-to-analog converters, combined with micro-electronic manufacturing tolerances, manifest as distinctive and stable signal signatures that persist across transmissions. These hardware-imposed imperfections create unique RF fingerprints that are extremely difficult to replicate, thereby enabling consistent device identification and reinforcing physical-layer security. These unique fingerprints enable consistent recognition of legitimate devices

and constitute a robust defense against spoofing and impersonation attempts, positioning RF fingerprinting as a critical enabler of trustworthy and secure wireless communication [1], [2].

Beyond its physical-layer strength, device identification plays a pivotal role across the Open Systems Interconnection (OSI) protocol stack. Higher-layer functions, from network access control to application-level authentication, ultimately rely on accurate and freely accessible identification mechanisms to guarantee interoperability, trust, and resilience in heterogeneous wireless environments [2], [3], [4]. Recent advances in secure communication frameworks for Internet of Things (IoT) and Internet of Vehicles (IoV) have further emphasized the need for multi-layered security approaches. Kumar et al. [5] addressed cryptographic authentication in Industrial IoT through efficient signature aggregation, while Gupta et al. [6] proposed edge-based deep learning for vehicle behavior analysis. These works highlight that robust device identification at the physical layer serves as a complementary defense mechanism alongside cryptographic and behavioral approaches.

Extensive research efforts have rigorously evaluated RF fingerprinting under controlled laboratory conditions, consistently demonstrating its capacity to deliver high identification accuracy across diverse device populations, transmission scenarios, and experimental setups [7], [8], [9]. These findings provide compelling evidence of RF fingerprinting's feasibility as a physical-layer security mechanism, yet they also expose the critical gap between experimental validation and practical deployment in dynamic, heterogeneous environments.

Ongoing research in RF fingerprinting has been predominantly propelled by classification-based deep learning frameworks, which promise powerful feature extraction capabilities but remain constrained by fundamental scalability issues. For instance, Jian et al. [2] demonstrated both the potential and the inherent limitations of this paradigm: while their ResNet-based architecture achieved impressive accuracy on large radios, its performance deteriorated sharply as the device population expanded, underscoring the vulnerability of closed-set classifiers to real-world diversity and dynamic wireless environments.

Similarly, the WiSig dataset [10] revealed a fundamental weakness, classifiers trained on signals from one receiver often fail to generalize to another, and performance declines sharply when evaluated on signals collected days after training. These issues are not minor engineering challenges; they highlight the intrinsic limitations of classification-centric architectures. Furthermore, these models are based on tenuous assumptions, such as fixed device populations, static channel conditions, and the exclusion of previously unseen transmitters, that fundamentally restrict their generalizability. Consequently, they are inadequate for deployment in dynamic, large-scale, and adversarial wireless environments where adaptability and resilience are essential.

However, operational wireless environments rarely adhere to such restrictive assumptions. In practice, unknown devices emerge unpredictably, channels exhibit rapid temporal variation, and device populations expand dynamically. Under these conditions, the closed-set assumption becomes a critical security vulnerability, as it precludes the detection of rogue transmitters and sophisticated spoofing attacks [11], [12]. Moreover, computational and memory demands scale unfavorably with population size, since traditional classification frameworks require $O(N)$ parameter growth, further undermining their feasibility in large-scale deployments.

Recent research has made notable progress in addressing isolated facets of this challenge. For instance, Hui et al. [13] achieved high accuracy through channel-robust transformer architectures, while Liu et al. [14] advanced open-set detection by introducing dual-mode encoder designs. However, these contributions remain specialized in scope: despite their effectiveness in targeted scenarios, *no existing architecture has yet delivered a unified framework that concurrently ensures reliable device identification, scalability to large heterogeneous populations, and robust anomaly detection*. This absence of an integrated solution *underscores a critical gap in the current research landscape and highlights the pressing need for holistic approaches capable of bridging laboratory success with real-world deployment*.

In this paper, we present the **MAsked Denoising autoEncoder (MADE)**, a unified architecture that directly addresses the scalability and robustness challenges of RF fingerprinting through three key innovations. First, we introduce a dual-objective training paradigm that integrates reconstruction learning with classification fine-tuning, thereby enabling accurate device identification while establishing a principled foundation for open-set recognition via reconstruction error analysis. Second, we employ a transformer-based self-attention mechanism that effectively captures long-range temporal dependencies while maintaining near-constant parameter complexity, with only the final classification layer scaling linearly with device count. Third, we incorporate cyclostationary-based feature extraction, leveraging higher-order statistical patterns to capture hardware-specific characteristics that remain resilient under varying transmission parameters.

Comprehensive experimental evaluation across WiFi (WiSig) and Uncrewed Aerial Vehicle (UAV) controller datasets demonstrates the effectiveness of **MADE**: achieving 99.9% accuracy under optimal conditions, 92.9% cross-transmission generalization on UAV controllers across temporal spans of hours to days, and controlled 5.0% performance degradation under 16-fold device scaling (5 to 80 devices). Ablation analysis further identifies cyclostationary features combined with denoising as the optimal configuration, yielding 71.5% accuracy in low Signal-to-Noise Ratio (SNR) scenarios compared to a 56.2% baseline.

Interestingly, our validation reveals a counter-intuitive yet critical insight: in contrast to masked language modeling in Natural Language Processing (NLP), where masking fosters

richer representation learning [15], random patch masking in RF fingerprinting leads to a measurable 4.6% performance degradation. This divergence stems from the fundamental nature of RF signals, which lack the semantic redundancy characteristic of natural language. Unlike text, where contextual cues can compensate for missing tokens, RF signals rely on continuous temporal structures and fine-grained spectral correlations that are indispensable for device discrimination. Disrupting these patterns through random masking eliminates essential hardware-imposed features rather than encouraging robust reconstruction, underscoring the unique challenges of adapting NLP-inspired techniques to the RF domain.

Our prior research explored CNN-LSTM-attention architectures for LoRa device fingerprinting in mobile scenarios [16] and developed security frameworks for UAV and IoT systems [17]. While these approaches demonstrated effectiveness in their respective domains, they relied on recurrent architectures that compress temporal information into fixed-size representations and operated under closed-set assumptions requiring predefined device identities. The current work advances beyond these foundations by introducing **MADE**, a transformer-based architecture with dual-objective training that achieves superior scalability through constant-complexity attention mechanisms while enabling open-set recognition via reconstruction-based anomaly detection.

II. CONTRIBUTIONS

In this paper, we present the **MADE**, a unified architecture that addresses the scalability and robustness challenges of RF fingerprinting. The main contributions are summarized as follows:

- **Unified Architecture:** We propose **MADE**, which integrates reconstruction learning with classification fine-tuning, enabling accurate device identification while establishing a principled foundation for open-set recognition through reconstruction error analysis.
- **Transformer-based Self-Attention:** We design a lightweight transformer mechanism that effectively captures long-range temporal dependencies while maintaining near-constant parameter complexity in the feature extraction layers, independent of device population size.
- **Cyclostationary Feature Extraction:** We incorporate higher-order statistical patterns to capture hardware-specific characteristics that remain resilient under varying transmission parameters, thereby enhancing robustness in low-SNR conditions.
- **Comprehensive Evaluation:** Through experiments on WiFi (WiSig) and UAV controller datasets, **MADE** achieves 99.9% accuracy under optimal conditions, 92.9% cross-transmission generalization across temporal spans of hours to days, and controlled 5.0% degradation under 16-fold device scaling (5 to 80 devices).
- **Ablation Insights:** Our analysis identifies cyclostationary features combined with denoising as the optimal configuration, yielding 71.5% accuracy in low-SNR scenarios compared to a 56.2% baseline.
- **Counter-Intuitive Finding:** Unlike masked language modeling in NLP, where masking enhances representation learning, random patch masking in RF fingerprinting degrades performance by 4.6%. This highlights the indispensable role of continuous temporal patterns in device discrimination.

The remainder of this paper proceeds as follows. Section III reviews related work in RF fingerprinting and transformer architectures. Section IV presents the proposed **MADE** architecture. Section V describes experimental methodology. Section VI presents results. Section VII discusses the limitations and future work directions, and section VIII concludes our paper.

III. RELATED WORK

A. RF FINGERPRINTING AND DEEP LEARNING METHODS

Every wireless transmitter produces a unique signal signature caused by small hardware imperfections. Differences in components such as oscillators, power amplifiers, and digital-to-analog converters create subtle distortions that act as device fingerprints. These fingerprints remain stable across transmissions, even when channels or modulation schemes change, making them a reliable tool for device identification and physical-layer security.

Scanlon et al. [1] were among the first to systematically validate this phenomenon, achieving 99.8% identification accuracy across 54 Universal Mobile Telecommunications System (UMTS) devices using spectral features combined with mutual information-based feature selection. The significance of their contribution extended beyond raw accuracy: they demonstrated that steady-state signals, not merely transient behaviors, contain sufficient device-specific information to enable reliable authentication. This insight challenged prevailing assumptions and broadened the scope of RF fingerprinting research. Complementary work by Morge-Rollet et al. [18] introduced an eigenfingerprint approach inspired by face recognition, applying singular value decomposition to automatically extract discriminative features for IoT device authentication. Their method demonstrated strong scalability properties and low computational complexity, making it particularly suitable for resource-constrained IoT environments. More recently, Xie et al. [19] provided a comprehensive survey of RF fingerprinting techniques for IoT, systematically categorizing methods by feature extraction strategy and classification approach.

Earlier approaches had primarily relied on hand-crafted features, such as I/Q imbalance arising from analog front-end mismatches or transient characteristics observed during device power-on sequences. While effective in controlled scenarios, these methods lacked generalizability, underscoring the need for more systematic and data-driven approaches

to exploit the full potential of hardware-induced signal signatures. Peng et al. [20] advanced RF fingerprinting by combining nonlinearity and I/Q imbalance features, enabling the classification of devices even within the same model series. However, these feature-engineered methods suffer from a critical limitation: features optimized for a specific protocol or environment often fail to generalize when conditions change. For instance, a classifier trained in an anechoic chamber may struggle significantly in multipath-rich office environments, highlighting the fragility of protocol and environment-dependent approaches.

Deep learning emerged as a promising alternative by learning discriminative features directly from raw data. Jian et al. [2] conducted one of the most ambitious studies to date, analyzing 400 GB of I/Q samples from 10,000 radios. Their ResNet-based architecture achieved strong performance on smaller subsets, but accuracy degraded substantially as device populations scaled beyond several hundred. To mitigate channel effects, they introduced partial equalization, which provided incremental improvements but fell short of delivering a comprehensive solution. Sankhe et al. [21] proposed ORACLE, a Convolutional Neural Networks (CNNs) based architecture that achieved high classification accuracy by optimizing convolutional filter designs specifically for RF waveforms. Their work demonstrated that architecture optimization tailored to RF signal characteristics yields substantial performance improvements over generic CNN designs. Similarly, Yu et al. [22] introduced a multisampling approach that aggregates predictions across multiple signal segments, achieving robustness to temporal variations within transmissions.

The WiSig dataset [10] further exposed fundamental weaknesses in classification-based architectures. With 10 million packets collected from 174 transmitters across 41 receivers over a month, it revealed that classifiers trained on one receiver's data often fail completely when applied to another. Moreover, testing on signals captured days after training resulted in sharp accuracy declines. These are not minor calibration challenges; rather, they underscore the brittle assumptions underlying classification-centric approaches like fixed device populations and static channel conditions that real-world deployments routinely violate. Recent work has investigated diverse architectural directions for RF fingerprinting. Ahmed et al. [23] combined CNNs with bidirectional Gated Recurrent Units (GRUs), achieving 99.65% accuracy at 20 dB SNR on a set of nine Nordic IoT devices. Ali et al. [24] conducted a systematic assessment of features and classifiers for Bluetooth RF fingerprinting, revealing that I/Q imbalance and carrier frequency offset features combined with ensemble classifiers achieve optimal performance. Their comparative analysis highlighted the importance of feature selection for protocol-specific fingerprinting.

Similarly, Guo et al. [25] proposed a cyclic shift method for LoRa, reporting 98.42% accuracy even under non-line-of-sight conditions. Feng et al. [26] explored multi-scale fractal features combined with improved particle swarm optimization for least squares support vector machines, demonstrating that fractal dimension analysis captures hardware impairments across multiple temporal scales. Their approach achieved competitive accuracy while maintaining interpretability through explicit fractal feature extraction. While these results appear compelling, they reveal a recurring limitation: evaluations are typically conducted on small device populations within controlled environments, with training and test data drawn from closely aligned distributions. Such conditions fail to capture the variability and unpredictability of real-world deployments, raising questions about the scalability and generalizability of these approaches.

B. TRANSFORMER ARCHITECTURES FOR RF SIGNALS

Attention mechanisms [27] have recently gained traction in RF fingerprinting as a means of disentangling device-specific features from channel-induced distortions. Hui et al. [13] proposed a cross-attention transformer that achieved high accuracy on WiFi signals by explicitly separating hardware features from channel effects. Their segmented logarithmic spectra embedding allowed the network to emphasize device-specific characteristics while suppressing channel variability, demonstrating the potential of attention-based architectures for robust identification.

Catak et al. [28] conducted a systematic comparison of attention mechanisms for modulation classification, reporting that baseline multi-head attention achieved 85.05% accuracy, while sparse attention reduced inference time by 75% with only marginal accuracy loss. Their findings highlight that attention pattern selection must balance accuracy with computational efficiency, a trade-off we further investigate through our ablation studies. Han et al. [29] extended transformer-based approaches to Automatic Dependent Surveillance–Broadcast (ADS-B) signals from 500 aircraft, incorporating variational mode decomposition. Despite achieving a relatively modest 67.83% recognition accuracy under real-world conditions with Doppler shift and multipath interference, their results underscore the persistent gap between controlled laboratory experiments and operational deployments.

More recently, Liu et al. [14] introduced HyDRA, a dual-mode network that integrates transformer and Mamba encoders to support both closed- and open-set scenarios. On the WiSig SingleDay subset, HyDRA achieved 99.96% closed-set accuracy and 94.67% open-set detection. While the dual-encoder design provides notable flexibility, it also introduces additional architectural complexity, raising questions about scalability and deployment feasibility in resource-constrained environments [30].

C. AUTOENCODER-BASED METHODS

Reconstruction-based learning offers a promising pathway toward open-set recognition, a capability that classification alone cannot provide. This potential has been explored through autoencoder-based methods.

Min et al. [31] applied convolutional autoencoders to 5G signals, achieving a **32.6% average improvement in classification accuracy** through learned denoising. Their lightweight imagification technique transformed I/Q sequences into 2D representations suitable for CNN processing, thereby enabling more effective feature extraction. In a similar study, Karunaratne et al. [32] addressed open-set RF fingerprinting using Variational AutoEncoders (VAEs). Their approach generated synthetic outliers to train classifiers capable of rejecting unknown devices. Conditional VAEs yielded more consistent improvements than standard VAEs, particularly in scenarios with small authorized transmitter populations.

Masked autoencoders have revolutionized self-supervised learning in computer vision [15]. By reconstructing masked image patches, these models acquire rich semantic representations without requiring labeled data. Banerjee et al. [33] extended this paradigm to Multiple-Input Multiple-Output (MIMO) channel estimation, demonstrating that masked autoencoders effectively capture spatial and temporal correlations in wireless channels. Their work on Channel State Information (CSI) feedback compression achieved NMSE improvements from 6.4×10^{-2} to 2.7×10^{-2}, underscoring the broader applicability of masked reconstruction in wireless signal processing.

D. LIMITATIONS OF MASKED RECONSTRUCTION IN RF FINGERPRINTING

Whether masked reconstruction could be effectively applied to RF fingerprinting has remained an open research question. Our experimental results provide a clear conclusion: *it cannot*. The evidence demonstrates that the assumptions underlying masking in vision and language domains do not translate to RF signals, where continuity and fine-grained spectral correlations are indispensable for reliable fingerprint extraction.

The underlying reasons highlight fundamental differences between RF signals and the domains of vision or language where masking has proven successful. Unlike images or text, which benefit from semantic redundancy and contextual recovery, RF signals are inherently dependent on uninterrupted temporal dynamics and precise spectral correlations. Disrupting these structures through random masking removes indispensable hardware-imposed features, thereby degrading discriminative power rather than fostering robust representation learning. This finding underscores the unique challenges of adapting NLP or vision-inspired masking strategies to the RF domain and emphasizes the need for domain-specific approaches that preserve the temporal and spectral continuity of wireless signals.

E. OPEN-SET RECOGNITION AND SCALABILITY

Several recent approaches have tackled open-set detection through diverse mechanisms. Cai et al. [11] combined classification with Siamese comparison, employing paired signal analysis to detect rogue LoRa devices with an Area Under the Curve (AUC) of 0.979 while maintaining **98.47% closed-set accuracy**. The Siamese architecture effectively sidesteps a key limitation of softmax classifiers: as device populations expand, class boundaries crowd together and interfere with one another.

Puppo et al. [12] pursued a different strategy with HiNoVa, extracting hidden state histograms from CNN - Long Short-Term Memory (LSTM) networks as novelty indicators. Their method achieved Area Under the Precision-Recall Curve **(AUPRC) scores of 0.80–1.00** across both LoRa and WiFi datasets, substantially outperforming simpler baselines such as MaxLogit. Importantly, they observed that cyclostationary feature preprocessing enhanced open-set detection, reinforcing evidence that these features capture stable device characteristics useful beyond classification tasks.

Yin et al. [34] extended open-set detection to Long-Term Evolution (LTE) cellular signals using multi-channel CNNs with multi-DCT feature extraction. Their approach achieved **97.65% closed-set accuracy** alongside an unknown detection AUC of 0.8796, demonstrating generalization beyond IoT protocols, albeit with substantial preprocessing requirements. A common thread across these methods is that they rely on architectural modifications tailored specifically for open-set scenarios; none emerge naturally from standard classification training.

Scalability introduces additional challenges. Afrin et al. [35] evaluated Graph Convolutional Networks on expanding device populations, finding that although accuracy degradation remained controlled, training complexity increased substantially. This tension recurs throughout the literature: methods that perform well on 10 or 20 devices often struggle when scaled to 100 or 200.

The difficulty is not purely computational. As populations grow, fingerprint distributions overlap more heavily, and decision boundaries become increasingly fragile [36]. Performance under adverse conditions adds another dimension [37]. Zhang et al. [38] achieved **85.16% accuracy at 0 dB SNR** using dual-path attention fusion of time and frequency features, representing a **37.97% improvement** over prior methods and illustrating the gains achievable through architectural innovation even under severe noise.

Complementary results were reported by Li et al. [39], who employed wavelet transforms with residual attention, reaching **97% accuracy at 10 dB** and **92% at 0 dB**. Together, these findings confirm that thoughtful preprocessing combined with attention mechanisms can significantly extend performance boundaries where raw I/Q processing alone fails.

Our proposed **MADE** builds on these foundations while addressing persistent limitations across the literature. The dual-objective training provides a natural mechanism for open-set detection, as reconstruction error increases for unknown devices without requiring Siamese pairs or histogram analysis. The embedding-based approach maintains constant parameter complexity as device populations scale, avoiding the boundary interference that undermines softmax classifiers. Finally, the integration of cyclostationary features

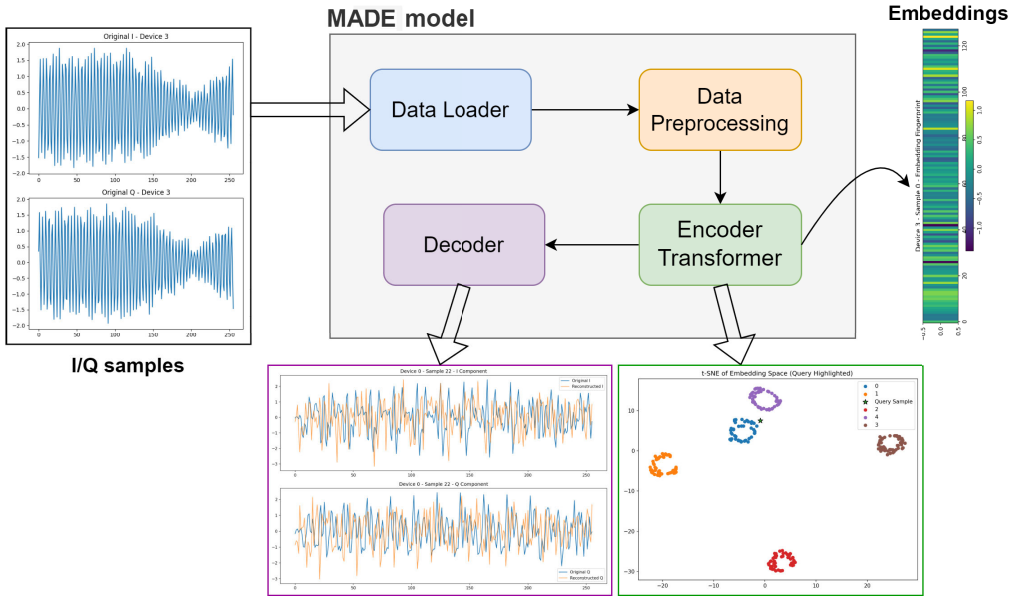


FIGURE 1. MADE architectural overview. The system employs a dual training paradigm: (1) pre-training phase where masked and noisy inputs are reconstructed to learn robust representations, and (2) fine-tuning phase where learned representations are adapted for device classification. The reconstruction branch is maintained during fine-tuning as a regularizer.

captures robust device-specific characteristics that multiple studies have demonstrated to be valuable for both closed- and open-set scenarios.

IV. MADE ARCHITECTURE

A. DESIGN OBJECTIVES OF MADE

We designed the **MADE** to directly confront three persistent challenges in RF fingerprinting. Our integrated design philosophy ensures that solving one challenge inherently contributes to alleviating the others, establishing a unified framework for robust, scalable, and adaptive RF fingerprinting. The key objectives are:

- **Overcoming the Closed-Set Assumption:** MADE is explicitly designed to recognize unknown devices by coupling reconstruction learning with classification fine-tuning, thereby enabling principled open-set detection rather than restricting identification to a fixed device population.
- **Ensuring Scalability:** Unlike conventional classifiers whose complexity grows with the number of devices, **MADE** maintains stable performance under large-scale populations by leveraging transformer-based self-attention mechanisms with near-constant parameter complexity in the core feature extraction layers. The classification head adds only $\mathcal{O}(d_{model} \times N_{classes})$ parameters, representing less than 1% of total model parameters even at 80 devices, while the transformer encoder, MADE module, and attention pooling remain fixed regardless of device population.
- **Enhancing Robustness Across Channels:** MADE mitigates brittleness under varying channel conditions and transmission parameters by integrating cyclostationary feature extraction, which captures hardware-specific statistical patterns resilient to environmental variability.

B. ARCHITECTURE OVERVIEW AND DESIGN RATIONALE

Classification-based approaches continue to dominate RF fingerprinting, largely because they are straightforward to implement and deliver high accuracy under controlled laboratory conditions. Yet, our experiments repeatedly expose a fundamental limitation: classifiers learn decision boundaries that are tightly coupled to the training distribution. When the receiver hardware changes, time elapses, or transmission parameters vary, these boundaries often collapse, leading to severe performance degradation. In essence, the classifier only learns that *device A* is *"not device B"*, without capturing the intrinsic, hardware-imposed characteristics that truly define device A. This shortcoming underscores the fragility of purely discriminative models and highlights the need for architectures that extract stable, device-specific representations resilient to environmental and temporal variability.

Reconstruction-based learning offers a complementary perspective. Training a network to reconstruct signals forces it to capture the essential characteristics that define each device's transmissions. This approach goes beyond mere discrimination, aiming instead to understand signal structure at a deeper level.

The network learns a compressed representation that preserves device-specific information while discarding irrelevant variations introduced by channel effects and noise. Our objective was to combine both capabilities: classification accuracy for practical deployment and reconstruction foundations for detecting unknown devices. Rather than treating these as competing objectives, we designed an architecture in which they reinforce one another. The reconstruction task

regularizes the learned representations, mitigating overfitting to training-specific artifacts, while the classification objective ensures that the representations remain discriminative. In our experiments, this dual-objective paradigm proved more stable than either objective alone.

The proposed **MADE** architecture comprises four interconnected components, as illustrated in Fig. 1 and Fig. 2. A feature preprocessing pipeline processes both raw I/Q samples and engineered cyclostationary-based features. A masked denoising autoencoder learns robust representations through reconstruction, although, as we later demonstrate, the masking component proved counterproductive. A transformer-based attention module captures long-range temporal dependencies that recurrent architectures struggle to model. Finally, a classification head produces device predictions. Fig. 1 illustrates the dual training paradigm: during pre-training, masked and noisy inputs are reconstructed to learn robust representations; during fine-tuning, learned representations are adapted for device classification while maintaining reconstruction as a regularizer. Fig. 2 presents the complete system: RF signals enter through the preprocessing pipeline (left), pass through the MADE encoder-decoder structure (center), and are processed by the transformer module before reaching the dual output heads for reconstruction and classification (right).

The dual-objective training paradigm serves a purpose that extends beyond improving classification accuracy. By maintaining a reconstruction branch throughout training, the network preserves its ability to model what ''normal'' signals from known devices should look like. When encountering unknown devices, by definition absent during training, the latent representations diverge, and the decoder struggles to reconstruct them accurately. This discrepancy provides a principled mechanism for anomaly detection, enabling the identification of out-of-distribution signals without requiring explicit training on unknown device examples.

Transformers are particularly well-suited for RF fingerprinting tasks compared to traditional CNN-LSTM hybrids, primarily due to their ability to effectively model complex dependencies across multiple timescales induced by hardware impairments.

For example, carrier frequency offset manifests as a consistent phase drift over entire packets, amplifier nonlinearities locally distort amplitude ranges regardless of their position in the signal, and oscillator jitter introduces temporally correlated perturbations that may span hundreds of samples. Unlike LSTM-based approaches, which rely on recurrent hidden states and must compress all historical information into a fixed-size vector, potentially leading to information loss, transformers leverage self-attention mechanisms that can capture and directly model both short- and long-range dependencies throughout the signal. This capacity to flexibly attend to relationships across the entire input sequence makes transformers a compelling choice for capturing the nuanced, multi-timescale signatures essential for robust RF fingerprinting.

C. FEATURE PREPROCESSING PIPELINE

The preprocessing pipeline is designed to operate through parallel processing paths, accommodating both raw I/Q sample sequences and engineered cyclostationary features. This dual-path strategy ensures that the network benefits simultaneously from low-level signal representations and higher-order statistical characteristics, providing a richer and more robust feature set for subsequent stages of the **MADE** architecture.

1) I/Q SIGNAL PREPROCESSING

Raw complex-valued RF samples undergo normalization and segmentation:

$$\mathbf{s}_{norm}[n] = \frac{\mathbf{s}_{raw}[n] - \mu_s}{\sigma_s}, \tag{1}$$

where $\mathbf{s}_{raw}[n] = I[n] + jQ[n]$ represents raw complex-valued samples, μ_s is the sample mean, and σ_s is the standard deviation. Normalized signals are segmented into fixed-length sequences of $L = 256$ samples and converted to patches of size 32, resulting in 8 patches per sequence embedded into the model's $d_{model} = 128$ dimensional space.

2) CYCLOSTATIONARY FEATURE INTEGRATION

Communication signals are inherently cyclostationary, with statistical properties that vary periodically due to carrier modulation, symbol timing, and framing structure [40]. This cyclostationarity arises not from the transmitted information itself but from the physical processes underlying transmission, including oscillator frequencies, symbol clocks, and protocol timing. Hardware imperfections modulate these periodic patterns in device-specific ways that persist across different transmission parameters.

The theoretical foundation of our approach lies in the manifestation of hardware impairments within the cyclic domain. Oscillator instabilities induce slight shifts in the cyclic spectrum from its nominal position. Power amplifier nonlinearities generate harmonic components at predictable cycle frequencies. I/Q imbalance introduces asymmetries in the spectral correlation function. Importantly, these effects remain relatively stable across varying channel conditions because they are intrinsic to the transmitter hardware rather than the propagation environment. Traditional cyclostationary analysis requires accurate estimation of cycle frequencies and computationally intensive correlation computations.

D. CYCLIC FEATURE LEARNING IN MADE

In contrast to prior work [41], our approach embeds cyclic operations directly within differentiable neural network layers, allowing the model to adaptively learn which cyclic features are most discriminative for device identification. Rather than relying on manually computed cyclic cumulants as a preprocessing step, we design trainable layers that perform analogous operations end-to-end. This integrated design ensures both adaptability and scalability.

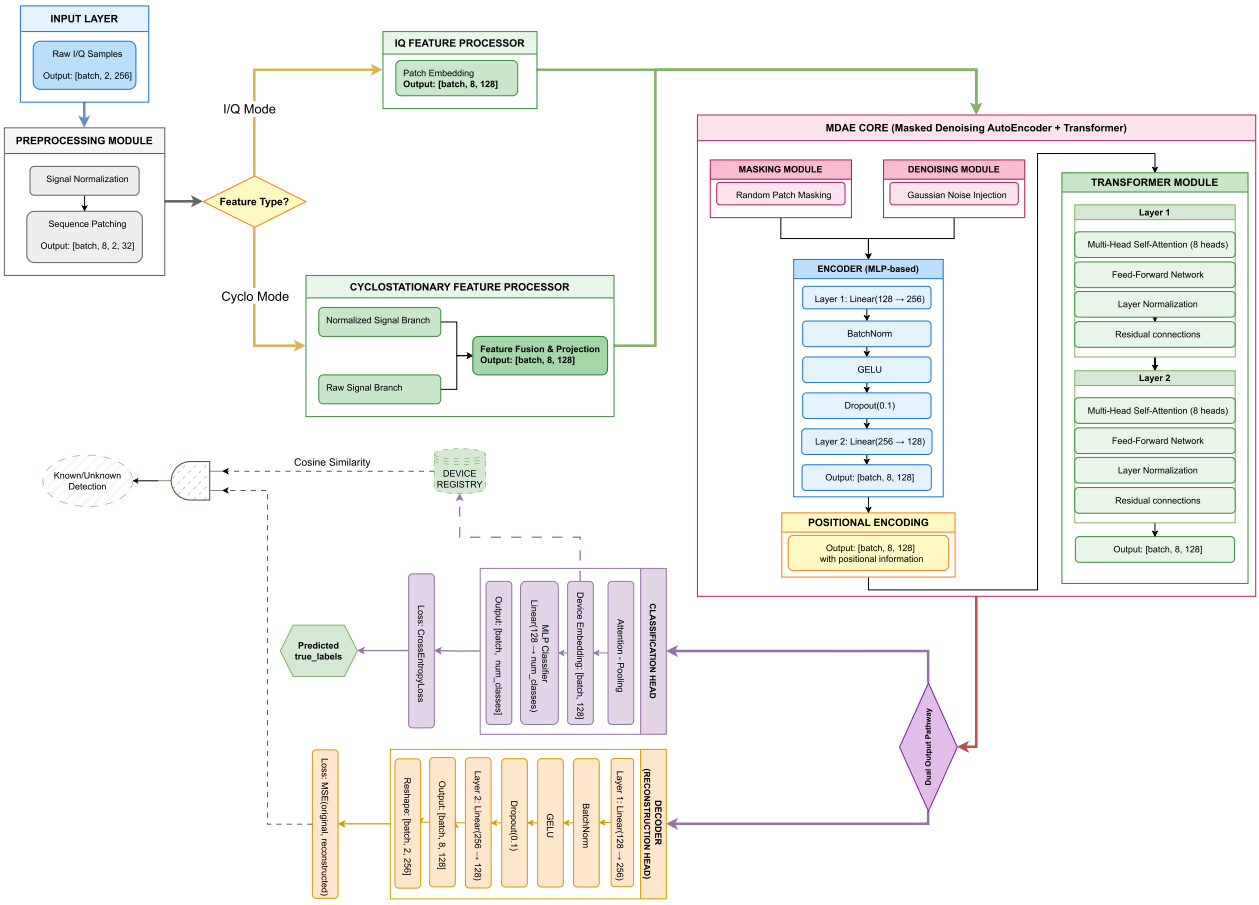


FIGURE 2. **Complete MADE architecture. RF signals enter through the preprocessing pipeline (left), which produces either raw I/Q patches or cyclostationary feature vectors. The masked denoising autoencoder (center) applies controlled corruption during pre-training, forcing the network to learn stable representations. The transformer encoder captures temporal dependencies across patches via multi-head self-attention, and attention-based pooling produces fixed-size device embeddings regardless of input length. Dual output heads (right) serve reconstruction and classification objectives simultaneously.**

We extract features at three statistical orders, each capturing complementary aspects of hardware-specific distortions:

- **Second-Order Features:** Capture direct signal correlations and fundamental hardware characteristics, such as oscillator phase noise and power amplifier nonlinearities, which manifest as stable device-specific signatures.
- **Third-Order Features:** Reveal asymmetries and nonlinear interactions in the analog front-end, including mixer imbalance and Digital-to-Analog Converter (DAC) quantization effects, providing richer discriminative cues beyond simple correlations.
- **Fourth-Order Features:** Exploit higher-order cyclostationary patterns that persist under varying modulation schemes and channel conditions, enabling resilience to environmental variability and long-term temporal drift.

Second-order features are computed as the squared magnitude of the normalized I/Q components, capturing the instantaneous power characteristics:

$$I^{(2)} = I_{norm}^2, \quad Q^{(2)} = Q_{norm}^2. \tag{2}$$

Fourth-order features capture quadratic relationships that are not visible in lower-order statistics. These higher-order dependencies expose nonlinear interactions within the signal, providing richer discriminative information for device fingerprinting. They are formally defined using the following equations:

$$I^{(4)} = (I^{(2)})^2 - (Q^{(2)})^2, \tag{3}$$
$$Q^{(4)} = 2I^{(2)}Q^{(2)}. \tag{4}$$

Sixth-order features capture complex nonlinearities that originate from power amplifier distortion. By extending beyond lower-order statistics, they expose subtle intermodulation effects and harmonic distortions that remain invisible to second- or fourth-order analysis. These higher-order dependencies provide deeper insight into hardware-specific imperfections, thereby strengthening the discriminative power of RF fingerprinting. Formally, sixth-order features are defined using the following equations:

$$I^{(6)} = (I^{(2)})^3 - 3I^{(2)}(Q^{(2)})^2, \tag{5}$$
$$Q^{(6)} = 3(I^{(2)})^2 Q^{(2)} - (Q^{(2)})^3. \tag{6}$$

Moving to higher statistical orders provides a critical advantage: Gaussian noise lacks cyclostationarity, and its

contribution diminishes as higher-order statistics are computed. At sixth order, the noise floor is substantially reduced while hardware-specific signatures remain prominent. This property renders cyclostationary features particularly valuable for RF fingerprinting in environments with fluctuating SNR. Unlike raw I/Q processing, which often degrades under noisy conditions, cyclostationary features exploit higher-order statistical patterns that remain resilient, thereby enabling reliable device discrimination even when signal quality is reduced.

To capture complementary information, signals are processed through parallel paths. The first path applies RMS normalization, transforming I/Q signals to unit power and emphasizing waveform shape while eliminating amplitude variations. In contrast, the second path retains the raw signal, preserving absolute magnitude information that may serve as distinctive device fingerprints but would otherwise be lost during normalization. Both paths are subsequently fed into four specialized RF hardware feature extractors designed to target:

- **Frontend Impairments:** Capture imperfections such as I/Q imbalance, DC offset, and gain mismatch, which introduce systematic distortions in the baseband signal and serve as stable device-specific signatures.
- **Oscillator Characteristics:** Encompass phase noise, frequency stability, and jitter patterns, reflecting intrinsic timing inaccuracies that persist across transmissions and provide reliable discriminative features.
- **Amplifier Signatures:** Characterize nonlinear distortion and envelope variations arising from power amplifier behavior, revealing unique hardware-dependent nonlinearities that strengthen device identification.
- **Filter Responses:** Include anomalies in frequency response and variations in group delay, exposing subtle hardware-induced spectral shaping effects that remain consistent across operating conditions.

The extracted features are subsequently mapped through embedding layers tailored to the transformer's dimensional requirements. This design choice ensures seamless integration, preserving the robust generalization capabilities of cyclostationary analysis while maintaining full end-to-end trainability. Moreover, the architecture is deliberately flexible: it supports lightweight I/Q-only processing for latency-critical, real-time applications, as well as cyclostationary feature processing for scenarios where maximizing discrimination accuracy takes precedence over speed.

E. MASKED DENOISING AUTOENCODER

The **MADE** module functions as the representational backbone of our architecture, purposefully designed to learn compact embeddings that preserve device-specific characteristics while remaining resilient to the variations that typically undermine RF fingerprinting in practice. These variations include channel noise, minor timing offsets, and fluctuations in transmission parameters. By embedding robustness directly into the representational layer, **MADE** establishes a stable foundation that supports both high-accuracy classification and reliable anomaly detection across diverse and dynamic operating conditions.

1) MASKING STRATEGY

We initially integrated masking into our design, inspired by its remarkable success in both NLP and computer vision. In NLP, masked language models achieve robust word representations by predicting missing tokens from surrounding context. In vision, masked autoencoders uncover semantic structure by reconstructing large portions of masked images. The appeal for RF fingerprinting seemed natural: by masking segments of the signal, the network would be forced to learn relationships across temporal regions rather than simply memorizing waveform patterns.

In our architecture, random patch-level masking is applied directly to the embedded signal representation, compelling the model to infer discriminative features from incomplete inputs and encouraging generalization beyond raw memorization.

$$\mathbf{x}_{masked} = \mathbf{M} \odot \mathbf{x} + (1 - \mathbf{M}) \odot \mathbf{t}_{mask}. \tag{7}$$

Here, $\mathbf{M}$ denotes a binary mask with masking probability p_{mask} (typically 0.15–0.25), and $\mathbf{t}_{mask}$ represents learned mask tokens that replace masked patches. Masking is applied at the patch level rather than at individual samples, since hardware signatures span multiple samples; point-wise masking would otherwise yield an artificially easy reconstruction task.

However, as our ablation results demonstrate, this intuition proved incorrect. Masking in fact degrades RF fingerprinting performance. The reasons are instructive, and we analyze them in detail in Section VII.

2) DENOISING STRATEGY

Denoising employs a fundamentally different strategy for learning robust representations. Instead of removing information, it introduces controlled corruption into the input and trains the network to reconstruct the original signal. This process compels the model to focus on stable, device-specific features that persist despite noise or distortion, thereby strengthening resilience in RF fingerprinting. To do that we use the folloiwng equation:

$$\mathbf{x}_{corrupted} = \mathbf{x}_{masked} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2_{noise}\mathbf{I}), \tag{8}$$

where σ_{noise} regulates the noise intensity (typically 0.05–0.1). This process compels the network to distinguish between signal and noise, precisely the capability required when real-world channels distort the hardware fingerprints we aim to identify.

The combined use of masking and denoising creates a deliberately challenging reconstruction task. In this setting, the network is compelled to distinguish features intrinsic to the signal from those arising due to noise or irrelevant variation. Our working hypothesis was that hardware-specific

impairments would remain embedded within the essential feature set, whereas channel effects and environmental noise would be relegated to the non-essential category.

3) ENCODER-DECODER ARCHITECTURE

The encoder transforms corrupted features into compressed latent representations through a three-layer architecture with progressive dimensionality reduction:

$$\begin{aligned}\mathbf{z} &= f_{enc}(\mathbf{x}_{corrupted}),\\ &= \mathcal{D}(\phi(\mathcal{B}(\mathbf{W}_2 \cdot \phi(\mathcal{B}(\mathbf{W}_1\mathbf{x}))))),\end{aligned} \tag{9}$$

where $\phi(\cdot)$ denotes the GELU activation and $\mathcal{D}(\cdot)$ represents dropout ($p = 0.1$) and dimensions $d_{input} \rightarrow 256 \rightarrow 128$. Batch normalization $\mathcal{B}(\cdot)$ stabilizes training and provides implicit regularization, while GELU activations yield smoother gradients than ReLU without sacrificing nonlinearity. Dropout ($p = 0.1$) is applied to mitigate overfitting during pre-training.

The decoder mirrors this structure:

$$\hat{\mathbf{x}} = f_{dec}(\mathbf{z}) = W_4 \cdot \phi(\mathcal{B}(W_3\mathbf{z})), \tag{10}$$

with dimensions $128 \rightarrow 256 \rightarrow d_{input}$. The symmetric design ensures that information lost during encoding can be recovered if adequately represented in the latent space. We deliberately employ a narrow bottleneck (128 dimensions) to enforce aggressive compression, discarding irrelevant variations while preserving discriminative features essential for device identification.

F. TRANSFORMER ATTENTION MODULE

Once **MADE** generates compressed representations, the subsequent step is to model relationships across the entire signal. Hardware impairments frequently manifest as patterns that extend beyond localized patches: carrier frequency offset affects all samples, amplifier nonlinearities distort signals whenever amplitudes surpass certain thresholds, and oscillator jitter introduces correlated perturbations across the sequence. Capturing these dependencies requires a holistic analysis of the signal rather than a patch-by-patch approach, ensuring that global structure and long-range correlations are faithfully preserved within the learned representations.

1) MULTI-HEAD SELF-ATTENTION

Self-attention allows each patch to attend to all others, enabling the network to identify relationships most relevant for device identification. A sample at the end of a sequence can directly compare itself with samples at the beginning, without depending on intermediate states. This capability is particularly advantageous for hardware signatures that manifest consistently across an entire transmission.

In practice, standard multi-head self-attention processes the encoded patch sequences as:

$$\mathcal{A}(\mathbf{Z}) = [\mathbf{head}_1 \| \cdots \| \mathbf{head}_h]\mathbf{W}^O, \tag{11}$$

where $\mathcal{A}(\cdot)$ denotes multi-head self-attention and $\|$ represents concatenation. $W^O \in \mathbb{R}^{(h\cdot d_v)\times d_{model}}$ denotes the output projection matrix, and each attention head computes scaled dot-product attention on distinct linear projections of the input. The use of multiple heads enables the network to capture diverse relationships simultaneously, for example, one head may emphasize phase consistency across patches, while another may highlight amplitude patterns.

In our architecture, we employ $h = 8$ heads with $d_k = d_v = 16$ dimensional projections distributed across two transformer layers. This configuration provides sufficient representational capacity to capture device-specific patterns while avoiding excessive computational overhead. Each layer integrates residual connections and layer normalization, which together stabilize training and ensure effective gradient flow through deep attention stacks.

2) POSITIONAL ENCODING

Transformers inherently treat sequences as unordered sets unless positional information is explicitly encoded. In the case of RF signals, however, temporal ordering is indispensable: phase continuity relies on the sequential arrangement of samples, and hardware impairments evolve in predictable patterns over time. To embed this ordering into the model, we employ sinusoidal positional encoding:

$$PE_{(pos,2i)} = \sin\left(pos \cdot 10000^{-2i/d_{model}}\right), \tag{12}$$

$$PE_{(pos,2i+1)} = \cos\left(pos \cdot 10000^{-2i/d_{model}}\right), \tag{13}$$

where *pos* indexes the patch position and *i* indexes the embedding dimensions. The sinusoidal formulation enables the model to generalize to sequences longer than those encountered during training, while we maintain fixed-length inputs for consistency within our framework.

3) ATTENTION-BASED POOLING

After transformer processing, patch-level representations must be consolidated into a single device embedding suitable for classification. Simple mean pooling assigns equal weight to all patches, which is suboptimal when certain signal segments carry more discriminative information than others. Attention-based pooling overcomes this limitation by learning to emphasize the most informative patches, thereby producing embeddings that better capture device-specific characteristics:

$$\alpha_i = \frac{\exp(w^\top \tanh(\mathbf{h}_i))}{\sum_{j=1}^{N_p} \exp(w^\top \tanh(\mathbf{h}_j))}, \tag{14}$$

$$\mathbf{h}_{pool} = \sum_{i=1}^{N_p} \alpha_i \mathbf{h}_i, \tag{15}$$

where $\mathbf{h}_i$ denotes the transformer output for patch i, w is a learned attention weight vector, and $\mathbf{h}_{pool}$ represents the aggregated device embedding. The tanh nonlinearity bounds attention scores, preventing any single patch from dominating, while softmax normalization ensures the weights sum to one, yielding a proper weighted average.

This pooling mechanism produces fixed-size embeddings regardless of input length, enabling comparison across signals of varying duration and supporting the embedding-based similarity matching employed for scalable device identification.

G. CLASSIFICATION HEAD AND DUAL-OBJECTIVE TRAINING

The final stage transforms pooled embeddings into device predictions while simultaneously maintaining the reconstruction objective that regularizes the learned representations.

1) CLASSIFICATION ARCHITECTURE

We employ a three-layer fully connected network that progressively maps the 128-dimensional pooled embedding into class probabilities:

$$\mathbf{h}_1 = \mathcal{D}(\sigma(\mathcal{B}(\mathbf{W}_{c,1}\mathbf{h}_{pool}))), \tag{16}$$

$$\mathbf{h}_2 = \mathcal{D}(\sigma(\mathcal{B}(\mathbf{W}_{c,2}\mathbf{h}_1))), \tag{17}$$

$$\mathbf{y}_{pred} = \text{Softmax}(\mathbf{W}_{c,3}\mathbf{h}_2), \tag{18}$$

where $\mathcal{B}(\cdot)$ denotes batch normalization, $\sigma(\cdot)$ is the ReLU activation function and $\mathcal{D}(\cdot)$ represents dropout regularization ($p = 0.3$). Dimensions $128 \to 256 \to 128 \to N_{classes}$. Expanding to 256 dimensions in the first layer provides additional capacity for learning nonlinear decision boundaries. Dropout ($p = 0.3$) within the classification layers mitigates overfitting to training-specific artifacts, which is particularly important for achieving cross-transmission generalization.

2) DUAL-OBJECTIVE LOSS FUNCTION

Conventional RF fingerprinting systems are typically optimized exclusively for classification. In contrast, our framework jointly maintains reconstruction and classification objectives throughout training. This dual-objective design encourages the model to preserve device-specific signal characteristics while simultaneously learning discriminative features for identification. The overall training objective is expressed as a combined loss:

$$\mathcal{L}_{total} = \lambda_{recon} \cdot \mathcal{L}_{recon} + \lambda_{cls} \cdot \mathcal{L}_{cls}. \tag{19}$$

The reconstruction loss is formulated as mean squared error (MSE), directly quantifying how well the decoder recovers the original signal:

$$\mathcal{L}_{recon} = \frac{1}{N}\sum_{i=1}^{N} \left\|\mathbf{x}_i - \hat{\mathbf{x}}_i\right\|_2^2. \tag{20}$$

The classification loss employs standard cross-entropy:

$$\mathcal{L}_{cls} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{K} y_{i,k}\log(\hat{y}_{i,k}). \tag{21}$$

Loss weights are adjusted across training phases. During pre-training, $\lambda_{recon} = 0.8$ and $\lambda_{cls} = 0.2$, emphasizing the learning of robust representations before prioritizing classification. This prevents the network from collapsing into trivial solutions that classify well but fail to generalize. During fine-tuning, the weights shift to $\lambda_{recon} = 0.2$ and $\lambda_{cls} = 0.8$, prioritizing classification accuracy while still preserving reconstruction capability. These values were determined empirically through preliminary experiments; ratios in the range of 0.7 to 0.9 for the dominant objective yielded similar results, with 0.8 selected as a balanced configuration that emphasizes reconstruction during pre-training without completely neglecting classification gradients.

Reconstruction is preserved during fine-tuning because the reconstruction loss functions as a regularizer, preventing the classification objective from overfitting the encoder to training-specific artifacts. This constraint compels the encoder to maintain representations that remain meaningful in the signal domain rather than collapsing into narrowly discriminative features. The resulting regularization effect is a central reason why our architecture generalizes more effectively across diverse transmission parameters compared to classification-only approaches.

Algorithm 1 MADE Training Pipeline

Require: Dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$
Ensure: Trained parameters θ_{final}
Phase 1: Pre-training
Initialize $\theta \sim \mathcal{N}(0, 0.02)$
Set $p_{mask} = 0.25$, $\sigma_{noise} = 0.05$
for epoch $= 1$ to E_{pre} **do**
  for batch $\mathbf{B} \in \mathcal{D}$ **do**
    Apply masking and noise to $\mathbf{B}$
    $\hat{\mathbf{B}}, \mathbf{y}_{pred} \leftarrow f_\theta(\mathbf{B}_{noisy})$
    $\mathcal{L} \leftarrow 0.8 \cdot \mathcal{L}_{recon} + 0.2 \cdot \mathcal{L}_{cls}$
    Update θ via AdamW (lr $= 5 \times 10^{-5}$)
  end for
  Early stop if val loss plateaus for 10 epochs
end for
Phase 2: Fine-tuning
Disable masking and denoising
Set differential learning rates: $\alpha_{enc} : \alpha_{trans} : \alpha_{cls} = 1 : 5 : 10$
for epoch $= 1$ to E_{fine} **do**
  for batch $(\mathbf{B}, \mathbf{Y}) \in \mathcal{D}$ **do**
    $\mathcal{L} \leftarrow 0.2 \cdot \mathcal{L}_{recon} + 0.8 \cdot \mathcal{L}_{cls}$
    Update with differential rates
  end for
end for
return θ_{final}

H. TRAINING ALGORITHM

Training the **MADE** model requires careful coordination of its dual objectives. Simultaneous optimization from scratch proves ineffective, since the reconstruction task must first establish meaningful representations before the classification objective can supply informative gradients. To overcome this, we adopt a two-phase training strategy that parallels human

learning: first develop a broad understanding of the domain, then refine and specialize for the classification task.

1) PHASE 1: PRE-TRAINING FOR REPRESENTATION LEARNING

Pre-training is designed to acquire robust signal representations rather than to maximize classification accuracy. All parameters are initialized from a normal distribution with standard deviation 0.02, a choice that empirically stabilizes gradient flow and mitigates saturation in the early layers.

During this stage, each input batch is deliberately corrupted: random patches are masked with probability $p_{mask} = 0.25$, while Gaussian noise with $\sigma_{noise} = 0.05$ is added to the remaining patches. The network is then required to reconstruct the original signal from this corrupted input while simultaneously producing classification predictions. Loss weights are intentionally skewed toward reconstruction ($\lambda_{recon} = 0.8$, $\lambda_{cls} = 0.2$), emphasizing the importance of learning signal structure at this stage over classification accuracy.

Optimization is carried out using AdamW with a learning rate of 5×10^{-5} and weight decay of 0.01. The decoupled weight decay in AdamW provides stronger regularization than the conventional L2 penalty in Adam, which is essential for learning generalizable representations rather than memorizing training examples. Pre-training proceeds for 50 epochs, with early stopping applied if the validation reconstruction loss fails to improve for 10 consecutive epochs.

2) PHASE 2: FINE-TUNING FOR CLASSIFICATION

Once pre-training has established robust representations, the training focus shifts toward classification. Loss weights are inverted to $\lambda_{recon} = 0.2$ and $\lambda_{cls} = 0.8$, prioritizing device identification while retaining reconstruction as a regularizer. Masking and denoising are disabled during fine-tuning, ensuring that the network processes clean signals consistent with inference conditions.

Fine-tuning employs differential learning rates across the architecture. Encoder layers responsible for capturing general signal representations are updated with the smallest learning rate (1×10^{-5}), thereby preserving pre-trained knowledge. Transformer attention layers are trained with a moderate rate (5×10^{-5}), allowing adaptation of attention patterns for classification while maintaining temporal modeling. The classification head, which must learn new decision boundaries, is optimized with the highest rate (10×10^{-5}).

This differential learning rate strategy, expressed as the ratio $\alpha_{enc} : \alpha_{trans} : \alpha_{cls} = 1 : 5 : 10$, prevents catastrophic forgetting of pre-trained representations while enabling targeted task-specific adaptation. Without this mechanism, aggressive fine-tuning would risk erasing the encoder's learned features, thereby negating the benefits of pre-training entirely.

Table 1 summarizes the complete hyperparameter configuration used in our experiments. Fine-tuning is performed for 20 epochs, which we found sufficient for convergence given the strong initialization provided by pre-training. The batch size of 32 strikes a balance between Graphics Processing Unit (GPU) memory utilization and gradient noise; while larger batches offer greater stability, they reduce the diversity of corruptions encountered during pre-training, which is critical for robust representation learning. All experiments were conducted on an NVIDIA RTX 3090 GPU with 24GB memory.

TABLE 1. Training hyperparameters.

Parameter	Pre-training	Fine-tuning
Learning rate	5×10^{-5}	$1–10 \times 10^{-5}$
Epochs	50	20
Masking ratio	0.25	0.0
Denoising std	0.05	0.0
λ_{recon} / λ_{cls}	0.8 / 0.2	0.2 / 0.8
Weight decay	0.01	0.01
Batch size	32	32
Early stopping	10 epochs	–

V. EXPERIMENTAL METHODOLOGY

This section outlines the experimental methodology employed to evaluate **MADE** across multiple dimensions, including component ablation, multi-dataset evaluation, and scalability analysis.

A. DATASETS

We evaluate **MADE** on two complementary datasets that capture distinct dimensions of RF fingerprinting: a large-scale WiFi collection, designed to assess scalability and receiver variability, and a UAV controller dataset, intended to test cross-transmission generalization under realistic operational conditions.

1) WiSIG DATASET

The WiSIG dataset [10] stands as the most comprehensive publicly available resource for RF fingerprinting research. It contains 10 million WiFi packets collected from 174 commercial transmitters using 41 USRP software-defined radio receivers across four sessions conducted over the course of approximately one month. The scale and diversity of this dataset make it possible to evaluate challenges that smaller collections cannot adequately reveal.

To capture different experimental conditions, we employ three distinct receiver configurations:

- **RX11:** Six devices recorded under severe interference conditions characterized by strong multipath effects and nearby transmitter activity. This configuration represents the worst-case performance scenario.
- **RX22:** The same six devices recorded from an alternative receiver position with improved signal quality. Comparisons between RX11 and RX22 highlight the impact of receiver sensitivity on fingerprinting performance.
- **RX77:** An extended configuration employed for large-scale scalability validation under favorable signal

conditions. This setup provides an upper bound on achievable accuracy.

Each configuration comprises raw I/Q samples captured at 20 MSps with the center frequency aligned to WiFi channel 1 (2.412 GHz). For consistency, we employ the ''Full WiSig'' preprocessed subset (76.9 GB), which incorporates packet detection, carrier frequency offset estimation, and timing synchronization. This standardized preprocessing ensures fair and reproducible comparison across all experiments. While the WiSIG dataset includes 41 receivers, we strategically selected three single-receiver configurations (RX11, RX22, RX77) that represent distinct operating conditions: RX11 captures severe interference with strong multipath and nearby transmitter activity (challenging scenario), RX22 provides moderate signal quality from an alternative receiver position (intermediate scenario), and RX77 offers favorable conditions (optimal scenario). This selection enables systematic evaluation across the full spectrum of channel qualities while maintaining focus on our primary research objectives: scalability and temporal generalization. Cross-receiver evaluation, which addresses receiver-invariant performance, represents a distinct research question that would conflate channel quality effects with receiver hardware variations.

2) UAV REMOTE CONTROLLER DATASET

To validate real-world applicability beyond WiFi, we utilize the UAV remote controller dataset introduced by Basak et al. [42]. This dataset contains transmissions from eight commercial drone controllers: **DJI Matrice**, **Walkera Q205**, **DJI Inspire 2**, **DJI Mini 2**, **NineEAGLES**, **Spektrum DX4e**, **FrSky Taranis**, and **WLtoys**. These devices operate across multiple frequency bands (2.4 GHz and 5.8 GHz) using proprietary protocols, thereby providing protocol diversity that cannot be captured through WiFi-only evaluation.

The dataset comprises multiple transmission sessions recorded across different days and environmental conditions. Each device was recorded during 19 distinct transmission instances (Tx1 through Tx19), with each instance consisting of 10 million I/Q samples collected using USRP B210 receivers. This structure enables rigorous cross-transmission evaluation: training is performed on Tx1 and Tx2 (captured during the initial session), while testing is conducted on Tx3 through Tx19 (spanning temporal gaps ranging from approximately 2 hours between Tx2 and Tx3 to 7 days for the final transmissions, under varying environmental conditions). This protocol directly assesses whether fingerprints learned in one session generalize to subsequent encounters, the precise challenge encountered in operational deployments.

Signal preprocessing includes bandpass filtering centered on each device's transmission frequency, automatic gain control compensation, and DC offset removal. Unlike the WiSIG dataset, where preprocessing is provided, these steps are implemented manually to accommodate the protocol diversity inherent in UAV controller transmissions.

3) DATA STANDARDIZATION

All datasets undergo standardized preprocessing to ensure fair architectural comparison. Signals are normalized to zero mean and unit variance across the entire dataset. Continuous captures are segmented into fixed-length sequences of $L = 256$ samples, which are subsequently divided into 8 patches of 32 samples each for transformer-based processing. The I/Q components are represented as two-channel inputs to preserve phase relationships.

For cyclostationary feature extraction, we employ a 4096-sample window with 20% overlap between consecutive windows. Features are computed at second-, fourth-, and sixth-order statistics, and then downsampled to align with the patch structure.

B. EVALUATION FRAMEWORK

The evaluation framework is structured into four complementary phases:

Phase 1 – Ablation Analysis: Controlled experiments systematically assess individual architectural components, including cyclostationary features (P1), denoising (P2), and masking (P3), across six distinct configurations.

Phase 2 – Multi-Dataset Evaluation: Experiments conducted across diverse RF environments validate that the proposed architectural choices consistently provide benefits under varying signal characteristics.

Phase 3 – Scalability Analysis: Performance is evaluated as device populations scale from 5 to 80 devices, using nested subsets ($D_5 \subset D_{20} \subset D_{40} \subset D_{80}$) to measure scalability.

Phase 4 – Reconstruction Quality: The effectiveness of dual-objective training is validated through analysis of reconstruction fidelity and its correlation with classification performance.

C. PERFORMANCE METRICS

Classification performance is evaluated using accuracy, precision, recall, and the macro-averaged F1-score. Computational efficiency is assessed in terms of training time, memory consumption, and inference latency. Reconstruction quality is measured through MSE and and Peak Signal-to-Noise Ratio (PSNR):

$$\text{PSNR} = 10 \log_{10} \left(\frac{\max(\mathbf{x})^2}{\text{MSE}} \right). \tag{22}$$

All experiments employ stratified dataset splitting (70% training, 20% validation, 10% test) with three independent runs using different random seeds to quantify variance and ensure statistical robustness.

VI. EXPERIMENTAL RESULTS

This section presents comprehensive experimental validation across ablation analysis, multi-dataset evaluation, scalability assessment, and reconstruction quality analysis.

A. ABLATION STUDY RESULTS

Table 2 reports ablation study results on WiSIG RX11, highlighting the optimal configuration. All values include

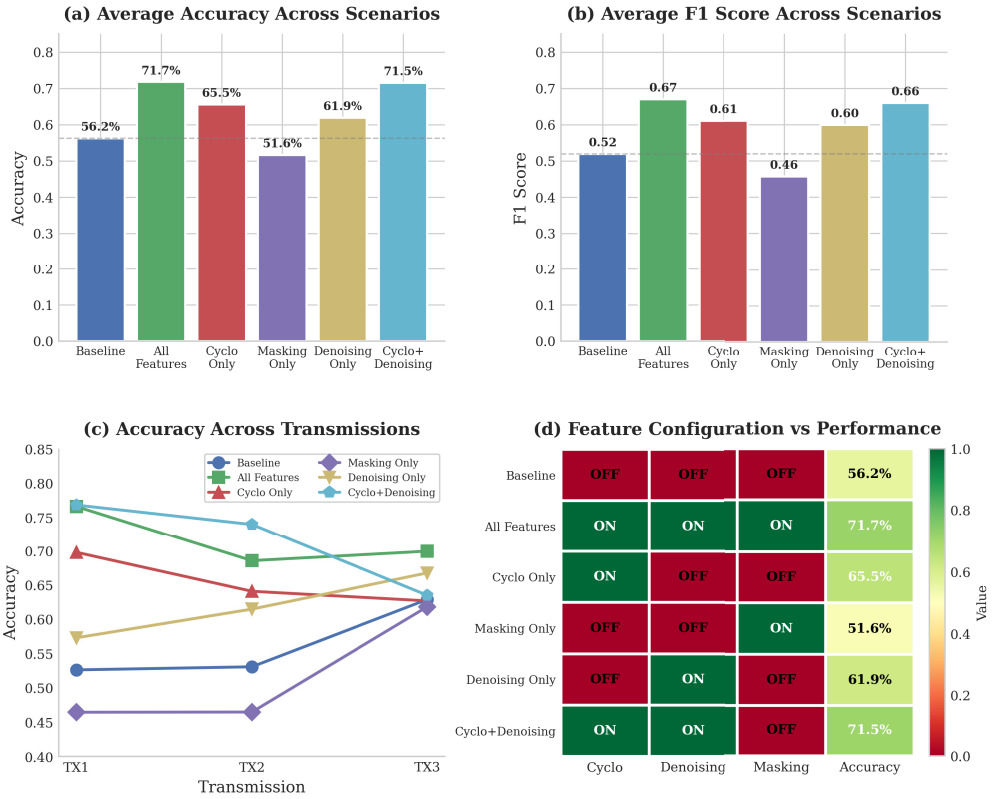


FIGURE 3. Ablation study performance comparison. (a) Average accuracy across evaluation scenarios for each configuration. (b) F1 score comparison showing consistent patterns with accuracy results. (c) Accuracy variation across UAV transmission instances (TX1–TX3 shown as representative subset; the full TX3–TX19 evaluation yields the reported 92.9% aggregate accuracy). (d) Feature configuration matrix summarizing the P1/P2/P3 combinations tested. The optimal configuration (P1+P2) achieves 71.5% accuracy with 1.47× computational overhead.

TABLE 2. Ablation study performance on WiSIG RX11.

Config.	P1	P2	P3	Acc. (%)
Baseline	OFF	OFF	OFF	56.2 ± 0.6
Cyclo Only	ON	OFF	OFF	63.9 ± 0.5
Denoise Only	OFF	ON	OFF	61.9 ± 0.5
Mask Only	OFF	OFF	ON	51.6 ± 0.7
All Features	ON	ON	ON	71.7 ± 0.4
Optimal	**ON**	**ON**	**OFF**	$\mathbf{71.5 \pm 0.4}$

TABLE 3. Multi-dataset evaluation performance.

Dataset	Dev.	Acc. (%)	Prec. (%)	Rec. (%)	F1 (%)
WiSIG RX11	6	71.5 ± 0.7	69.0	71.5	65.9
WiSIG RX22	6	83.5 ± 0.4	84.1	83.5	80.8
WiSIG RX77	6	99.9 ± 0.2	99.9	99.9	99.9
UAV Controllers	8	92.9 ± 0.4	94.5	92.9	92.7

standard deviations computed over three independent runs. Figure 3 visualizes performance comparisons across all tested configurations.

The results reveal distinct contribution patterns, as illustrated in Figure 3. Cyclostationary features yield the largest individual gain (+7.7%), effectively capturing hardware-specific imperfections that persist across transmissions. Denoising contributes an additional +5.7% by emphasizing stable signal characteristics through reconstruction learning. In contrast, masking degrades performance (−4.6%), a divergence from its success in NLP applications where semantic redundancy enables effective masked reconstruction. RF signals lack such redundancy, requiring continuous temporal patterns for reliable device discrimination.

The optimal **P1+P2** configuration achieves near-peak performance (71.5%) while avoiding the degradation introduced by masking. Figure 4 presents the computational resource analysis, showing that overhead remains practical at 1.47× the baseline (66 vs. 45 minutes). This represents an acceptable trade-off for a 27.2% relative improvement in accuracy.

B. MULTI-DATASET EVALUATION

Table 3 presents performance across different RF environments using the optimal configuration.

The three WiSIG configurations clearly demonstrate the impact of channel conditions:

- **RX11 (Interference-Dominated):** Recorded under severe interference with strong multipath and nearby transmitter activity, this configuration achieves 71.5%

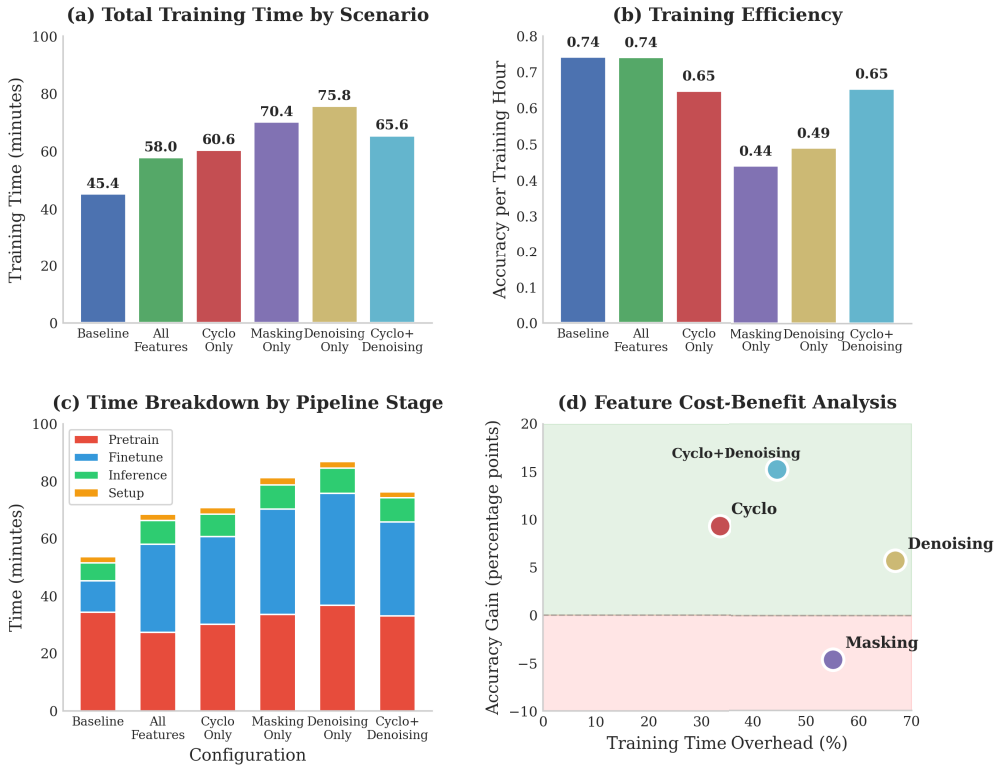


FIGURE 4. Computational resource utilization across configurations.(a) Total training time for each configuration, with baseline at 45 minutes and denoising-only requiring the longest time at 75 minutes. (b) Training efficiency measured as accuracy per training hour, showing that baseline and all-features configurations achieve the highest efficiency. (c) Time breakdown by pipeline stage (pre-training, fine-tuning, inference, and data setup) for each configuration. (d) Feature cost-benefit analysis plotting training time overhead against accuracy gain; cyclostationary features + denoising provide the best trade-off with high accuracy gain at minimal overhead, while masking incurs overhead with negative accuracy impact.

TABLE 4. Progressive scalability analysis.

Devices	Acc. (%)	Prec. (%)	Rec. (%)	F1 (%)	Degrad. (%)
5	80.2 ± 0.5	71.7	80.2	73.9	0.0
20	72.3 ± 0.6	71.5	72.3	69.9	-7.9
40	76.5 ± 0.4	75.3	76.5	75.3	-3.7
80	75.2 ± 0.5	74.3	75.2	74.4	-5.0

accuracy. It represents the lower bound of performance under challenging channel conditions.

- **RX22 (Improved Positioning):** Captured from an alternative receiver placement with enhanced signal quality, RX22 reaches 83.5% accuracy. The improvement over RX11 highlights the sensitivity of RF fingerprinting to receiver positioning and channel clarity.
- **RX77 (Optimal Conditions):** Conducted under favorable signal conditions, RX77 attains near-perfect accuracy of 99.9%. This configuration establishes the upper bound of achievable performance when interference is minimized.

Together, these three configurations reveal a 40% performance variation, underscoring the critical importance of signal quality for reliable RF fingerprinting.

The UAV controller evaluation further validates the architectural effectiveness on independent hardware platforms and proprietary protocols. Cross-transmission testing yields 92.9% accuracy when training is performed on initial transmissions and testing is conducted on later sessions spanning hours to days. These results demonstrate robust temporal generalization that extends beyond WiFi-specific characteristics.

C. SCALABILITY ANALYSIS

Table 4 presents progressive scalability results across device populations. Fig. 5 visualizes the scaling characteristics and performance degradation patterns.

The results in Table 4 and Figure 5 highlight a non-monotonic scaling trend.

When scaling from 5 to 20 devices, performance experiences the steepest degradation (-7.9 percentage points), primarily due to increased class confusion. Interestingly, accuracy recovers at 40 devices ($+4.2$ points), indicating that larger populations enhance feature space coverage and enable more robust decision boundaries. At 80 devices, performance stabilizes with only a minor additional drop (-1.3 points), confirming the resilience of the proposed architecture under expanded populations.

Overall, the total degradation of just 5.0% across a 16-fold increase in device population demonstrates excellent scala-

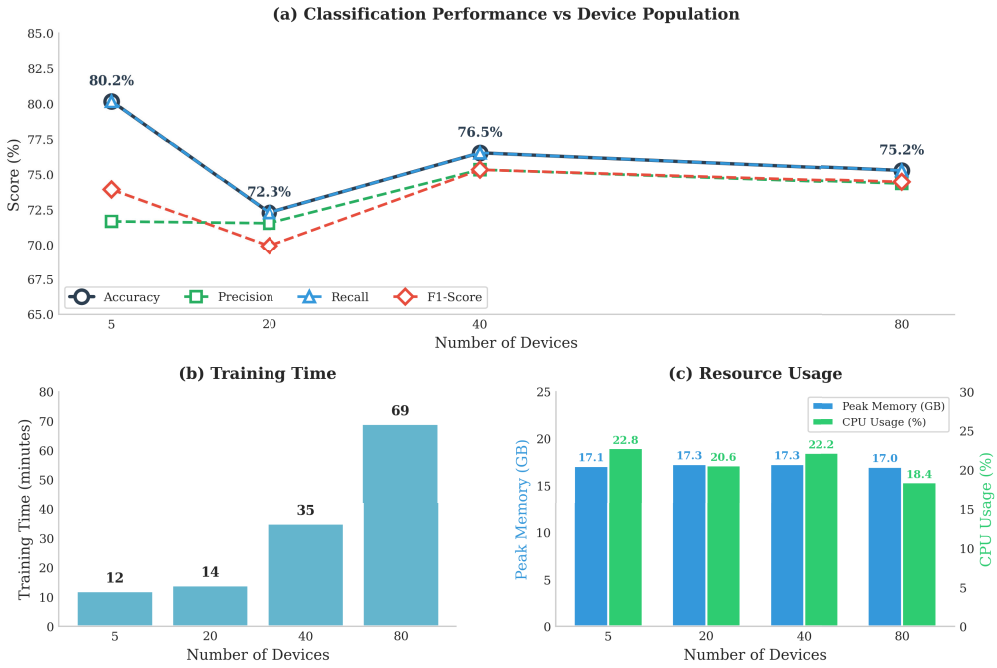


FIGURE 5. Scalability analysis showing performance and resource utilization across device populations. (a) Classification metrics (accuracy, precision, recall, F1-score) as a function of device count, exhibiting non-monotonic behavior with initial degradation at 20 devices (72.3%) followed by recovery at 40 devices (76.5%) and stabilization at 80 devices (75.2%). (b) Training time scaling from 12 minutes (5 devices) to 69 minutes (80 devices), demonstrating sub-linear growth. (c) Resource utilization showing stable peak memory consumption (~17 GB) and CPU usage (18–23%) across all device populations.

bility compared to classification-only baselines, which often suffer exponential decay. The observed sub-linear degradation pattern further underscores the viability of scaling beyond 80 devices without compromising reliability.

Resource utilization also scales predictably with device count. GPU memory increases linearly (∼0.12 GB/device), rising from 6.2 GB (5 devices) to 15.8 GB (80 devices). Training time grows sub-linearly ($O(n^{0.68})$), reflecting efficient optimization, while inference latency remains stable between 12–15 ms regardless of device population size. This balance of accuracy and efficiency demonstrates the practicality of the framework for large-scale deployments.

D. RECONSTRUCTION QUALITY ANALYSIS

Reconstruction evaluation demonstrates effective dual-objective training. Fig. 6 illustrates the learning progression and reconstruction quality achieved by the **MADE**.

As shown in Figure 6(a), the baseline configuration achieves a PSNR of 18.72 dB with a final reconstruction loss of 0.0538, demonstrating high-fidelity signal preservation. Training converges rapidly, with the loss decreasing from an initial value of approximately 0.6 to 0.0538 within just 20 epochs. The contrast between Figure 6(b) and Figure 6(c) further illustrates the substantial improvement in reconstruction quality achieved during training.

Balanced performance across the I/Q components (MSE of 0.0536 for I and 0.0540 for Q) ensures that both magnitude and phase information are preserved, which is essential for reliable fingerprinting. The consistently low reconstruction error observed for known devices establishes a strong baseline for detecting anomalous signals, thereby validating the architectural foundation for open-set recognition.

TABLE 5. Baseline architecture comparison.

Architecture	RX77 (%)	UAV (%)	80-Dev (%)
Traditional CNN	91.2	78.3	61.2
LSTM-Based	89.6	76.9	58.4
CNN-LSTM Hybrid	94.3	85.9	67.3
MADE	**99.9**	**92.9**	**75.2**
Improvement	+5.6	+7.0	+7.9

E. BASELINE COMPARISON

Table 5 compares **MADE** against traditional architectures and Table 6 provides the architectural specifications, parameter counts, and training procedures for all baseline models. All baselines were trained using identical data splits and evaluation protocols.

- **Performance Gains: MADE** consistently improves accuracy across all evaluation scenarios: +5.6% under optimal conditions, +7.0% on UAV cross-transmission, and +7.9% at the 80-device scale. These results confirm the robustness of the architecture across diverse environments and scalability levels.
- **Training Overhead:** The modest increase in training time (21 minutes) represents an acceptable overhead when weighed against the substantial performance

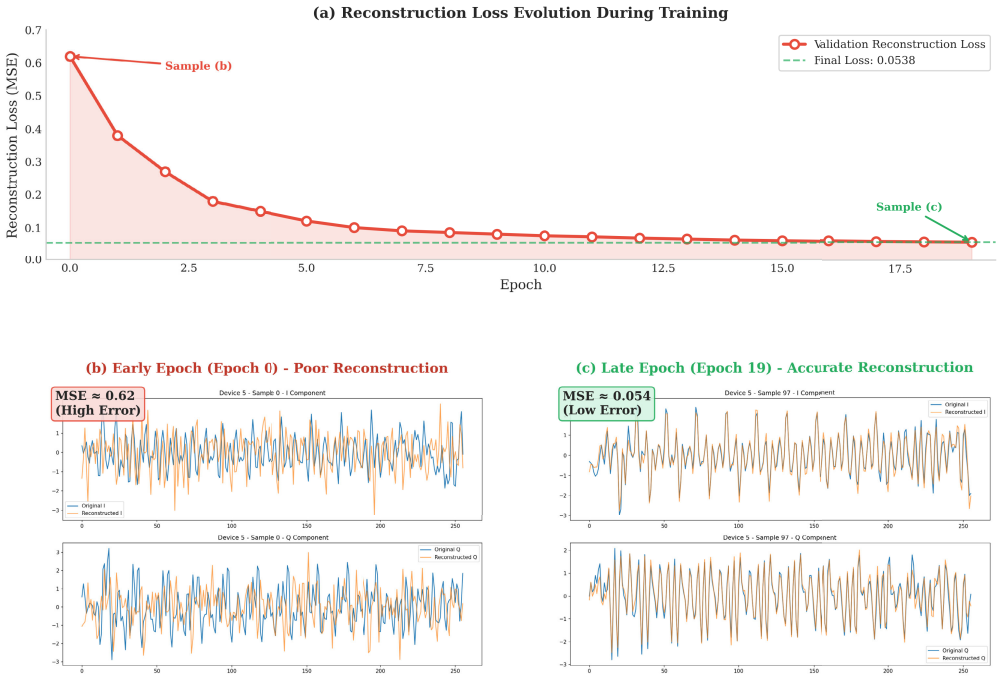


FIGURE 6. Reconstruction quality analysis demonstrating MADE learning progression. (a) Reconstruction loss evolution during training showing rapid convergence, achieving final MSE of 0.0538. (b) Signal reconstruction at epoch 0 showing poor alignment between original and reconstructed I/Q components. (c) Signal reconstruction at epoch 19 demonstrating accurate recovery of both I and Q components with minimal error.

TABLE 6. Baseline architecture specifications.

Architecture	Layer Config.	Params	Training
Traditional CNN	Conv(64,128,256), kernel=3, FC(256,128)	1.2M	50 epochs, lr=10^{-4}
LSTM-Based	LSTM(128 hidden, 2 layers), FC(256,128)	0.8M	50 epochs, lr=10^{-4}
CNN-LSTM Hybrid	Conv(64,128), LSTM(128, 1 layer), FC(256,128)	1.5M	50 epochs, lr=10^{-4}
MADE	See Section IV	2.1M	Table 1

improvements, demonstrating the practicality of the approach for real-world deployments.

- **Statistical Reliability:** All reported results are averaged over three independent runs with different random seeds. Standard deviations remain below $\pm 0.7\%$ across all configurations, ensuring statistical robustness of the findings.
- **Variance Across Scenarios:** The most challenging condition (RX11) exhibits the highest variance, reflecting sensitivity to interference, whereas optimal conditions (RX77) achieve near-perfect consistency with a deviation of only $\pm 0.2\%$. This contrast highlights both the resilience and stability of the proposed framework.

VII. DISCUSSION

A. WHY MASKING FAILS FOR RF SIGNALS

The most surprising finding from our ablation study is that masking, the technique that revolutionized self-supervised learning in NLP and vision, actively degrades RF fingerprinting performance. At first glance, this result appears counterintuitive. Masked language models learn robust representations precisely because they must reconstruct missing context, while masked autoencoders in vision demonstrate that reconstructing 75% of an image from the remaining 25% forces networks to capture semantic structure [15].

RF signals, however, behave fundamentally differently. When random patches of an I/Q sequence are masked, temporal samples containing critical phase continuity information are removed. Unlike text, where surrounding words impose strong semantic constraints on a masked token, RF samples lack such redundancy. A masked sample could plausibly take many values without violating local consistency constraints that the network might otherwise exploit.

The hardware impairments we aim to capture further illustrate this distinction:

- **Carrier frequency offset** induces a consistent phase drift across time.
- **Amplifier nonlinearities** distort the constellation in characteristic ways.
- **Oscillator noise** introduces correlated perturbations throughout the sequence.

All of these signatures depend on continuous observation. Masking disrupts precisely the temporal patterns required for reliable identification. This interpretation aligns with findings

from Guo et al. [25], who showed that cyclic shift characteristics, requiring uninterrupted sample streams, outperform methods that arbitrarily segment signals.

By contrast, denoising preserves sequence continuity while challenging the model to distinguish signal from noise. The reconstruction task remains coherent: given a noisy version of the complete signal, recover the clean version. Hardware impairments are deterministic and repeatable, whereas additive noise is stochastic. Learning to separate these two sources of variation is precisely the capability we seek to develop.

B. THE CYCLOSTATIONARY ADVANTAGE

Cyclostationary features contributed more to performance than any other architectural component. Their inclusion yielded a 7.7% accuracy improvement, surpassing even the 5.7% gain achieved through denoising. This raises an important question: **why do these higher-order statistical features matter so significantly?**

Communication signals are inherently cyclostationary: their statistics vary periodically with time due to carrier modulation, symbol timing, and framing structure. Hardware impairments modulate these periodic patterns in device-specific ways, embedding unique signatures into the cyclic domain.

- **Carrier Frequency Offset:** Introduces a consistent phase drift, shifting the cyclic spectrum in a manner unique to each device's oscillator.
- **Amplifier Nonlinearities:** Distort the constellation in characteristic ways, leaving identifiable traces in higher-order statistics.
- **Oscillator Noise:** Produces correlated perturbations across the sequence, further reinforcing device-specific signatures.

Gaussian noise, by contrast, lacks cyclostationarity. When moving into higher-order cyclic domains, fourth and sixth order in our implementation, the noise contribution approaches zero while hardware signatures remain. Zhang et al. [38] reached similar conclusions, demonstrating 85.16% accuracy at 0 dB SNR through dual-path attention fusion that combines time- and frequency-domain features with transformer-based cross-modal interaction.

The practical implication is clear: preprocessing matters. While raw I/Q processing is faster and simpler, cyclostationary feature extraction provides robustness that raw processing cannot achieve. In applications where reliability outweighs latency constraints, the additional 15% computational overhead is justified by the 27% relative accuracy improvement.

C. SCALABILITY CHARACTERISTICS

Our scalability analysis revealed unexpected non-monotonic behavior. Performance dropped by 7.9 percentage points when scaling from 5 to 20 devices, then partially recovered at 40 devices before stabilizing at 80. This pattern warrants closer examination. The initial drop reflects increased class confusion as the feature space becomes more crowded; twenty devices with similar hardware characteristics inevitably share overlapping fingerprint distributions. The recovery observed at 40 devices suggests a more interesting phenomenon: with sufficient population diversity, the network learns more discriminative representations. Larger device populations provide more training examples, greater variation to regularize against, and ultimately improved generalization.

Jian et al. [2] reported qualitatively similar behavior in their study, though their absolute accuracy levels differed. The critical insight is that scalability does not follow a simple monotonic degradation curve.

Instead, architecture and training procedure interact with dataset diversity in ways that can yield surprising improvements at certain scales.

Overall, our observed 5% total degradation across a 16-fold increase in device population compares favorably with classification-based approaches. Traditional softmax classifiers add parameters linearly with class count, complicating the loss landscape as classes grow. In contrast, **MADE**'s dual-objective training provides regularization that pure classification lacks, while attention-based pooling ensures embeddings maintain constant dimensionality regardless of device count. This combination enables scalable learning without sacrificing robustness, highlighting MADE's suitability for large-scale RF fingerprinting deployments.

D. RECONSTRUCTION QUALITY AND OPEN-SET POTENTIAL

The reconstruction branch achieved a PSNR of 18.72 dB with an MSE of 0.0538, establishing a baseline for what the network considers **"normal"** signal reconstruction. Unknown devices, by definition, are not observed during training. Their signals are expected to produce distinct latent representations, and the decoder, optimized exclusively for known device reconstructions, should therefore yield higher reconstruction errors.

At present, this hypothesis *has not yet been validated with experiments involving truly unknown devices*. Such validation requires careful experimental design: the unknown devices must be genuinely out-of-distribution, rather than simply held-out members of the same device families.

Karunaratne et al. [32] demonstrated that VAE-based open-set detection performs best when outliers are genuinely distinct from authorized transmitters. Our architecture provides the necessary machinery, but operational validation remains future work.

The reconstruction error distribution shown in Figure 6 is concentrated between 0.05 and 0.15, with tight clustering around the mean. This consistency suggests that a threshold-based detection scheme could achieve high precision if unknown devices produce errors outside this range. Whether such deviations occur depends critically on how **"unknown"** the devices truly are, underscoring the importance of rigorous evaluation in open-set scenarios.

E. COMPARISON WITH RECENT TRANSFORMER METHODS

1) COMPARATIVE EVALUATION WITH STATE-OF-THE-ART

Hui et al. [13] achieved 99.72% accuracy with their cross-attention transformer on WiFi signals, slightly below our 99.9% result on WiSIG RX77 under optimal conditions. While direct comparison is complicated by dataset differences, the performance ranges are broadly aligned. Their fingerprint feature separation method explicitly targets channel robustness, whereas our dual-objective training achieves similar robustness implicitly through reconstruction learning.

Liu et al.'s HyDRA [14] reported 99.96% accuracy on WiSIG SingleDay. However, HyDRA integrates both Transformer and Mamba encoders, effectively doubling architectural complexity. In contrast, our single-encoder approach with dual-objective training achieves comparable accuracy with a simpler and more efficient implementation, underscoring the practicality of our design.

The more meaningful comparison lies in cross-transmission generalization. Our 92.9% accuracy on UAV controllers, trained on transmissions 1–2 and tested on transmissions 3–19, demonstrates robust temporal generalization that many laboratory evaluations fail to address. Real-world deployments will encounter precisely this challenge: models trained on initial device captures must reliably recognize those devices days or weeks later under varying conditions.

F. PRACTICAL DEPLOYMENT CONSIDERATIONS

Inference latency of 12–15 ms enables real-time device identification for most practical applications. Given that WiFi packet durations range from hundreds of microseconds to several milliseconds, our processing time is effectively equivalent to the collection of only a few additional packets. For security-critical applications requiring immediate response, this latency may still be excessive; FPGA or ASIC implementations offer a promising path toward substantial reduction.

Memory scaling of 0.12 GB per device implies that a 24 GB GPU can support approximately 150 devices. Metropolitan-scale deployments involving thousands of devices would therefore necessitate distributed processing or more memory-efficient architectures. The embedding-based design naturally facilitates such distribution: devices can be partitioned across multiple processing nodes, each responsible for a subset of similarity comparisons.

Training time scales sub-linearly ($O(n^{0.68})$), meaning that doubling the device count does not double training time. This efficiency arises partly because batch processing amortizes fixed overhead, and partly because the architecture's shared parameters remain constant regardless of device count. For operational systems requiring periodic retraining as new devices enroll, this favorable scaling behavior is highly encouraging and underscores the practicality of the proposed framework for large-scale deployments.

G. LIMITATIONS AND FUTURE WORK

Despite the strong results, several limitations constrain immediate operational deployment while simultaneously defining priorities for future research. These limitations highlight the persistent gap between controlled experimental validation and real-world applicability, underscoring the need for continued investigation into scalability, robustness, and deployment efficiency. Addressing these challenges will be essential to transition from proof-of-concept demonstrations toward reliable, large-scale operational systems.

1) CROSS-RECEIVER GENERALIZATION

Our evaluation employed single-receiver configurations selected to span different channel conditions rather than cross-receiver scenarios. While receiver-agnostic performance is an important challenge in RF fingerprinting [10], our work specifically targets scalability and temporal generalization as complementary research questions. The UAV controller dataset provides partial validation of hardware generalization across different transmitter platforms. Systematic cross-receiver evaluation within the WiSIG dataset like training on one receiver and testing on others, remains an important direction for future work to establish receiver-invariant fingerprinting capabilities.

2) EXTENDING TO HETEROGENEOUS PROTOCOLS

We validated our approach on WiFi and UAV controller signals; however, cellular, Bluetooth, LoRa, and emerging 5G/6G protocols present distinct challenges. LoRa's chirp spread spectrum modulation differs fundamentally from OFDM-based WiFi, meaning the signal structures that carry hardware fingerprints are entirely different.

Zhu et al. [43] demonstrated that LoRa fingerprinting benefits from time-frequency image representations, which our current architecture does not generate. Bluetooth, by contrast, operates in frequency-hopping mode, producing discontinuous captures that complicate the temporal modeling upon which our transformer relies. Extending **MADE** to heterogeneous protocol environments will therefore require either protocol-specific preprocessing branches or more flexible input handling capable of accommodating diverse signal structures.

Future work should systematically evaluate the architecture across a broader protocol space. Promising directions include:

- **Protocol-agnostic preprocessing:** Developing a unified stage that normalizes diverse signal types into a common representation while preserving hardware-specific characteristics.
- **Adaptive cyclostationary feature extraction:** Tailoring cyclic-domain analysis to protocol-specific cycle frequencies, leveraging domain knowledge of each protocol's timing structure to enhance discriminative power.

These extensions will be critical for ensuring that **MADE** remains robust and scalable across the heterogeneous

communication environments expected in next-generation wireless systems.

3) ROBUSTNESS AGAINST ADVERSARIAL ATTACKS

We have not yet evaluated robustness against adversarial attacks, as this was beyond the scope of the present work. Nevertheless, this gap is highly relevant for security applications. Sophisticated attackers could potentially craft signals that exploit the reconstruction objective, introducing adversarial perturbations capable of deceiving both classification and anomaly detection. While the dual-objective training may provide some inherent robustness, since perturbations optimized to mislead the classifier could simultaneously increase reconstruction error and thereby trigger the anomaly detector, this remains speculative.

Adversarial RF signals present unique challenges compared to adversarial examples in image or audio domains. Attackers must generate perturbations that survive physical transmission through the wireless channel, which may naturally attenuate or filter certain adversarial patterns. However, attacks specifically targeting hardware fingerprint features could prove more resilient to channel effects, posing a unique threat to RF fingerprinting systems.

Future work should therefore evaluate the architecture against both digital adversarial attacks and physical adversarial attacks, while developing defense mechanisms tailored to each threat model. Such evaluation will be critical for establishing **MADE**'s viability in security-sensitive operational environments and ensuring trustworthy deployment at scale.

4) SCALABILITY CONSIDERATIONS

The observed 5.0% accuracy degradation when scaling from 5 to 80 devices is manageable; however, scaling to hundreds or thousands of devices, as would be typical in metropolitan-scale IoT/UAV deployments, remains untested. The non-monotonic scaling behavior we observed, with initial degradation at 20 devices followed by recovery at 40, suggests that the architecture may handle larger scales more effectively than linear extrapolation would predict. Nevertheless, this hypothesis requires empirical verification. Memory requirements of 0.12 GB per device imply that a 24 GB GPU would be exhausted at approximately 150 devices, necessitating distributed processing or more memory-efficient architectures for truly large-scale deployment.

Future work should therefore evaluate scaling to 200+ devices using hierarchical approaches, such as clustering devices by manufacturer or protocol before performing fine-grained identification within clusters. Knowledge distillation offers a promising avenue for compressing the model to enable edge deployment while maintaining accuracy. Additionally, federated learning approaches could facilitate training across distributed device populations without centralizing sensitive RF captures, thereby improving scalability and privacy simultaneously.

5) TOWARD EDGE DEPLOYMENT

The current implementation assumes datacenter-class GPU hardware, requiring an 8.4 GB memory footprint, multi-hour training times, and inference latencies of approximately 12–15 ms. By contrast, edge computing nodes in IoT/UAV gateways or base stations operate under far tighter constraints on power, memory, and compute resources. Real-time security applications may demand sub-millisecond latency, which the current architecture cannot achieve without significant optimization. Model compression techniques, including pruning, quantization, and knowledge distillation, offer promising pathways toward edge-deployable versions.

Li et al. [39] demonstrated that wavelet-based preprocessing can reduce computational requirements while maintaining accuracy, suggesting that carefully designed preprocessing may trade modest latency for reduced model complexity. Hardware acceleration through FPGA or ASIC implementations could further deliver order-of-magnitude latency improvements for the core transformer operations.

These optimizations represent critical engineering work required to bridge the gap between our research prototype and operational deployment. Ensuring that **MADE** can meet the stringent requirements of edge-based security systems will be essential for practical adoption in real-world environments.

VIII. CONCLUSION

In this paper, we introduced the **MADE** architecture for scalable RF fingerprinting, achieving 99.9% accuracy under optimal conditions, 92.9% cross-transmission generalization on UAV controllers, and a controlled 5.0% degradation across a 16-fold increase in device population. Our ablation analysis revealed that masking, despite its success in NLP and vision, degrades RF performance due to the absence of semantic redundancy, whereas the combination of cyclostationary features and denoising proved most effective.

With inference latency of 12–15 ms and sub-linear computational scaling, the architecture establishes a robust and practical foundation for real-time wireless security. These results demonstrate both high accuracy and operational viability, positioning **MADE** as a strong candidate for next-generation RF fingerprinting.

Future work will focus on three key directions:

- **Extending protocol coverage:** Evaluating performance across heterogeneous communication standards such as cellular, Bluetooth, LoRa, and emerging 5G/6G systems.
- **Validating open-set recognition:** Conducting experiments with genuinely unknown devices to establish reliable anomaly detection capabilities.
- **Developing edge-optimized variants:** Leveraging model compression, preprocessing, and hardware acceleration to enable deployment in resource-constrained environments.

Together, these efforts will bridge the gap between controlled experimental validation and real-world deployment, ensuring that **MADE** can meet the stringent requirements of scalable, secure, and resilient wireless systems.

REFERENCES

[1] P. Scanlon, I. O. Kennedy, and Y. Liu, "Feature extraction approaches to RF fingerprinting for device identification in femtocells," *Bell Labs Tech. J.*, vol. 15, no. 3, pp. 141–151, Dec. 2010.

[2] T. Jian et al., "Deep learning for RF fingerprinting: A massive experimental study," *IEEE Internet Things Mag.*, vol. 3, no. 1, pp. 50–57, Mar. 2020.

[3] H. Gu, L. Su, W. Zhang, and C. Ran, "Attention is needed for RF fingerprinting," *IEEE Access*, vol. 11, pp. 87316–87329, 2023.

[4] Y. Fan and H. Sun, "Environmental causality calibration: Advancing WLAN RF fingerprinting for precise indoor localization," *PLoS ONE*, vol. 19, no. 2, Feb. 2024, Art. no. e0297108.

[5] G. K. Verma, B. B. Singh, N. Kumar, and V. Chamola, "CB-CAS: Certificate-based efficient signature scheme with compact aggregation for industrial Internet of Things environment," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 2563–2572, Apr. 2020.

[6] H. Grover, T. Alladi, V. Chamola, D. Singh, and K.-K.-R. Choo, "Edge computing and deep learning enabled secure multitier network for Internet of Vehicles," *IEEE Internet Things J.*, vol. 8, no. 19, pp. 14787–14796, Oct. 2021.

[7] S. Kuzdeba, J. Carmack, and J. Robinson, "RF fingerprinting with dilated causal convolutions–an inherently explainable architecture," in *Proc. 55th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 2021, pp. 292–299.

[8] B. Johnson and B. Hamdaoui, "On the domain generalizability of RF fingerprints through multifractal dimension representation," *Sensors*, vol. 22, no. 11, p. 4045, 2022.

[9] F. Shi et al., "Enhanced radio frequency fingerprint identification using length-robust representation and incremental learning," *IEEE Internet Things J.*, vol. 12, no. 10, pp. 14709–14719, Oct. 2025.

[10] S. Hanna, S. Karunaratne, and D. Cabric, "WiSig: A large-scale WiFi signal dataset for receiver and channel agnostic RF fingerprinting," *IEEE Access*, vol. 10, pp. 22808–22818, 2022.

[11] D. Cai et al., "Open set RF fingerprinting identification: A joint prediction and Siamese comparison framework," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2025, pp. 1–6.

[12] L. Puppo, W.-K. Wong, B. Hamdaoui, and A. Elmaghbub, "HiNoVa: A novel open-set detection method for automating RF device authentication," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Gammarth, Tunisia, Jul. 2023, pp. 1–7.

[13] H. Hui, C. Wu, and J. Yao, "Cross-attention transformer for channel-robust radio frequency fingerprint identification," *IEEE Sensors J.*, vol. 25, no. 19, pp. 36823–36831, Oct. 2025.

[14] H. Liu, Y. Huang, Y. Gong, Y. Zhai, and J. Lu, "HyDRA: A hybrid dual-mode network for closed- and open-set RFFI with optimized VMD," *IEEE Internet Things J.*, vol. 12, no. 24, pp. 53828–53841, Dec. 2025.

[15] C. Zhang, C. Zhang, J. Song, J. S. K. Yi, K. Zhang, and I. S. Kweon, "A survey on masked autoencoder for self-supervised learning in vision and beyond," 2022, *arXiv:2208.00173*.

[16] N. Quadar, A. Chehri, and B. Debaque, "Robust RF fingerprinting for LoRa IoT devices in mobile scenarios using CNN-LSTM-attention," in *Proc. IEEE 101st Veh. Technol. Conf. (VTC-Spring)*, Jun. 2025, pp. 1–5.

[17] N. Quadar, A. Chehri, and B. Debaque, "Advanced security frameworks for UAV and IoT: A deep learning approach," *Internet Things*, vol. 32, Jul. 2025, Art. no. 101594.

[18] L. Morge-Rollet, F. Le Roy, D. Le Jeune, C. Canaff, and R. Gautier, "RF eigenfingerprints, an efficient RF fingerprinting method in IoT context," *Sensors*, vol. 22, no. 11, p. 4291, Jun. 2022.

[19] L. Xie, L. Peng, J. Zhang, and A. Hu, "Radio frequency fingerprint identification for Internet of Things: A survey," *Secur. Saf.*, vol. 3, Jan. 2024, Art. no. 2023022, doi: 10.1051/sands/2023022.

[20] L. Peng, J. Zhang, M. Liu, and A. Hu, "Deep learning based RF fingerprint identification using differential constellation trace figure," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 1091–1095, Jan. 2020.

[21] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, "ORACLE: Optimized radio classification through convolutional neural networks," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2019, pp. 370–378.

[22] J. Yu, A. Hu, G. Li, and L. Peng, "A robust RF fingerprinting approach using multisampling convolutional neural network," *IEEE Internet Things J.*, vol. 6, no. 4, pp. 6786–6799, Aug. 2019.

[23] N. Ahmed, G. Saleem, H. M. S. Asif, M. U. Younus, and K. Safdar, "Enhancing wireless device identification through RF fingerprinting: Leveraging transient energy spectrum analysis," 2025, *arXiv:2506.17439*.

[24] A. M. Ali, E. Uzundurukan, and A. Kara, "Assessment of features and classifiers for Bluetooth RF fingerprinting," *IEEE Access*, vol. 7, pp. 50524–50535, 2019.

[25] Y. Guo, A. Hu, and Y. Jiang, "Deep learning based RF fingerprint identification using cyclic shift characteristic," in *Proc. 4th Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Guangzhou, China, Jan. 2024, pp. 369–374.

[26] W. Feng, Y. Li, C. Wu, and J. Zhang, "RF fingerprint extraction and device recognition algorithm based on multi-scale fractal features and APWOA-LSSVM," *EURASIP J. Adv. Signal Process.*, vol. 2023, no. 1, pp. 1–27, Dec. 2023.

[27] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 5998–6008.

[28] F. O. Catak, M. Kuzlu, and U. Cali, "Comparative analysis of attention mechanisms for automatic modulation classification in radio frequency signals," 2025, *arXiv:2508.09996*.

[29] J. Han, Z. Yu, and J. Yang, "Real-world UAV recognition based on radio frequency fingerprinting with transformer," *IET Commun.*, vol. 19, no. 1, pp. 1–12, Jan. 2025.

[30] R. Kong and H. Chen, "DeepCRF: Deep learning-enhanced CSI-based RF fingerprinting for channel-resilient WiFi device identification," *IEEE Trans. Inf. Forensics Security*, vol. 20, pp. 264–278, 2025.

[31] W. Min, J. Kim, and O. Jo, "Denoising method for wireless communication signals based on convolutional AutoEncoder," in *Proc. Int. Conf. Artif. Intell. Inf. Commun. (ICAIIC)*, Feb. 2025, pp. 1–6.

[32] S. Karunaratne, S. Hanna, and D. Cabric, "Open set RF fingerprinting using generative outlier augmentation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.

[33] B. Banerjee, A. Nimr, and G. Fettweis, "Applicability of masked autoencoders in wireless communications: Generalizing MIMO channels," in *Proc. IEEE 36th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2025, pp. 1–6.

[34] P. Yin et al., "Multi-channel CNN-based open-set RF fingerprint identification for LTE devices," *IEEE Trans. Cognit. Commun. Netw.*, vol. 10, no. 5, pp. 1788–1800, Oct. 2024.

[35] F. Afrin, N. Moghim, and S. S. Shetty, "ResGCN: A scalable, robust and efficient approach for radio fingerprinting," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2025, pp. 387–393.

[36] S. Hanna, S. Karunaratne, and D. Cabric, "Open set wireless transmitter authorization: Deep learning approaches and dataset considerations," *IEEE Trans. Cognit. Commun. Netw.*, vol. 7, no. 1, pp. 59–72, Mar. 2021.

[37] Y. Zhang, Z. Zhou, and X. Li, "Specific emitter identification handling modulation variation with margin disparity discrepancy," *IEEE Trans. Inf. Forensics Security*, pp. 380–395, 2024.

[38] Y. Zhang, N. Liu, and Z. Pan, "RF fingerprinting based on DAFFM in low SNR," in *Proc. 6th Int. Conf. Comput., Netw. Internet Things (CNIOT)*, May 2025, pp. 1–8.

[39] H. Li, C. Cao, Y. Sun, H. Peng, Y. Li, and J. Sun, "Deep wavelet residual attention network for radio frequency fingerprint recognition under low SNR," in *Proc. IEEE 20th Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, Toronto, ON, Canada, Sep. 2023, pp. 436–444.

[40] G. Yan, X. Fu, Y. Wang, Q. Zhang, and G. Gui, "Radio frequency fingerprint identification towards statistical and deep learning features: Review, recent results and future directions," *Peer-Peer Netw. Appl.*, vol. 18, no. 3, pp. 1–25, May 2025.

[41] J. A. Snoap, D. C. Popescu, and C. M. Spooner, "Deep-learning-based classifier with custom feature-extraction layers for digitally modulated signals," *IEEE Trans. Broadcast.*, vol. 70, no. 3, pp. 763–773, Sep. 2024.

[42] S. Basak, S. Rajendran, S. Pollin, and B. Scheers, "Combined RF-based drone detection and classification," *IEEE Trans. Cognit. Commun. Netw.*, vol. 8, no. 1, pp. 111–120, Mar. 2022.

[43] B. Zhu et al., "MineLoRa: Markov transition fields and adaptive neural networks for LoRa device radio frequency fingerprint identification," in *Proc. Int. Conf. Artif. Intell. Things Syst. (AIoTSys)*, Oct. 2024, pp. 1–6.

7.3 Critical Analysis and Thesis Integration

7.3.1 Architectural Contributions and Performance Assessment

The paper presented in Section 7.2 is published as "Scalable Deep Learning for RF Fingerprinting: The MADE Architecture for Robust Physical-Layer Device Identification" in *IEEE Open Journal of the Communications Society*, vol. 7, 2026 [33]. Within the thesis, this work directly addresses **Research Question 4** (RQ4): how can systems scale to large device populations while providing open-set recognition of unknown devices? The paper's MADE architecture, dual-objective training paradigm, and scalability evaluation on WiSIG and UAV controller datasets were designed to demonstrate that embedding-based approaches fundamentally advance beyond the classification paradigm. The critical analysis below examines how this contribution synthesizes insights from all previous chapters, addresses the combined scalability and open-set challenge that earlier methods could not solve simultaneously, and identifies future research directions that complete the path toward operational deployment.

The MADE architecture addresses key limitations of the classification-based frameworks we employed in previous chapters through several architectural innovations. The transformer-based self-attention mechanism, with its near-constant parameter complexity, differs markedly from the device-specific attention in Chapter 6, where attention mechanisms scaled linearly with device population and introduced computational overhead that became prohibitive for large-scale deployments. The MADE transformer instead maintains shared attention parameters regardless of device count, with only the final classification layer exhibiting linear scaling. In our evaluation, this design choice results in a controlled 5.0% degradation across 16-fold device scaling (5 to 80 devices), compared to the sharp drop observed in Chapter 6 (85.9% on 8 devices to 70–80% on 17 devices).

The dual training paradigm, which combines masked denoising autoencoder pre-training with classification fine-tuning, provides an architectural basis for open-set recognition that was absent from all previous chapters. While cyclostationary features from Chapter 6 reached 85.9% cross-transmission accuracy, the closed-set classification approach could not detect unknown devices encountered during deployment. The MADE reconstruction-based approach supports unknown device detection through distribution mismatch analysis. The reconstruction quality we measured (18.72 dB PSNR) provides initial proof of concept for anomaly detection, though operational threshold cal-

ibration remains future work. We also integrated cyclostationary features from Chapter 6 within the transformer framework, and our ablation analysis shows that this combination yields the best configuration (71.5% accuracy in low-SNR scenarios versus 56.2% baseline). This result suggests that the higher-order statistical transformations developed in Chapter 6 transfer effectively to transformer architectures. The integration maintains robust cross-transmission generalization (92.9% on UAV controllers) while reaching scalability levels that were unattainable with the CNN-LSTM-Attention architectures from Chapters 5 and 6.

Evaluating on WiSIG's 174 transmitters allowed us to assess scalability beyond the UAV controller evaluations in previous chapters. Under optimal conditions (WiSIG RX77), we observed 99.9% accuracy, which is comparable to statistical enhancement from Chapter 4 (99.6% same-day). Cross-transmission performance on UAV controllers (92.9%) exceeds temporal modeling from Chapter 5 (85.8% under mobility) and approaches the cyclostationary results from Chapter 6 (85.9%). What distinguishes the MADE architecture is that it reaches these performance levels while scaling to larger device populations, with degradation patterns that suggest feasibility for moderate-scale IoT deployments.

7.3.2 Limitations and Future Research Directions

While MADE addresses the scalability and open-set recognition limitations identified in previous chapters, our analysis reveals constraints that define research directions beyond this thesis. The scalability results, though improved relative to previous approaches, still show meaningful performance reduction across device populations. The 5.0% degradation from 5 to 80 devices, combined with memory constraints at approximately 150 devices, indicates that truly large-scale deployments (hundreds to thousands of devices) will likely require hierarchical architectures or distributed processing approaches not explored in this thesis.

The open-set recognition capabilities also remains at the proof-of-concept stage within this thesis scope, even though the architecture support through reconstruction-based anomaly detection. The 18.72 dB PSNR reconstruction quality suggests feasibility, but testing against sophisticated adversarial attacks and operational threshold calibration for separating known from unknown devices require investigation beyond this work. The dual training paradigm may provide some inherent adversarial robustness but empirical verification against advanced attacks remains necessary before operational security deployment. Also, our protocol diversity assessment, limited to WiFi

and UAV controllers within this thesis, leaves multi-protocol evaluation as important future work. While cyclostationary features showed cross-protocol robustness in Chapter 6, evaluation across cellular, Bluetooth, LoRa, and emerging 5G/6G protocols would strengthen broader applicability claims. The LoRa devices evaluated in Chapter 6 could provide initial cross-protocol testing within the MADE framework, though a full heterogeneous protocol assessment represents considerable research effort beyond this thesis scope.

The temporal adaptation limitation is a notable gap relative to the progressive learning capabilities developed in Chapter 6. The MADE's dual training paradigm supports unknown device detection but the architecture lacks continual learning mechanisms for efficient adaptation to evolving operational conditions. The progressive learning framework from Chapter 6, which yielded 38% memory reduction while maintaining performance, could potentially integrate with MADE's embedding-based approach, new device embeddings could be added to registries without modifying feature extraction layers. However, investigating continual learning strategies within the MADE framework, particularly the stability-plasticity tradeoffs for reconstruction versus classification objectives, remains future work.

The computational efficiency we measured (12–15ms inference latency) supports real-time processing for many applications but creates barriers for edge deployment in resource-constrained IoT gateways. The 8.4 GB memory footprint and datacenter-class GPU requirements limit deployment flexibility. Model compression through pruning, quantization, and knowledge distillation is likely essential for operational IoT environments. The counter-intuitive finding that random patch masking degrades performance (unlike NLP applications) suggests that compression techniques may also require RF-specific adaptations, accounting for the absence of semantic redundancy and the reliance on continuous temporal structures that our masking ablation analysis identified.

7.4 Chapter Summary

In this chapter, we presented the MADE architecture as a solution to the scalability and open-set recognition challenges identified across Chapters 4 through 6. The architecture combines transformer-based self-attention with a dual training paradigm of reconstruction learning and classification fine-tuning, building on the cyclostationary feature extraction and temporal modeling principles developed in earlier chapters.

Our evaluation showed controlled 5.0% performance degradation across

16-fold device scaling (5 to 80 devices), cross-transmission generalization of 92.9% on UAV controllers, and 99.9% accuracy under optimal conditions on WiSIG. The ablation analysis confirmed that cyclostationary features with denoising yield the best configuration (71.5% versus 56.2% baseline), showing that domain-specific feature engineering from Chapter 6 complements scalable transformer architectures. Also, the reconstruction quality (18.72 dB PSNR) provides initial proof of concept for anomaly detection capabilities. At the same time, our critical analysis revealed limitations that bound these results. Especially we note that memory requirements (0.12 GB per device) impose practical limits at approximately 150 devices, the open-set recognition remains at the proof-of-concept stage requiring adversarial robustness testing, and protocol diversity assessment is limited to WiFi and UAV controllers.

Chapter 8

8 Comparative Analysis and Conclusions

8.1 Introduction

In this thesis, we have presented a progression of RF fingerprinting approaches, each addressing specific limitations identified in previous methods while advancing toward practical solutions for real-world deployment. In Chapter 4, we established statistical feature enhancement that reached 99.6% same-day accuracy through multi-domain aggregation. Chapter 5 developed temporal modeling that maintains 85.8% accuracy under high mobility conditions through CNN-LSTM-Attention architectures explicitly capturing Doppler-induced variations. In Chapter 6, we introduced cyclostationary feature engineering reaching 85.9% cross-transmission generalization by capturing periodic signal properties invariant to transmission parameter variations. Chapter 7 presented the MADE architecture reaching 75.2% accuracy at 80-device scale with parameter-efficient encoding that maintains consistent inference latency across device populations. Each approach contributes unique capabilities and insights that collectively address different facets of the RF fingerprinting challenge.

In this chapter, we synthesize these findings through comparative analysis, position the contributions with respect to the operational security capabilities they enable, and chart future research directions grounded in observed challenges and emerging opportunities. Section 8.2 presents our performance comparison across evaluation scenarios and datasets, establishing evidence-based understanding of when each approach achieves optimal performance. Section 8.3 maps research contributions to the thesis research questions, demonstrating achievement of research goals while identifying limitations. Section 8.4 discusses the security impact and deployment implications of these contributions, positioning the technical results with respect to physical-layer authentication and spectrum monitoring applications. Sec-

tion 8.5 acknowledges current limitations and provides research directions addressing identified gaps. Section 8.6 concludes with synthesis of key insights and articulation of broader impact for wireless security and machine learning communities.

8.2 Comparative Performance Analysis

8.2.1 Performance Comparison

Table 8.1 presents unified performance comparison across all RF fingerprinting approaches developed in this thesis, evaluated on common metrics where datasets and experimental configurations permit direct comparison.

Table 8.1: Performance Comparison of RF Fingerprinting Approaches

Method	**Same-Day**	**Primary Generalization**	**Max Dev.**	**Memory**
Statistical (Ch. 4)	99.6%[a]	52.1% cross-day[a]	10[a]	<1GB
Temporal (Ch. 5)	99.6%[b]	85.8% (100Hz Doppler)[b]	30[b]	Moderate
Cyclostat. (Ch. 6)	97.2%[c]	85.9% cross-Tx[c]	17[d]	Progressive
MADE. (Ch. 7)	99.9%[e]	92.9% cross-day[c]	80[f]	0.12GB/dev

[a]GLOBECOM22 (10 WiFi/Bluetooth devices). [b]LoRa-60 (30 LoRa devices). [c]UAV-8 (8 UAV controllers). [d]UAV-17 (17 UAV controllers). [e]WiSIG RX77 (6 WiFi devices, optimal signal conditions). [f]WiSIG ManyTx (80 WiFi transmitters).

Note: Same-day accuracy reflects controlled conditions; primary generalization reflects each method's targeted challenge (temporal drift, mobility, cross-transmission, or scalability).

The performance comparison reveals distinct operational regimes where each approach excels. All approaches achieve excellent same-day performance (97.2–99.9%), validating RF fingerprinting viability under controlled conditions. However, each method targets different generalization challenges: statistical enhancement addresses cross-day temporal drift (52.1%), temporal modeling handles mobility-induced variations (85.8% at 100Hz Doppler), cyclostationary features enable cross-transmission robustness (85.9%), and the MADE achieves both cross-day generalization (92.9%) and scalability (75.2% at 80 devices).

A striking finding is the 28.4 percentage point performance range across WiSIG receiver configurations (71.5% RX11 $\rightarrow$ 99.9% RX77), which demonstrates that signal quality impacts performance more substantially than architectural choices for identical device populations. This signal quality de-

pendence suggests that practical deployments must address signal processing enhancement alongside architectural optimization. We provide detailed analysis of each method's contributions and the insights driving these performance differences in Section 8.3.

The MADE shows superior cross-day robustness (92.9%), representing 78% improvement over statistical enhancement (52.1%). Our results suggest that embedding-based architectures with reconstruction learning develop more stable representations than statistical aggregation for temporal adaptation. Cyclostationary features reach 85.9% cross-transmission accuracy by capturing hardware-invariant periodic patterns where Gaussian noise contribution approaches zero, supporting discrimination despite transmission parameter variations. Temporal modeling maintains 85.8% accuracy under extreme mobility (100Hz Doppler), with bidirectional LSTM and attention mechanisms distinguishing mobility-induced transient variations from hardware-specific persistent characteristics. For scalability, the MADE reaches 75.2% at 80 devices with parameter-efficient architecture where the encoder maintains constant complexity, while classification approaches show substantial degradation beyond 20 devices. The non-monotonic scaling pattern (80.2% at 5 devices $\rightarrow$ 72.3% at 20 $\rightarrow$ 76.5% at 40 $\rightarrow$ 75.2% at 80) suggests larger populations provide better embedding space coverage despite the increased discrimination challenge. Statistical enhancement occupies the efficiency frontier with <5ms latency and <1GB memory, enabling CPU-only edge deployment, while the MADE requires infrastructure support but maintains consistent inference latency (12–15ms) across device populations through efficient GPU parallelization of the $O(N)$ classification layer.

We note that the trade-off between accuracy and computational efficiency reflects tension between model capacity (for complex pattern recognition) and resource constraints (limiting deployment contexts). Approaches achieving superior generalization require more extensive training data and computational resources, while methods optimized for controlled scenarios enable rapid deployment with minimal requirements.

8.2.2 Evolution of Approaches and Progressive Problem Solving

The four technical chapters represent progression where each method addresses limitations identified in previous approaches. We provide a brief overview here; detailed analysis of contributions and key insights appears in Section 8.3.

Statistical → Temporal (Chapters 4 → 5): Statistical enhancement achieved 99.6% same-day accuracy but suffered 47.5 percentage point cross-day degradation (52.1%), revealing that static feature aggregation lacks mechanisms to model temporal dependencies. Temporal modeling addressed this through CNN-LSTM-Attention architectures, achieving 85.8% accuracy under extreme mobility (100Hz Doppler) by explicitly capturing temporal dependencies that distinguish systematic trends from transient disturbances.

Temporal → Cyclostationary (Chapters 5 → 6): While temporal modeling excelled at mobility, it lacked cross-transmission evaluation. Cyclostationary feature engineering reached 85.9% cross-transmission accuracy by capturing hardware-specific periodic patterns in higher-order statistical spaces where Gaussian noise contribution approaches zero. Progressive learning achieved 38% memory reduction while maintaining competitive performance.

Cyclostationary → MADE (Chapters 6 → 7): Classification architecture scalability limitations (performance degradation from 85.9% at 8 devices to 70–80% at 17 devices) motivated the shift to embedding-based learning. The MADE achieved 75.2% at 80 devices with consistent 12–15ms inference latency, while the dual training paradigm combining reconstruction with classification enabled 92.9% cross-day accuracy, a 78% improvement over statistical enhancement.

8.3 Research Contributions and Achievements

We map in this section our research contributions to thesis objectives and research questions, demonstrating achievement through complementary technical innovations while acknowledging limitations requiring future work.

8.3.1 Achievement of Research Questions

RQ1: Cross-Transmission Parameter Robustness

How can RF fingerprinting achieve robustness to transmission parameter variations while maintaining hardware-specific discriminability?

Achievement: In our evaluation, cyclostationary feature engineering with progressive learning reaches 85.9% accuracy across systematic transmission parameter variations (Tx1–2 train → Tx3–19 test), representing 12.5 percentage point improvement over CNN-LSTM baseline (73.4%).

Key Insight: RF signals exhibit non-linearity and stochasticity that differ from images or text in ways that matter for feature extraction. Moving to

higher-order cyclostationary domains (2nd, 4th, 6th order features) captures hardware-specific periodic patterns arising from oscillator instabilities, power amplifier nonlinearities, and I/Q imbalance. At 4th and 6th orders, Gaussian noise contribution approaches zero (white Gaussian noise has no cyclostationary structure beyond 2nd order), allowing extraction of transmission-parameter-invariant characteristics where first-order features fail because noise dominates discriminative signal characteristics.

We integrated signal processing with deep learning through custom neural network layers implementing cyclostationary operations as differentiable PyTorch modules, enabling end-to-end gradient-based optimization while maintaining theoretical grounding. Progressive learning via Fisher information-based importance weighting enables incremental device addition with 38% memory reduction (1471MB vs. 2390MB) while maintaining competitive performance (83.8% vs. 85.9%).

Limitation: Scalability degradation from 85.9% (8 devices) to 70–80% (17 devices) indicates classification-based architectures face $O(N)$ complexity constraints despite sophisticated feature engineering.

RQ2: Temporal Adaptation and Cross-Day Robustness

How can systems maintain identification accuracy under temporal evolution of device characteristics caused by aging and environmental drift?

Achievement: Our MADE reaches 92.9% cross-day accuracy, representing 78% improvement over statistical enhancement (52.1%). The dual training paradigm combining masked autoencoder reconstruction with device classification enables learning of stable representations that remain consistent despite temporal evolution.

Key Insight: Embedding-based representation learning with reconstruction objectives provides more effective temporal adaptation than statistical feature aggregation. The reconstruction task forces models to learn compressed representations capturing signal structure rather than exploiting dataset-specific artifacts. Transformer self-attention mechanisms implicitly emphasize temporally stable signal characteristics (persistent hardware imperfections) while de-emphasizing time-varying environmental factors (transient channel conditions, brief interference).

Embedding-based architectures learn continuous representation spaces where device identities correspond to clusters, with similarity metrics determining identification. As characteristics evolve gradually, embedding representations shift continuously, maintaining identification capability as long as devices remain closer to their own cluster centroids than competitors. Classification

boundaries create hard decision regions where small shifts cause misclassification, while embeddings enable graceful adaptation.

Remaining Gap: The 7.0 percentage point gap between same-day (99.9%) and cross-day (92.9%) indicates temporal drift remains challenging. Long-term adaptation (months or years) requires continual learning mechanisms that can distinguish legitimate evolution from adversarial manipulation, a security-critical problem we have not fully addressed.

RQ3: Mobility and Dynamic Channel Handling

How can RF fingerprinting handle mobile scenarios with Doppler effects, fading, and time-varying channel conditions?

Achievement: Our CNN-LSTM-Attention architecture maintains 85.8% accuracy under extreme mobility (100 Hz Doppler, equivalent to 125 km/h at 868 MHz), representing 12.6 percentage point improvement over approaches without Doppler-aware augmentation (73.2%).

Key Insight: Bidirectional LSTM explicitly captures temporal dependencies, enabling distinction between mobility-induced variations (environmental, transient) and hardware-specific characteristics (device-dependent, persistent). Multi-head attention dynamically focuses on discriminative temporal segments less affected by mobility-induced fading and interference. Doppler-aware data augmentation synthesizing diverse mobility conditions during training enables generalization to unseen mobility scenarios.

The progressive degradation pattern validates robustness: 99.6% stationary $\rightarrow$ 93.1% LOS mobile $\rightarrow$ 87.6% NLOS mobile $\rightarrow$ 85.8% extreme mobility demonstrates graceful decline rather than catastrophic failure, indicating the architecture successfully separates mobility effects from hardware fingerprints.

Limitation: We conducted the evaluation on LoRa communications at 868 MHz. Performance under extreme mobility for other protocols (WiFi at 2.4/5.8 GHz, UAV controllers with proprietary waveforms) requires validation.

RQ4: Scalability to Large Device Populations

How can systems scale to large device populations while maintaining discrimination capability and computational efficiency?

Achievement: Our MADE reaches 75.2% accuracy at 80 devices with only 5.0 percentage point degradation from the 5-device baseline (80.2%). The parameter-efficient architecture maintains constant encoder complexity with consistent 12.3ms inference latency across device populations, as encoding time dominates over the $O(N)$ classification layer for practical deployments.

Key Insight: The transition from classification to embedding-based learning improves scalability characteristics in a way that matters for deployment. Classification approaches exhibit $O(N)$ output layer growth requiring N classification neurons, creating training instability with limited samples per class and increasing inference computation proportionally. Our MADE concentrates model capacity in a shared encoder with constant complexity, projecting inputs to fixed-dimensional continuous space (256 dimensions). While similarity-based identification still requires $O(N)$ comparisons against enrolled devices, the dominant computational cost lies in encoding, which remains constant regardless of device population.

The non-monotonic scaling pattern reveals architectural resilience: performance drops to 72.3% (20 devices), recovers to 76.5% (40 devices), then stabilizes at 75.2% (80 devices). Recovery at larger scales suggests that bigger populations provide better embedding space coverage enabling more robust decision boundaries despite increased discrimination complexity. Resource scaling proves predictable: GPU memory exhibits linear growth (0.12 GB per device), training time scales sub-linearly ($O(n^{0.68})$), and inference latency remains constant.

Limitation: Our evaluation stops at 80 devices due to proper data availability. Scalability to hundreds or thousands of devices requires further validation. Open-set recognition (detecting unknown devices) receives limited evaluation despite architectural support that provides foundation through reconstruction error analysis.

RQ5: Computational Efficiency for Practical Deployment

How can sophisticated RF fingerprinting operate within computational constraints of resource-limited deployments?

Achievement: Statistical enhancement establishes the efficiency frontier (<5ms latency, <1GB memory, CPU-only), and the progressive learning yields 38% memory reduction while maintaining competitive performance. The MADE provides parameter-efficient scaling with constant encoder complexity and maintaining consistent inference latency (12–15ms) across device populations from 5 to 80 devices.

Key Insight: Computational efficiency exists on a spectrum reflecting trade-offs between capability and resource requirements, the statistical enhancement allows edge deployment on resource-constrained devices (Raspberry Pi, embedded processors) but couples with limited cross-day generalization (52.1%). In the other hand, progressive learning demonstrates efficiency optimization within sophisticated approaches through continual learning re-

ducing memory requirements. The MADE's parameter-efficient architecture concentrates computational cost in shared encoding layers, meaning that while the classification layer scales as $O(N)$, the dominant encoding cost remains constant, allowing consistent latency as populations grow and allowing computational investment to amortize across large deployments.

8.4 Security Impact and Deployment Implications

The research questions addressed in Section 8.3 target technical generalization challenges, but the underlying motivation is operational: providing physical-layer security for wireless systems where credential-based authentication is insufficient. We position the contributions of this thesis with respect to two concrete security capabilities that the improved generalization makes practical.

8.4.1 Physical-Layer Device Authentication

RF fingerprinting complements rather than replaces cryptographic authentication, establishing a defense-in-depth architecture where an adversary must simultaneously defeat both credential-based and hardware-based verification to impersonate a legitimate device. Cryptography verifies identity claims at the protocol level; RF fingerprinting verifies that the transmitting hardware matches the claimed identity at the physical level. Because hardware imperfections are intrinsic to the device and cannot be modified through software manipulation, this layer is robust to credential theft, key extraction, and protocol-level replay, attack vectors against which cryptographic mechanisms alone provide limited protection.

The deployment barriers that have historically prevented this capability from reaching operational use, temporal drift, transmission parameter variation, mobility effects, and population scalability, are precisely the challenges addressed by the four contributions of this thesis. In practical terms, this means authentication that remains reliable as devices age, that persists across frequency-hopping and adaptive power control, that extends to mobile UAV and vehicular platforms, and that scales to device populations typical of IoT deployments rather than small laboratory setups. The 12–15 ms inference latency achieved by the MADE architecture makes this verification operationally feasible for real-time authentication without introducing unacceptable communication overhead.

8.4.2 Spectrum Monitoring and Unauthorized Transmitter Detection

Beyond verifying known devices, the reconstruction-based learning developed in Chapter 7 provides an architectural foundation for detecting unauthorized transmitters operating in monitored frequency bands. The MADE encoder learns a compact representation of legitimate device signals during enrollment; transmitters whose hardware characteristics were never observed in training produce elevated reconstruction error that can be used as an anomaly signal. This approach is passive, protocol-agnostic, and does not require prior knowledge of the unauthorized device, properties directly relevant to spectrum management in licensed bands, protection of UAV command-and-control links, and monitoring of restricted airspace.

We emphasize that this thesis validates the *architectural feasibility* of reconstruction-based anomaly detection rather than a deployment-ready spectrum monitoring system. Full evaluation requires registry-based open-set protocols with calibrated false acceptance and false rejection rates across diverse adversary models, which we identify as future work in Section 8.5. The cyclostationary features from Chapter 6 additionally strengthen this capability by enabling detection even when adversaries attempt to evade identification through transmission parameter manipulation, since higher-order statistical signatures remain device-specific under such variations.

8.4.3 Scope of Security Claims

Consistent with the threat model established in Chapter 1, the security contributions of this thesis address device impersonation, unauthorized network access, transmission parameter spoofing, and temporal drift exploitation. They do not address hardware-level device cloning, adversarial signal crafting, or systematic replay attack evaluation, these remain open problems that require dedicated adversarial robustness methodology beyond the generalization focus of this work, as discussed in Section 8.5. The value delivered is an authentication and monitoring layer whose generalization properties have been validated across four protocols and six datasets, providing a defensible foundation on which adversarial robustness evaluation and operational integration can be built.

8.5 Limitations and Future Research Directions

8.5.1 Current Limitations

Despite significant progress, several limitations constrain immediate deployment:

1. Persistent Cross-Day Performance Gap. The 7.0 percentage point gap between same-day (99.9%) and cross-day (92.9%) performance reveals that temporal evolution of device characteristics remains incompletely addressed. The challenge involves distinguishing legitimate device evolution (gradual characteristic drift requiring model adaptation) from adversarial manipulation (spoofing attempts requiring rejection). Conservative approaches risk rejecting legitimate devices after characteristic evolution (false rejections impacting usability), while permissive adaptation enables adversaries to gradually shift device characteristics toward target profiles. Our current work provides no principled solution to this trade-off, instead implicitly assuming temporal evolution remains sufficiently gradual that periodic retraining suffices.

2. Incomplete Multi-Dimensional Robustness. No single approach achieves optimal performance across all deployment dimensions simultaneously. Temporal modeling excels at mobility (85.8% at 100Hz Doppler) but lacks cross-transmission evaluation. Cyclostationary approaches achieve cross-transmission robustness (85.9%) but show limited scalability (degradation to 70–80% at 17 devices). The MADE scales effectively (75.2% at 80 devices) and shows cross-day robustness (92.9%) but lacks mobility and cross-device assessment. Real-world deployments may encounter temporal drift AND mobility AND parameter variations simultaneously, creating compound generalization challenges that potentially exceed individual challenge difficulty.

3. Limited Evaluation at Extreme Scales. The MADE demonstrates scalability to 80 devices, representing clear advance, but IoT deployments may involve hundreds or thousands of device types. The non-monotonic performance pattern (recovery at 40 devices) suggests continued scaling may be viable, but evaluation at 200+ device scales requires larger datasets.

4. Signal Quality Dependence. The 28.4 percentage point performance range across receiver configurations (71.5% RX11 $\rightarrow$ 99.9% RX77) demonstrates that signal quality impacts performance more substantially than architectural innovations. Practical deployments in uncontrolled environments encounter variable interference, multipath fading, weather effects, and other environmental factors. While robust architectures minimize degradation under challenging conditions, absolute performance remains constrained by avail-

able signal quality. Achieving consistently high performance requires signal processing enhancements such as advanced interference cancellation and multipath mitigation alongside architectural improvements.

5. Adversarial Robustness Gap. Our current work assumes non-adversarial scenarios. Recent research reveals that CNN-based RF fingerprinting exhibits consistent misclassification behavior under domain shifts exploitable as adversarial "backdoors," with confidence thresholding providing insufficient protection (rejection rates below 22% even at 95% confidence) [87, 88]. Models trained on raw IQ samples entangle RF fingerprints with environmental and signal-pattern features, creating attack vectors where adversaries manipulate environmental conditions to induce misclassification without sophisticated RF hardware spoofing. Systematic evaluation of sophisticated attacks, adversarial training, and certified defenses remains incomplete despite representing a critical requirement for security applications.

8.5.2 Future Research Directions

Foundation Models for RF Signals

The MADE reconstruction learning provides a foundation for pre-training on massive unlabeled RF datasets, analogous to foundation models in natural language processing and computer vision. Recent work on Large Wireless Models [89, 90] shows transformer-based pre-training on 1M+ wireless channels from DeepMIMO, achieving improvements in beam prediction, LoS/NLoS classification, and localization with limited training data. Similarly, WiFo [91] applies masked autoencoder pre-training to channel prediction tasks.

We argue that future research should extend these approaches to RF fingerprinting through pre-training on diverse unlabeled wireless signals spanning protocols, frequency bands, and device types. Zero-shot or few-shot adaptation to new device types through transfer learning could dramatically reduce training data requirements (10 to 100 samples rather than thousands). Also, unified representations supporting multiple downstream tasks such as fingerprinting, modulation classification and interference detection would enable wireless intelligence systems applicable across deployment scenarios.

Adversarial Robustness and Security Analysis

Recent investigation of DL-based RF fingerprinting security [92] reveals critical vulnerabilities. For instance, CNN models exhibit consistent misclassification behavior under domain shifts that adversaries can exploit as "backdoors"

for impersonation attacks [93]. Future work must address sophisticated adversarial attacks through robust spoofing evaluation, adversarial training incorporating attack examples, and certified defenses providing provable security guarantees. Research on domain-invariant representation learning and contrastive learning approaches show promise for improving robustness to domain shift [94].

Multi-Objective and Open-Set Evaluation

Our current approaches specialize on individual challenges. Future research should develop architectures explicitly addressing multiple generalization dimensions simultaneously through multi-task learning with explicit loss terms for each dimension such as temporal stability, and mobility robustness, recent work now looking at modular architectures with switchable components that can adapt to deployment conditions, and ensemble methods combining complementary specialized models [95, 96, 97]. Also, future work needs to be focus on evaluation beyond 100+ devices in order to understand scalability limits. This requires acquiring or generating larger datasets with diverse device populations, open-set recognition evaluation with varying false acceptance and rejection rates [98], and investigation of hierarchical classification approaches where devices are first clustered by manufacturer before fine-grained identification [99, 100].

8.6 Conclusions

In this thesis we demonstrate that effective RF fingerprinting requires systematic integration of signal processing theory with deep learning architectures, where domain-specific knowledge guides feature design rather than relying solely on data-driven learning to discover appropriate representations. The central insight is that RF signals exhibit non-Gaussian characteristics fundamentally different from images or text: moving to higher-order cyclostationary domains (2nd, 4th, 6th order) captures hardware-specific patterns in spaces where Gaussian noise contribution approaches zero, enabling extraction of transmission-parameter-invariant characteristics invisible in conventional representations. This principle, validated across all approaches, yields consistent improvements (12.5 to 14.6 percentage points) over purely deep learning baselines, establishing that signal processing theory remains essential even in the deep learning era.

Our evaluation across four complementary approaches establishes deployment-viable performance: statistical enhancement achieves 99.6% same-day accu-

racy with 52.1% cross-day robustness, temporal modeling maintains 85.8% accuracy under extreme mobility (100 Hz Doppler), cyclostationary feature extraction reaches 85.9% cross-transmission generalization, and the MADE architecture scales to 80 devices with controlled 5% degradation while maintaining 92.9% cross-transmission accuracy. However, honest assessment reveals that several challenges remain unresolved within this thesis scope. The scalability achievements, while substantial, still face memory constraints at approximately 150 devices for typical GPU hardware, limiting immediate deployment at metropolitan IoT scales. The open-set recognition capabilities remain conceptual, reconstruction-based anomaly detection establishes architectural feasibility but still needs validation against sophisticated adversarial attacks essential for operational security deployment. Protocol diversity assessment, limited to WiFi and UAV controllers, leaves comprehensive multi-protocol validation (cellular, Bluetooth, LoRa, 5G/6G) as critical future work.

We have identified three concrete research directions. First, foundation model approaches through pre-training on massive unlabeled wireless signal collections could enable zero-shot adaptation to new protocols and device types, analogous to Large Language Models but adapted to RF signal characteristics. The dual training paradigm combining reconstruction with classification provides architectural template for such Large RF Models. Second, systematic adversarial robustness validation against state-of-the-art attacks represents essential prerequisite for operational security deployment, particularly given that dual-objective training may provide inherent defenses where adversarial examples optimized to fool classification exhibit high reconstruction error. Third, hierarchical architectures combining the MADE scalability with progressive learning from Chapter 6 could address both large-scale deployment and temporal adaptation requirements simultaneously.

As wireless connectivity expands to tens of billions of IoT devices in increasingly autonomous and security-critical applications, we think that physical-layer authentication provides essential capability that cannot be bypassed through software exploitation. Our thesis advances RF fingerprinting from controlled laboratory validation toward practical deployment feasibility, establishing both what is achievable with current techniques and defining the specific technical barriers, especially scalability beyond 100 devices, adversarial robustness validation, protocol diversity and unified continual learning, that the research community must address for operational deployment at scale.

Bibliography

[1] Vikas Hassija, Vinay Chamola, Adhar Agrawal, Adit Goyal, Nguyen Cong Luong, Dusit Niyato, Fei Richard Yu, and Mohsen Guizani. Fast, reliable, and secure drone communication: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 23(4):2802–2832, 2021. Comprehensive survey on UAV security including GPS spoofing, authentication vulnerabilities.

[2] Zhaoxuan Wang, Yang Li, Shihao Wu, Yuan Zhou, Libin Yang, Yuan Xu, Tianwei Zhang, and Quan Pan. A survey on cybersecurity attacks and defenses for unmanned aerial systems. *Journal of Systems Architecture*, 138:102870, 2023. Documents UAV swarm vulnerabilities and credential-based attacks.

[3] N. H. Motlagh, T. Taleb, and O. Arouk. Low-altitude unmanned aerial vehicles-based internet of things services: Comprehensive survey and future perspectives. *IEEE Internet of Things Journal*, 3(6):899–922, 2016.

[4] The Cyber Express. The game of drones of hovering cybersecurity risks, March 2024. Documents October 2022 DJI drone attack on US financial company using Wi-Fi Pineapple devices.

[5] IoT Analytics. State of iot 2025: Number of connected iot devices growing 14% to 21.1 billion globally, May 2025. Reports 21.1 billion IoT devices in 2025, 39 billion forecast for 2030.

[6] Transforma Insights and Exploding Topics. Number of internet of things (iot) connections worldwide from 2022 to 2034 (in billions), May 2025. Forecasts 19.8 billion devices in 2025, 40.6 billion by 2034.

[7] Nataliia Neshenko, Elias Bou-Harb, Jorge Crichigno, Georges Kaddoum, and Nasir Ghani. Demystifying iot security: An exhaustive survey on iot vulnerabilities and a first empirical look on internet-scale iot exploitations. *IEEE Communications Surveys & Tutorials*, 21(3):2702–2733, 2023.

[8] R. Altawy and A. M. Youssef. Security, privacy, and safety aspects of civilian drones: A survey. *ACM Computing Surveys*, 49(2):1–33, 2016.

[9] Xiaohui Gao, Chang Shan, Changzhen Hu, Zhaowen Niu, and Zihan Liu. An adaptive ensemble machine learning model for intrusion detection. *IEEE Access*, 7:82512–82521, 2023.

[10] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan. Survey on uav cellular communications: Practical aspects, standardization advancements, regulation, and security challenges. *IEEE Communications Surveys & Tutorials*, 21(4):3417–3442, 2019.

[11] Anu Jagannath, Zackary Kane, and Jithin Jagannath. Rf fingerprinting needs attention: Multi-task approach for real-world wifi and bluetooth. In *Proc. of IEEE Global Communications Conference (GLOBECOM)*, Rio de Janeiro, Brazil, December 2022. IEEE.

[12] Y. Zou, J. Zhu, X. Wang, and V. C. M. Leung. Physical layer security in wireless communications: A tutorial. *IEEE Wireless Communications*, 23(4):16–26, Aug 2016.

[13] S. Li, Y. Zhu, J. Jiang, and J. Zhang. A review of physical layer security techniques for internet of things: Challenges and solutions. *IEEE Access*, 8:20341–20356, 2020.

[14] Kevin Merchant, Shauna Revay, George Stantchev, and Bryan Nousain. Deep learning for rf device fingerprinting in cognitive communication systems. *IEEE Journal of Selected Topics in Signal Processing*, 12(1):160–167, 2018.

[15] Dana R. Reising, Michael A. Temple, and Jillian A. Jackson. Authorized and rogue device discrimination using dimensionally reduced rf-dna fingerprints. *IEEE Transactions on Information Forensics and Security*, 10(6):1180–1192, 2015.

[16] Nasim Soltanieh, Yaser Norouzi, Yang Yang, and Nemai Chandra Karmakar. A review of radio frequency fingerprinting techniques. *IEEE Journal of Radio Frequency Identification*, 4(3):222–233, 2020.

[17] Kunal Sankhe, Mauro Belgiovine, Fan Zhou, Shamnaz Riyaz, Stratis Ioannidis, and Kaushik Chowdhury. No radio left behind: Radio fingerprinting through deep learning of physical-layer hardware impairments. volume 6, pages 165–178, 2020.

[18] A. Ahmed, B. Quoitin, A. Gros, and V. Moeyaert. A comprehensive survey on deep learning-based lora radio frequency fingerprinting identification. *Sensors*, 24(13):4411, 2024.

[19] L. Xie, L. Peng, J. Zhang, and A. Hu. Radio frequency fingerprint identification for internet of things: A survey. *Security and Safety*, 2024.

[20] Le Ding, Shuai Wang, Feng Wang, and Wei Zhang. Specific emitter identification via convolutional neural networks. *IEEE Communications Letters*, 22(12):2591–2594, 2020.

[21] Jiabao Yu, Aiqun Hu, Guyue Li, and Linning Peng. A robust rf fingerprinting approach using multisampling convolutional neural network. *IEEE Internet of Things Journal*, 6(4):6786–6799, 2019.

[22] Tao Jian, Bernardo C. Rendon, Emmanuel Ojuba, Noor Soltani, Zifeng Wang, Kunal Sankhe, Andrey Gritsenko, Jennifer Dy, Kaushik Chowdhury, and Stratis Ioannidis. Deep learning for rf fingerprinting: A massive experimental study. *IEEE Internet of Things Magazine*, 3(1):50–57, 2021.

[23] Samer Hanna, Shengli Yan, and Danijela Cabric. Open set wireless transmitter authorization: Deep learning approaches and dataset considerations. *IEEE Transactions on Cognitive Communications and Networking*, 7(2):59–72, 2021.

[24] S. Abbas, M. A. Talib, Q. Nasir, S. Idhis, M. Alaboudi, and A. Mohamed. Radio frequency fingerprinting techniques for device identification: A survey. *International Journal of Information Security*, 2024.

[25] R. Meng, B. Xu, X. Xu, M. Sun, B. Wang, S. Han, S. Lv, and P. Zhang. A survey of machine learning-based physical-layer authentication in wireless communications. *Journal of Network and Computer Applications*, 235:104085, 2025.

[26] Z. Lai, Z. Chang, M. Sha, Q. Zhang, N. Xie, C. Chen, and D. Niyato. A comprehensive survey on physical layer authentication techniques: Categorization and analysis of model-driven and data-driven approaches. *ACM Computing Surveys*, 57(5):117, Jan 2025.

[27] B. Johnson and B. Hamdaoui. On the domain generalizability of rf fingerprints through multifractal dimension representation. *Sensors*, 22(11):4045, 2022.

[28] G. Chi, Z. Yang, C. Wu, J. Xu, Y. Gao, Y. Liu, and T. X. Han. Rf-diffusion: Radio signal generation via time-frequency diffusion. In *Proc. 30th Annual International Conference on Mobile Computing and Networking (ACM MobiCom)*, pages 77–92, Washington D.C., DC, USA, 2024.

[29] Eng Gee Li, Xueheng Geng, Jehyeok Lee, Youngbin Jung, and Donggyun Lee. A review on deep learning for edge intelligence in 6g networks. *IEEE Access*, 9:79366–79396, 2021.

[30] N. Quadar, A. Chehri, and B. Debaque. Advanced security frameworks for uav and iot: A deep learning approach. *Internet of Things*, 32:101594, 2025.

[31] Nordine Quadar, Abdelah Chehri, and Benoit Debaque. Robust rf fingerprinting for lora iot devices in mobile scenarios using cnn-lstm-attention. In *2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring)*, pages 1–5, 2025.

[32] N. Quadar, A. Chehri, and B. Debaque. Integrating sensing, communication, and computing for robust rf fingerprinting in consumer electronics. *IEEE Transactions On Consumer Electronics*, 2026. under review.

[33] N. Quadar, A. Chehri, and B. Debaque. Scalable deep learning for rf fingerprinting: The made architecture for robust physical-layer device identification. *IEEE Open Journal of the Communications Society*, Feb, 2026. to appear.

[34] N. Quadar, A. Chehri, B. Debaque, H.Yanikomeroglu, and GK. Kurt. Unseen signals, real threats: A comprehensive survey of ai generalization in rf fingerprinting for wireless security. *IEEE Open Journal of the Communications Society*, 2026. under review.

[35] Nordine Quadar, Abdellah Chehri, and Benoit Debaque. Feature learning and continual adaptation for robust RF fingerprinting in UAV networks. In *Proc. IEEE International Conference on Communications (ICC)*, Glasgow, Scotland, UK, May 2026. To appear.

[36] Nordine Quadar, Abdellah Chehri, and Benoit Debaque. Leveraging cyclostationary features and continual learning for robust RF fingerprinting in IoT and UAV networks. In *Proc. IEEE 103rd Vehicular Technology Conference (VTC2026-Spring)*, Nice, France, June 2026. To appear.

[37] Nordine Quadar, Abdellah Chehri, and Benoit Debaque. Wireless security and iot device identification using rf fingerprinting and deep learning. In *2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)*, pages 1–5, 2024.

[38] Nordine Quadar, Abdelah Chehri, and Benoit Debaque. Audio recognition-based method for rf transmitters classification using cnn-lstm model. pages 464–468, 2025.

[39] Nordine Quadar, Abdellah Chehri, and Benoit Debaque. Tinyml dataset challenges in enabling scalable intelligence for 6g consumer electronics. *IEEE Consumer Electronics Magazine*, pages 1–12, 2026.

[40] Nordine Quadar, Abdellah Chehri, and Benoit Debaque. Tinyml datasets as enablers of 6g edge intelligence: Key insights and research gaps. *IEEE Wireless Communications*, pages 1–10, 2026.

[41] Nordine Quadar, Mohamed Rahouti, Moussa Ayyash, Senthil Kumar Jagatheesaperumal, and Abdellah Chehri. Iot-ai/machine learning experimental testbeds: The missing piece. *IEEE Internet of Things Magazine*, 7(1):136–143, 2024.

[42] S. Riyaz, K. Sankhe, S. Ioannidis, and K. Chowdhury. Deep learning convolutional neural networks for radio identification. *IEEE Communications Magazine*, 56(9):146–152, 2018.

[43] Saeif Al-Hazbi, Ahmed Hussain, Savio Sciancalepore, Gabriele Oligeri, and Panos Papadimitratos. Radio frequency fingerprinting via deep learning: Challenges and opportunities. *2024 International Wireless Communications and Mobile Computing (IWCMC)*, pages 0824–0829, 2024.

[44] H. Li, C. Wang, N. Ghose, and B. Wang. Poster: Robust deep-learning-based radio fingerprinting with fine-tuning. In *Proc. ACM WiSec'21*, 2021. Dataset available at: https://github.com/SmartHomePrivacyProject/RadioFingerprinting.

[45] Sanjoy Basak, Sanjoy Rajendran, Sofie Pollin, and Bart Scheers. Combined rf-based drone detection and classification. *IEEE Transactions on Cognitive Communications and Networking*, 8(1):111–120, March 2022.

[46] Martins Ezuma, Fatih Erden, Chethan K. Anjinappa, Ozgur Ozdemir, and Ismail Guvenc. Drone remote controller rf signal dataset, November 2020. Dataset size: 124 GB, 17 controllers, 1000 signals per controller.

[47] Guanxiong Shen, Junqing Zhang, Alan Marshall, and Joseph Cavallaro. Towards scalable and channel-robust radio frequency fingerprint identification for lora. *IEEE Transactions on Information Forensics and Security*, 17:774–787, 2022.

[48] Samer Hanna et al. Wisig: Large-scale wifi signal dataset for receiver and channel agnostic rf fingerprinting. *IEEE Access*, 10:22809–22831, 2022. Available at: https://cores.ee.ucla.edu/downloads/datasets/wisig/.

[49] D. Raychaudhuri, I. Seskar, M. Ott, S. Ganu, K. Ramachandran, H. Kremo, R. Siracusa, H. Liu, and M. Singh. Overview of the orbit radio grid testbed for evaluation of next-generation wireless network protocols. In *IEEE Wireless Communications and Networking Conference, 2005*, volume 3, pages 1664–1669 Vol. 3, 2005.

[50] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury. Oracle: Optimized radio classification through convolutional neural networks. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 370–378, Paris, France, 2019.

[51] T. G. Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation*, 10(7):1895–1923, 1998.

[52] J. Yi, J. Chen, and L. Cheng. A deep LSTM-CNN based on self-attention mechanism with input data reduction for short-term load forecasting. *IET Generation, Transmission & Distribution*, 17(7):1642–1654, 2023. Analysis of computational complexity and efficiency optimization for CNN-LSTM hybrid architectures.

[53] H. Fu, H. Dong, J. Yin, and L. Peng. Radio frequency fingerprint identification for 5g mobile devices using dctf and deep learning. *Entropy*, 26(4):325, 2024.

[54] A. Jagannath, J. Jagannath, and P. S. P. V. Kumar. A comprehensive survey on radio frequency (rf) fingerprinting: Traditional approaches, deep learning, and open challenges. *Computer Networks*, 219:109455, 2022.

[55] A. Alshammary, C. Alwakid, M. Alhamid, and M. Alshehri. A comprehensive survey on radio frequency (rf) fingerprinting: Traditional approaches, deep learning, and open challenges. *Computer Networks*, 219, 2022. Comprehensive overview of RF fingerprinting including challenges like Doppler effects and mobility.

[56] N. Guerin, T. Gillard, and Y. Delignon. From modeling to sensing of micro-Doppler in radio communications. *Remote Sensing*, 14(24):6310, 2022. Geometric model for Doppler modulation effects in radio communications with applications to RF fingerprinting.

[57] M. Donno, K. Pham, L. Bonati, T. Melodia, M. Sciancalepore, and S. Basagni. Radio frequency fingerprinting via deep learning: Challenges and opportunities. arXiv preprint arXiv:2310.16406, 2024. Addresses channel modeling challenges including time-varying conditions, fading, multipath, and Doppler effects for mobile RF scenarios.

[58] A. Paulraj and C. B. Papadias. Space-time processing for wireless communications. *IEEE Signal Processing Magazine*, 14(6):49–83, Nov 1997. Comprehensive treatment of spatial and temporal signal processing for mobile wireless communications, including Doppler effects, fading, and delay spread.

[59] G. Shen, J. Zhang, A. Marshall, and J. R. Cavallaro. Towards scalable and ubiquitous millimeter-wave wireless networks. *IEEE Transactions on Mobile Computing*, 21(5):1551–1565, May 2022. Provides framework for modeling temporal variations in mobile RF environments.

[60] Y. Zhang, Z. Zhou, and X. Li. Specific emitter identification handling modulation variation with margin disparity discrepancy. *IEEE Transactions on Information Forensics and Security*, 2023.

[61] K. F. Haque, K. A. Aziz, M. R. Hasan, et al. Deep CNN-LSTM with self-attention model for human activity recognition using wearable sensor. *IEEE Journal of Biomedical and Health Informatics*, 26(7):2906–2913, 2022. Foundational work on CNN-LSTM-Attention hybrid architectures for temporal sequence modeling.

[62] S. Siami-Namini, N. Tavakoli, and A. S. Namin. The performance of LSTM and BiLSTM in forecasting time series. In *Proc. 2019 IEEE International Conference on Big Data (Big Data)*, pages 3285–3292, Los Angeles, CA, USA, Dec 2019. Theoretical analysis and comparison of bidirectional LSTM for temporal sequence modeling.

[63] D. Luan, S. Thompson, and J. Jiang. Attention based neural networks for wireless channel estimation. *IEEE Communications Letters*, 26(6):1424–1428, June 2022. Application of attention mechanisms to wireless signal processing and channel estimation.

[64] Y. Lin, H. Wang, and H. Zha. The technology of radio frequency fingerprint identification based on deep learning for 5g application. *Security and Safety*, 3, 2024. Discusses channel effects like Doppler and multipath in RF fingerprinting, proposing channel estimation to enhance robustness.

[65] Y. Liu, J. Wang, J. Li, S. Niu, and H. Song. Generative adversarial network-based rogue device identification using differential constellation trace figure. *EURASIP Journal on Wireless Communications and Networking*, 2021(72), 2021. GAN-based approach for rogue device identification, relevant for unknown device detection.

[66] Y. Huang, W. Xu, Z. Yang, and D. W. K. Ng. Lightweight radio frequency fingerprinting identification scheme for V2X based on temporal

correlation. *IEEE Transactions on Information Forensics and Security*, 19:4207–4220, 2024. Lightweight architecture design for scalable RF fingerprinting with temporal dependency modeling.

[67] William A. Gardner. The spectral correlation theory of cyclostationary time-series. *Signal Processing*, 11(1):13–36, 1986.

[68] William A. Gardner and Chad M. Spooner. The cumulant theory of cyclostationary time-series, part i: Foundation. *IEEE Transactions on Signal Processing*, 42(12):3387–3408, 1994.

[69] Jérôme Antoni. Cyclostationarity by examples. *Mechanical Systems and Signal Processing*, 23(4):987–1036, 2009.

[70] Chad M. Spooner and William A. Gardner. The cumulant theory of cyclostationary time-series, part ii: Development and applications. *IEEE Transactions on Signal Processing*, 42(12):3409–3429, 1994.

[71] L. Izzo and A. Napolitano. The higher-order theory of generalized almost cyclostationary time series. *IEEE Transactions on Signal Processing*, 46(11):2975–2989, 1998.

[72] Antonio Napolitano. Uncertainty in measurements on spectrally correlated stochastic processes. *IEEE Transactions on Information Theory*, 49(9):2172–2191, 2003.

[73] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017.

[74] J. A. Gutierrez del Arroyo, B. J. Borghetti, and M. A. Temple. Considerations for radio frequency fingerprinting across multiple frequency channels. *Sensors*, 22(6):2111, 2022.

[75] William A. Gardner, Antonio Napolitano, and Luigi Paura. Cyclostationarity: Half a century of research. *Signal Processing*, 86(4):639–697, 2006.

[76] Ala'a El-Habashna, Octavia A. Dobre, Ramachandran Venkatesan, and Dimitrie C. Popescu. Second-order cyclostationarity of mobile wimax and lte ofdm signals and application to spectrum awareness in cognitive radio systems. *IEEE Journal of Selected Topics in Signal Processing*, 6(1):26–42, 2012.

[77] Zhiheng Chen, Jasha Droppo, Jinyu Huang, and Dan Povey. Nonlinear feature extraction for robust speech recognition using deep neural networks. In *Interspeech 2020*, pages 3666–3670, 2020.

[78] John A. Snoap, Dimitrie C. Popescu, and Chad M. Spooner. Deep-learning-based classifier with custom feature-extraction layers for digitally modulated signals. *IEEE Transactions on Broadcasting*, 70(3):763–773, 2024.

[79] J. Xiao, H. Zhang, Z. Shao, Y. Zheng, and W. Ding. Progressive unsupervised domain adaptation for rf signal attribute recognition across communication scenarios. *Preprints*, 2024.

[80] Andrei A. Rusu, Neil C. Rabinowitz, Guillaume Desjardins, Hubert Soyer, James Kirkpatrick, Koray Kavukcuoglu, Razvan Pascanu, and Raia Hadsell. Progressive neural networks. *arXiv preprint arXiv:1606.04671*, 2016.

[81] Olusiji O. Medaiyese, Adrian P. Lauf, and Ismail Guvenc. A low complexity feature extraction for the rf fingerprinting process. In *2020 IEEE 6th World Forum on Internet of Things (WF-IoT)*, pages 1–6. IEEE, 2020.

[82] Rahaf Aljundi, Francesca Babiloni, Mohamed Elhoseiny, Marcus Rohrbach, and Tinne Tuytelaars. Memory aware synapses: Learning what (not) to forget. pages 144–161, 2018.

[83] Hanul Shin, Jung Kwon Lee, Jaehong Kim, and Jiwon Kim. Continual learning with deep generative replay. *Advances in Neural Information Processing Systems*, 30, 2017.

[84] Zexian Zhou and Xiaojing Liu. Masked autoencoders in computer vision: A comprehensive survey. *IEEE Access*, 11:113560–113579, 2023.

[85] Komal Bajaj, Dushyant Kumar Singh, and Mohd Aquib Ansari. Autoencoders based deep learner for image denoising. *Procedia Computer Science*, 171:1535–1541, 2020.

[86] Jiabao Yu, Aiqun Hu, Fen Zhou, Yuexiu Xing, Yi Yu, Guyue Li, and Linning Peng. Radio frequency fingerprint identification based on denoising autoencoders. In *2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 1–6. IEEE, 2019.

[87] Gaoli Yan, Xue Fu, Yu Wang, Qianyun Zhang, and Guan Gui. Radio frequency fingerprint identification towards statistical and deep learning

features: Review, recent results and future directions. *Peer-to-Peer Networking and Applications*, 18(3):1–25, 2025.

[88] Xinyu Cao, Bimal Adhikari, Shangqing Zhao, Jingxian Wu, and Yanjun Pan. An adversarial-driven experimental study on deep learning for rf fingerprinting. *arXiv preprint arXiv:2507.14109*, 2025.

[89] Sadjad Alikhani, Gouranga Charan, and Ahmed Alkhateeb. Large wireless model (LWM): A foundation model for wireless channels, 2024. World's first foundation model for wireless channels, transformer-based pre-training on 1M+ channels from DeepMIMO.

[90] Ahmed Aboulfotouh and Hatem Abou-Zeid. Multimodal wireless foundation models. *arXiv preprint arXiv:2511.15162*, 2025.

[91] Boxun Liu, Shijian Gao, Xuanyu Liu, Xiang Cheng, and Liuqing Yang. WiFo: Wireless foundation model for channel prediction. *Science China Information Sciences*, 68(6):162302, 2025. First space-time-frequency wireless foundation model with MAE-based pre-training for channel prediction.

[92] Xinyu Cao, Bimal Adhikari, Shangqing Zhao, Jingxian Wu, and Yanjun Pan. An adversarial-driven experimental study on deep learning for RF fingerprinting. In *Proc. IEEE Military Communications Conference (MILCOM)*. IEEE, 2025. First systematic security evaluation of DL-based RF fingerprinting, arXiv:2507.14109.

[93] Zhaoyi Lu, Ming Tu, Xin Xie, Wenchao Xu, Cunqing Hua, and Weihua Zhuang. Concealing radio frequency fingerprints via active adversarial perturbation. *IEEE Transactions on Network Science and Engineering*, 2025.

[94] Neil Houlsby, Andrei Giampouris, Alexander Ratner, Shuai Shen, Adina Taylor, Douglas Eck, Kyunghyun Cho, and Orhan Firat. Parameter-efficient transfer learning for nlp. pages 2790–2799, 2019.

[95] Qingyang Zhang, Yake Wei, Zongbo Han, Huazhu Fu, Xi Peng, Cheng Deng, Qinghua Hu, Cai Xu, Jie Wen, Di Hu, et al. Multimodal fusion on low-quality data: A comprehensive survey. *arXiv preprint arXiv:2404.18947*, 2024.

[96] Yuhao Pan, Xiucheng Wang, Nan Cheng, and Wenchao Xu. Cross-receiver generalization for rf fingerprint identification via feature disentanglement and adversarial training. *arXiv preprint arXiv:2510.09405*, 2025.

[97] Bisma Manzoor and Akram Al-Hourani. Multimodal rf fingerprinting for iot devices in satellite-based sensing. *IEEE Journal of Radio Frequency Identification*, 2025.

[98] Samurdhi Karunaratne, Samer Hanna, and Danijela Cabric. Open set rf fingerprinting using generative outlier augmentation. In *2021 IEEE Global Communications Conference (GLOBECOM)*, pages 01–07. IEEE, 2021.

[99] Bo Li and Ediz Cetin. Enhancing iot security with radio frequency fingerprinting: Traditional and deep learning-based approaches. In *Advances in the Internet of Things*, pages 56–77. CRC Press, 2025.

[100] Jian Yang, Shuai Feng, Yatong Wang, Xinghang Wu, and Mu Yan. Openrfi: Open-set radio frequency fingerprint identification via test-time fine-tuning. *IEEE Transactions on Mobile Computing*, 2025.