Securing Wi-Fi Networks: A Study on RF Fingerprinting and CNN-Based Intrusion Detection

by

K. M. Sazid Hasan (180021108)

Mohammed Shadman Sakib (180021302)

Shahriar Hassan (180021326)

BACHELOR OF SCIENCE

IN

ELECTRICAL AND ELECTRONIC ENGINEERING

Department of Electrical and Electronic Engineering

Islamic University of Technology (IUT)

Board Bazar, Gazipur-1704, Bangladesh.

June, 2023

CERTIFICATE OF APPROVAL

The thesis titled, "**Securing Wi-Fi Networks: A Study on RF Fingerprinting and CNN-Based Intrusion Detection**" accepted as partial fulfillment of the requirement for the Degree of BACHELOR OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING of Islamic University of Technology (IUT).

Approved by:

_____

Dr. Khondokar Habibul Kabir
(Supervisor)
Professor,
Department of Electrical and Electronic Engineering,
Islamic University of Technology (IUT),
Boardbazar, Gazipur-1704, Bangladesh.

Declaration of Candidate

It is hereby declared that this thesis report is only submitted to The Electrical and Electronic Engineering Department. Any part of it has not been submitted elsewhere for the award of any Degree or Diploma.

_____

K. M. Sazid Hasan

Student ID: 180021108

Date: _____

_____

Mohammed Shadman Sakib

Student ID: 180021302

Date: _____

_____

Shahriar Hassan

Student ID: 180021326

Date: _____

Table of Contents

List of Figures

List of Acronyms

AM	Acknowledged Mode
APN	Access Point Name
RF	Radio Frequency
ISM	Industrial, Scientific & Medical
SDR	Software Defined Radio
IQ	In-phase and Quadrature
IHMRF	Infinite Hidden Markov Random Field
IDS	Intrusion Detection System
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
SVM	Support Vector Machine
SAE	Stacked Auto Encoder
IoT	Internet of Things
CNN	Convolutional Neural Network
CSI	Channel State Information
STFT	Short Term Fourier Transform
USRP	Universal Software Radio Peripheral
API	Application Programming Interface
CFO	Carrier Frequency Offset
LoRa	Long Range
RFFI	Radio Frequency Fingerprint Identification
LO	Local Oscillator
IMRR	Image Rejection Ration
COTS	Commercial Off-The-Shelf
BER	Bit Error Rate
SNR	Signal-to-Noise Ratio

Acknowledgements

In the name of Allah, the Most Gracious, the Most Merciful. Before we commence, we would like to express our deep gratitude and highest praises to Almighty Allah, who has bestowed upon us the strength, knowledge, and ability to embark on this journey of intellectual exploration. It is His guidance and boundless blessings that have granted us the capacity to endure challenges and achieve this milestone. This work is a testimony to the truth that with faith, perseverance, and hard work, one can accomplish what they aspire to.

Foremost, we are profoundly grateful to our esteemed supervisor, Dr. Khondokar Habibul Kabir, whose sagacity, perspicuity, and unyielding support have been the guiding beacons throughout this journey. His deep insights into the subject matter, coupled with his fervor for imparting knowledge, have profoundly enriched our academic experience. His mentorship, marked by patience, optimism, and an unwavering belief in our capabilities, has been instrumental in the completion of this thesis. To him, we extend our heartfelt gratitude for his valuable contributions.

We owe an immense debt of gratitude to our families, whose unwavering support and prayers have been our constant source of strength. To our parents, who instilled in us the love for knowledge, nurtured our curiosity, and reinforced our determination, we express our deepest appreciation. Their sacrifices, encouragement, and enduring faith in our abilities have been the cornerstone of our achievements.

Lastly, we express our heartfelt thanks to our friends, who have shared in our journey, providing moral support, constructive criticism, and invaluable camaraderie. Their trust and encouragement, particularly during challenging periods, were vital in maintaining our morale and enthusiasm.

This accomplishment would not have been possible without the amalgamation of all these efforts and influences. While these acknowledgements do not fully express the depth of our gratitude, we wish to iterate our sincerest thanks to everyone who has been a part of our academic journey. May the Almighty Allah bless you all.

2023

Abstract

Advancements in wireless communication technology not only enhanced seamless connectivity and information exchange but also instigated the intrusion in RF networks, a pivotal challenge to the security of wireless communication networks. This thesis introduces a paradigm shift in wireless communication systems, to prevent unauthorized access attempts and malicious activities by classifying radio signals using Convolutional Neural Networks (CNNs). Diverging from conventional methods, an end-to-end deep learning model is proposed, capable of direct learning from raw time-domain signals, thereby preventing the requirement for manual feature engineering. The model is designed to extract and utilize rich, hierarchical feature representations from various radio signal types with diverse modulation techniques. Tested against a comprehensive radio signal dataset, the model demonstrates significantly enhanced classification accuracy and impressive generalization capabilities with unseen signals. The study explores the influence of different optimization algorithms on model performance, revealing how strategic parameter tuning can improve computational efficiency without compromising classification accuracy. The research not only advances the use of deep learning in radio signal classification, but also lays a foundation for future studies examining CNNs' noise resilience, interference handling, and adaptability to more intricate signals, thereby fostering the evolution of intelligent, autonomous systems in signal processing.

Chapter 1

Introduction

Over the past few years, the realm of wireless communications has undergone a substantial metamorphosis, propelled by the rapid progression of technological innovations. This evolving landscape has given rise to a multitude of complex requirements, one of the most critical being the accurate identification and classification of radio signals[1]. The precision of this process plays a cardinal role in a spectrum of applications, extending from intricate spectrum management tasks to the establishment and maintenance of secure communications systems.

One innovative solution that has emerged to enhance the accuracy and reliability of this process is radio fingerprinting. This unique approach seeks to identify the singular characteristics of individual radio devices, thus enabling their differentiation [2]. By identifying these unique characteristics, which can be thought of as the 'fingerprints' of a radio device, it becomes possible to distinguish one device from another with high precision. This technique is particularly relevant in modern communication systems, where the ability to identify and track devices is of paramount importance for numerous reasons, such as network optimization, security, and spectrum management.

However, the efficiency of traditional methods of radio fingerprinting is being increasingly called into question. These conventional methods, while effective in simpler scenarios, are proving insufficient to grapple with the intricate complexity and diversity of modern radio signals [3]. Modern radio devices are capable of utilizing a wide variety of transmission techniques, frequency bands, and signal modulations. As a result, the task of identifying and classifying radio signals has become significantly more complicated than ever before.

Furthermore, the rise of dynamic and adaptive communication systems has further amplified this complexity. Modern communication systems are designed to adapt to changing environmental conditions and user demands, altering their transmission parameters on the fly. This flexibility, while advantageous in terms of system performance and efficiency, introduces additional challenges for radio fingerprinting. Traditional

fingerprinting methods, which rely on static and unchanging signal characteristics, are often unable to accurately identify signals in these adaptive systems [4].

Due to their broadcast nature, wireless networks are more susceptible to assaults like data theft and forgery. As assaults get more sophisticated, it is a difficult problem to detect them, especially in Wi-Fi networks [5].

Hence, there is a growing need for advanced radio fingerprinting techniques that can address these challenges. These advanced methods should be capable of handling the increased diversity and dynamism of modern radio signals, accurately identifying and classifying signals even in the face of changing transmission parameters and evolving communication technologies. The development and refinement of such methods represent an exciting and crucial frontier in the field of wireless communications research.

This chapter provides a concise overview of the problem at hand and outlines the steps taken to address it. The chapter is organized as follows: Section 1.1 delves into the problem statement, while Section 1.2 elaborates on the research objectives. Section 1.3 explores the expected outcomes of this research, and Section 1.4 discusses the significance and motivation behind the thesis. In Section 1.5, a preview of the proposed research model is presented. Finally, Section 1.6 outlines the overall organization of the thesis.

1.1 Problem Statement

A technique called radio fingerprinting intrusion detection looks at the distinctive radio frequency (RF) fingerprints of intruding or malicious devices in wireless networks. Each device emits a unique RF signal pattern that can be recorded and analyzed in order to identify specific devices or spot unusual activity that might be an indication of an intrusion. The procedure entails gathering and examining RF signals sent by networked devices while keeping an eye on features including signal strength, frequency, and modulation. The detected RF signatures can be used to classify and identify network devices by comparing them to well-known reference fingerprints kept in a database. Radio fingerprinting intrusion detection systems frequently use machine learning algorithms to classify and analyze RF signatures. These algorithms have the ability to recognize patterns and distinguish abnormal behavior, allowing the identification of unauthorized devices or possibly harmful network activity. The capacity of radio fingerprinting intrusion detection to identify unwanted devices, detect device impersonation attacks, and improve overall wireless network security are some of its benefits. In dynamic or crowded wireless

situations, however, this approach can be constrained since RF transmissions might be impacted by things like interference or signal changes. Furthermore, prior familiarity with valid device fingerprints or training data for machine learning techniques may be necessary.

1.1.1 Research Gap

Based on the previous researches which took place in the industry for years, several research gaps can be identified in the field of RF fingerprinting and intrusion detection.

One significant research gap is the limited exploration of Convolutional Neural Networks (CNNs) for radio signal identification and classification in the context of intrusion detection. While there have been studies on similarity-based approaches and other techniques for device identification and RF fingerprinting, the specific application of CNNs in this domain remains relatively unexplored. This research gap suggests the need for further investigation into the potential benefits and performance of CNN-based methodologies in improving the accuracy and robustness of intrusion detection systems in Wi-Fi networks.

There is a need for studies focusing on the generalization capabilities of RF fingerprinting techniques. While some research has demonstrated the effectiveness of RF fingerprinting in specific scenarios, it is crucial to investigate how well these techniques can adapt to different environments, variations in signals, and various types of devices. Understanding the limits and strengths of RF fingerprinting in terms of generalization would provide valuable insights for practical deployments.

Also, the scalability and efficiency of RF fingerprinting methods require further exploration. With the increasing number of IoT devices and wireless networks, it is essential to develop scalable and efficient RF fingerprinting approaches capable of handling large-scale deployments. Research efforts should be directed towards addressing the computational demands and volume of data while maintaining high accuracy in RF fingerprinting systems.

1.1.2 Problem Identification

The field of RF fingerprinting and intrusion detection in Wi-Fi networks faces several challenges, including the limited exploration of CNN-based methodologies for radio signal identification and classification, the lack of extensive research on the

generalization capabilities of RF fingerprinting techniques, the need for scalable and efficient methods, and practical challenges in real-world deployments. Addressing these challenges is crucial for advancing RF fingerprinting and intrusion detection, improving the security and reliability of Wi-Fi networks.

1.1.3 Research Question

For our research work, we have formulated some questions based on our findings and the research gaps in the previous researches. These questions are addressed here.

- How the accuracy and efficiency of RF fingerprinting-based intrusion detection systems be improved in dynamic and congested wireless environments?
- What are the optimal parameters and configurations for classification and identification?
- How to deploy and integrate RF fingerprinting-based intrusion detection systems into existing network infrastructures effectively, considering factors like scalability, resource utilization, and compatibility with different wireless technologies?

1.1.4 Scope of the Research

We explored and showed how Convolutional Neural Networks (CNNs) may be used to identify radios in a static environment using raw IQ values. We have emphasized how well our model performs in particular configurations and environments. However, there is still a lot of potential for research and development in this area of study.

The main goal of our ongoing work is to improve the model's resilience and flexibility to different channel conditions and RF impairments. Additionally, we want to look into ways to optimize the neural network's topology so that it more closely matches the demands of this classification task.

In the subsequent sections, we will delve into the particular aspects that we plan to concentrate on during our future investigations.

1.2 Research Objective

In this section, we discuss about our tentative objectives to overcome the research gap. We utilized conventional classification techniques and other deep learning techniques to achieve the goals from the acquired knowledge of previous investigations. Our specific objectives for this research are:

- To enhance the accuracy and efficiency of RF fingerprinting-based intrusion detection systems in dynamic and congested wireless environments through improved algorithms and adaptive techniques.
- To determine the optimal parameters and configurations for classification and identification by conduct a comprehensive analysis and optimization process in RF fingerprinting systems.
- To explore efficient deployment strategies and assess their impact on system performance.

1.3 Research Outcome

Our extensive research on RF fingerprinting serves as a strong and reliable foundation, enabling us to make significant strides in the field and anticipate positive outcomes aligned with our objectives.

Firstly, by applying Convolutional Neural Networks (CNNs) to radio signal classification, this study introduces a new approach that moves away from traditional feature engineering methods. The outcome will be enhanced intrusion detection capabilities, achieving higher accuracy and improved efficiency in detecting and mitigating intrusions in challenging wireless environments, contributing to the advancement of wireless network security.

Secondly, our proposed end-to-end deep learning model will utilize optimal parameters and configurations for classification and identification in RF fingerprinting systems through a comprehensive analysis and optimization process, enabling improved accuracy and performance in intrusion detection and classification tasks.

Thirdly, through our research we want to find effective deployment strategies for integrating RF fingerprinting-based intrusion detection systems into existing network infrastructures, considering scalability, resource utilization, and compatibility with diverse

wireless technologies, leading to optimized system performance and enhanced network security.

1.4 Research Significance and Motivation

Our work on RF based intrusion detection system is important because it enhances network security, improves accuracy in detecting intrusions, provides automation and efficiency, ensures adaptability to evolving network environments, and offers practical application. Challenges include obtaining robust training datasets, addressing signal variations, and ensuring scalability and efficiency.

Firstly, our research focuses on the development of a radio fingerprinting-based intrusion detection system and classifier for predicting intruders in Wi-Fi networks. This work holds significant importance in the field of network security as it addresses the growing need for effective measures to protect against unauthorized access and malicious activities.

Secondly, our approach aims to improve the accuracy of intrusion detection compared to traditional methods. By leveraging radio fingerprinting techniques, we can extract unique characteristics of wireless signals, enabling more precise and robust identification of potential intrusions. This advancement in accuracy enhances the overall efficacy of WI-FI network security.

In addition, our research offers automation and efficiency to the process of intrusion detection. Through the utilization of machine learning algorithms, our classifier automates the analysis of radio fingerprints, allowing for real-time or near real-time predictions about intruders. This automation saves valuable time and resources, enabling timely responses to security threats.

Moreover, our work focuses on developing an adaptable and scalable intrusion detection system. WI-FI networks are subject to constant evolution, with new devices, protocols, and technologies being introduced regularly. By utilizing radio fingerprints, which capture intrinsic properties of wireless signals, our classifier can adapt to different network conditions and detect intrusions across varying Wi-Fi environments.

The practical application of our research is of paramount importance which further motivated us to work in this project. Intrusion detection plays a critical role in various settings, including homes, businesses, public spaces, and critical infrastructure. By providing a reliable and accurate classifier for Wi-Fi network intrusion detection, our

research offers a practical solution that can be implemented in diverse scenarios, enhancing network security and safeguarding sensitive data.

1.5 Overview of the Methodology

Our project introduces a novel deep learning model that operates directly on raw time-domain signals, eliminating the need for manual feature engineering. By adopting this approach, our model can automatically extract and leverage complex hierarchical features from diverse radio signal types, each with its unique modulation techniques.

To validate the effectiveness of our proposed model, we conduct extensive experiments using a comprehensive radio signal dataset. The results demonstrate a significant enhancement in classification accuracy compared to existing methods. Notably, our model exhibits remarkable generalization capabilities when confronted with previously unseen radio signal types.

Furthermore, we investigate the impact of different optimization algorithms on the performance of our model. Through empirical analysis, we provide evidence that careful parameter tuning can improve the computational efficiency of the model without compromising classification accuracy.

1.6 Organization of the Thesis

This thesis comprises seven comprehensive chapters that delve into various aspects of the research. It begins with an introduction, followed by a thorough literature review, theoretical background, data sources, methodology, analysis and interpretation of findings, and a concluding chapter discussing the implications of the research.

Chapter 2 builds upon the foundation set by previous researchers, exploring different techniques for interpreting and intercepting intruders in radio communication. It provides a critical analysis of the existing body of work.

In Chapter 3, we establish the necessary background knowledge to progress in our research. This section offers a concise overview of radio fingerprinting techniques, highlighting their potential as unique identifiers in wireless communication.

Chapter 4 outlines the research methodology employed for radio frequency-based intrusion detection. The procedures and models used in this study are thoroughly elucidated, ensuring a comprehensive understanding of the approach.

Chapter 5 presents the results obtained from simulations conducted on our models. Key findings are presented and discussed, shedding light on the outcomes of our research.

In Chapter 6, we evaluate the performance of our models and compare the accuracy achieved across different versions. This analysis provides insights into the effectiveness and reliability of our proposed approaches.

Lastly, Chapter 7 serves as the conclusion, summarizing our work and addressing the limitations of our study. Additionally, recommendations for further research are provided, offering potential directions for future exploration in this field.

Chapter 2

Literature Review

2.1 Related Works

The field of machine learning (ML), vast and varied as it is, boasts an extensive body of literature exploring its various theories and applications. This discourse, however, narrows its focus to those works that are directly relevant to the problem of Radio Frequency (RF) fingerprinting, primarily emphasizing supervised learning techniques. Supervised learning methodologies come to the fore when available datasets are furnished with appropriate labels.

However, the significance of unsupervised learning methods, particularly effective when initial label information is not readily available for the devices under scrutiny, should not be underestimated. For instance, the research highlighted in presents an intriguing perspective on an online classification algorithm founded on an infinite Hidden Markov Random field (iHMRF) framework. This algorithm leverages unsupervised clustering techniques and batch updates for wireless fingerprinting, thus marking an important development in the field.

Delving further into the unsupervised learning arena, the study cited in navigates transmitter characteristics through a non-parametric Bayesian approach. This approach, specifically deploying an infinite Gaussian Mixture Model, is used to passively classify multiple devices in an unsupervised fashion, underscoring the potential of such methodologies in RF fingerprinting.

An intrusion detection system (IDS) based on ensemble learning is suggested as a solution to the paper's discussion of the growing security risks in Wi-Fi networks. The AWID Wi-Fi intrusion dataset is used by the authors to pinpoint the crucial components of an efficient IDS [19].

Despite the merits of unsupervised learning, our approach involves creating genuine, device-specific data independently. In this context, labeling the device- specific datasets becomes an uncomplicated task. Given this simplicity and the availability of ground truth to guide model creation, we have elected to explore the supervised learning

paradigm. This paradigm involves the utilization of a large trove of labeled samples for training prior to the network deployment, promising greater precision and efficiency. Within the scope of this learning modality, two primary strategies emerge:

2.2 Approaches Based on Similarity Measurements

The cornerstone of similarity-based approaches lies in comparing the unique observed signatures of a particular device against a catalog of reference signatures stored in an exhaustive master database. An excellent demonstration of this technique is illustrated in the research presented in [20], which introduces a unique passive fingerprinting method. This innovative method attempts to identify the specific wireless device driver operating on a node that complies with IEEE 802.11 standards [21]. The process involves the collection of traces from probe request frames sent by the devices under scrutiny. Following the collection process, a supervised Bayesian approach is applied to analyze the accumulated traces meticulously, culminating in the creation of a unique fingerprint for each device driver [22].

The deep learning, dimensionality reduction, clustering, and RF fingerprinting algorithms used in the proposed model pipeline for intrusion detection in IoT devices. Ongoing research focuses on refining hyperparameters for dimension reduction and examining the approach's scalability. It shows promise in detecting previously unnoticed intruder classes or clusters.

The research [23] suggests a unique strategy dubbed deep-feature extraction and selection (D-FES) to address these problems. Weighted feature selection, inspired by shallow-structured machine learning, is combined with stacked feature extraction, which extracts meaningful representations from raw data. This method successfully reduces computational complexity and bias in machine learning models.

Another distinct is a passive black box technique. This technique is predicated on the use of the inter- arrival time of TCP or UDP packets to ascertain the type of access points by deploying wavelet analysis [24]. While these methods present a novel way of device identification, it's important to note that their efficacy is contingent on the availability of prior knowledge of vendor-specific features. This requirement could potentially limit the versatility and universal applicability of these techniques.

In addition to a robust method that involves stacking of processed received signals and the inclusion of artificial noise for repetitive symbols, this study introduces a

comprehensive model for RF fingerprinting. The suggested method gets rid of time-based impacts, improves identification robustness in time-variant channels, and works with any repeated symbol system. The RF fingerprints produced by this method show viability and robustness in AWGN (Additive White Gaussian Noise) or mild NLOS (Non-Line-of-Sight) loss conditions, and they are stable for about 22 months.

2.3 Approaches Based on Classification

2.3.1 Conventional Classification Techniques

Conventional classification approaches operate on the principle of identifying matches with pre-determined features, drawing on the system's domain knowledge. In other words, the principal features need to be identified beforehand. The method involves extracting a known preamble present within a packet and then computing spectral components derived from it [25]. These derived log-spectral-energy features are then fed into a k-nearest neighbors (k-NN) discriminatory classifier for device identification.

The study [26] presents a method for overcoming these difficulties that relies on deep learning for anomaly detection and categorization. This method enables autonomous feature learning to identify various forms of assaults and recognize network irregularities.

Another noteworthy method in this category is PARADIS, which introduces an innovative way of fingerprinting 802.11 devices based on the errors specific to the type of modulation used in the frame[21]. They utilize Support Vector Machine (SVM) and k-NN algorithms to achieve this, boasting an impressive accuracy rate of 99%. Furthermore, a technique for physical device and device-type classification, dubbed as GTID, is proposed in. This approach employs artificial neural networks to leverage variations in clock skews and the distinct hardware compositions of devices [27] – [29]. Despite the potential these techniques offer, a challenge arises when multiple diverse features are employed, as selecting the right set becomes nontrivial. This aspect also engenders scalability problems when dealing with a large number of devices, consequently leading to increased computational complexity during the training phase [30].

We intended to accumulate all this information and found the most suitable models to implement in our work.

2.3.2 Deep Learning Techniques

Deep learning provides a robust and potent framework for learning complex functions and makes the best use of large datasets [31]. It significantly amplifies the number of layers in addition to the neurons within each layer, providing a richer representation of the data. O'Shea and Corgan and O'Shea and Hoydis are notable for applying deep learning techniques at the physical layer, concentrating specifically on modulation recognition using IQ samples and convolutional neural networks [25]. They succeed in classifying 11 different modulation schemes. However, it's crucial to remember that their approach does not facilitate device identification; it simply identifies the type of modulation employed by the transmitter [32].

A ladder network-based deep learning strategy is presented [33] that enables the system to autonomously learn the attributes required for spotting network anomalies and precisely classifying attacks. The inclusion of focal loss as a loss function improves the model's discriminative performance, especially for difficult samples.

There is a strong need for a vast network of linked devices as a result of the integration of IoT and 5G technologies in smart cities [34]. However, as smart cities grow in size, concerns have been raised regarding the security and privacy of IoT devices since hackers can use flaws to launch a variety of assaults.

The research [35] suggests using feature selection to determine the most crucial features for identifying impersonation attacks in order to enhance detection. For the AWID dataset, a deep learning technique called Stacked Auto Encoder (SAE) is utilized as a classifier. An artificial neural network (ANN) is used for feature selection.

In the context of the Internet of Things (IoT), the authors of this study [36] discuss the requirement for scalable, precise, energy-efficient, and impermeable authentication mechanisms. They suggest ORACLE, a cutting-edge system that uses convolutional neural networks (CNNs) to "fingerprint" and distinguish a particular radio device from a wide range of devices. Through the use of deep learning algorithms, ORACLE discovers the specific hardware limitations created by the radio circuitry on physical-layer I/Q samples.

In another preliminary work, they utilized raw IQ samples and a Convolutional Neural Network (CNN) to identify low-end Software-Defined Radio (SDR) radios. As per our knowledge, ORACLE marks a milestone in being the first work that trains a CNN for bit-similar device identification, allowing the same classifier to operate in unknown or dynamic channel conditions without necessitating new trials.

They point out that because it is challenging to identify device-specific hardware flaws on wireless channels, CNN-based fingerprinting can be less accurate. The authors [37] give an expansive open dataset of over 7TB of wireless data that was gathered from 20 devices with identical RF circuitry in a variety of settings and channel circumstances to address the research gaps. Using this dataset and an extra 400GB dataset provided by DARPA, they carry out a thorough assessment of the effect of the wireless channel on CNN-based fingerprinting techniques.

In this study [38], the authors investigate the application of the Short-Time Fourier Transform (STFT) for the analysis of the specific frequency patterns of RF signals from mobile devices and precise localization of these devices inside a building. Seven cell phones were put in various locations on a single floor, and the researchers recorded raw signals and channel state information (CSI) frames from each one. They used three software-defined radios (SDRs) to simultaneously receive transmissions.

From all these deep learning techniques, we wanted to implement such a model that can identify and classify the radio signal at the same time, more efficiently.

Chapter 3

Background Study

3.1 Radio Fingerprinting

Radio fingerprinting, in its essence, refers to the meticulous process of discerning and categorizing unique characteristics, colloquially termed as 'fingerprints,' that are innately present in every radio device [6]. These unique fingerprints can be traced and delineated across multiple layers of a signal's composition, extending from inherent hardware traits that are encapsulated at the physical layer, right up to the distinct transmission patterns observable at higher, more abstract layers. Consequently, the process of fingerprinting entails an intricate, multi-layered analysis aimed at recognizing and cataloging unique patterns, as well as identifying deviations that can function as differentiating features or hallmarks of a specific radio device.

The distinct fingerprints borne by each device can often be traced back to inherent discrepancies in their hardware components. For instance, variations in oscillator frequencies, characteristics of the power amplifier, and the employed modulation schemes all contribute to a device's unique signature [7]. These hardware eccentricities often manifest as identifiable patterns within the transmitted signals. Thus, through careful extraction, analysis, and understanding of these patterns, a unique identity or 'fingerprint' for each radio transmitter can be formulated.

This meticulously constructed 'fingerprint' then holds immense potential for a wide array of applications. A primary use case lies in device authentication, where the unique identity aids in verifying the legitimacy of a radio device, thereby fortifying the authentication process [8]. Additionally, it proves instrumental in facilitating advanced intrusion detection mechanisms. By maintaining an inventory of 'known' fingerprints, any deviant or unregistered fingerprint can be rapidly flagged as a `potential security threat, thereby bolstering the overall cybersecurity framework.

By providing a mechanism to identify and authenticate individual devices, radio fingerprinting aids in maintaining a robust, secure, and trustworthy communication

network. It ensures that only authenticated devices gain access to network resources and that any potential intruders or malicious devices are quickly identified and isolated.

Thus, radio fingerprinting, by leveraging the inherent hardware discrepancies and unique transmission patterns of radio devices, provides a sophisticated and reliable mechanism to enhance the security and integrity of wireless communication networks. However, to exploit its full potential, it is essential to continue advancing our understanding of the complex signal characteristics and improving our methodologies in the intricate process of radio fingerprinting.

3.2 Wi-Fi Signal Characteristics

The IEEE 802.11b standard provides the foundational specifications for the physical layer of Wi-Fi-enabled devices. The moniker "Wi-Fi" is ubiquitously employed to refer to wireless local area networks that operate in accordance with this standard. The specifics of the Wi-Fi physical layer are comprehensively delineated in referenced literature, yet it's valuable to outline some fundamental spectral and time domain characteristics here for a deeper understanding [9].

Wi-Fi devices predominantly function within the 2.4 GHz band, designated for industrial, scientific, and medical (ISM) usage [10]. The operation of these devices across different channels within this band, including the number of active channels and the center frequency for each, is subject to allocation by relevant regulatory entities in regions such as North America, Europe, and Japan. These allocations are clearly specified in the standard, ensuring consistent and regulated usage across the spectrum [11].

One of the key aspects mandated by the standard is the acceptable tolerance for the transmitted center-frequency, which should not exceed ±25 parts per mill ppm). This regulation ensures a high degree of accuracy in the center frequency of the transmitted signal, thereby reducing potential interference and enhancing overall signal integrity.

Moreover, the standard sets forth stringent requirements regarding the spectral emissions of Wi-Fi signals. It states that transmitted spectral products should not exceed -30 dBr (decibels relative to the sin(x)/x peak) for frequencies ranging between 11 and 22 MHz off from the center frequency. Additionally, for frequencies deviating by more than 22 MHz from the center frequency, the spectral products must be kept under -50 dBr. As a result, the effective bandwidth of the Wi-Fi signals extends to 22 MHz, enabling efficient usage of the allocated spectrum. A unique characteristic of the IEEE 802.11b standard is

its specific mention of a gradual scheme for the power-up and power-down of the transmitting device [12]. This particular provision is intended to prevent the inadvertent spread of power to adjacent channels when the device is switching on or off. By enforcing a gradual ramping up and down of the power, the standard mitigates the risk of causing interference with other devices operating on neighboring channels. This thoughtful inclusion underscores the meticulous attention to detail that the standard has adopted to ensure the smooth functioning and coexistence of multiple devices within the Wi-Fi network ecosystem [9], [13], [14].

3.3 Unique Identifier Patterns within IQ Samples

In the domain of RF identification and fingerprinting, the concept of unique signatures extracted from In-phase and Quadrature (IQ) samples holds great promise. A typical wireless communication system relies on IQ modulation, where each data symbol is represented by a point in a predefined constellation plot. In an ideal scenario, these points are fixed and precisely known, however, in reality, they are subject to various transmitter hardware imperfections that cause deviations from the ideal positions [15].

These deviations are primarily caused by three types of hardware imperfections: IQ imbalance, nonlinear distortion, and phase noise. IQ imbalance refers to the mismatch in amplitude or phase between the in-phase and quadrature components. Nonlinear distortion, often the result of power amplifiers operating in non-linear regions, can lead to a constellation diagram that varies from the ideal form. Phase noise, typically induced by instability in the oscillator, can cause a rotation or jitter in the constellation points.

Interestingly, these hardware imperfections, while usually seen as challenges in achieving optimal signal quality, serve as robust and unique signatures for a given transmitter. This is because these deviations are often intrinsic to the hardware itself, remaining consistent over time and being resilient to most environmental changes.

Moreover, it is possible to intentionally modify certain aspects of these inherent hardware signatures, through programmable interfaces like SDR platforms [16], [17]. This manipulation not only further enhances the uniqueness of the signatures but also has the potential to improve the efficiency of device classification models, providing a new dimension for research and application in the field of wireless security identification. [18]

Chapter 4

Methodology

4.1 Identification Scheme in Radio Communication

This section of our discussion commences with a meticulous investigation into the RF hardware impairments that yield variations within IQ samples, consequently establishing a unique signature, or 'fingerprint', for each individual device. Our research principally centers on two such hardware impairments: IQ imbalance and DC offset. These specific impairments have been chosen due to their two distinctive characteristics. Firstly, they are unaffected by environmental conditions, thereby ensuring the consistency and repeatability of the resulting device signatures [39]. Secondly, they are not constrained to a specific transmitter-receiver pair. These properties distinguish them from other potential impairments, such as the relative phase offset, which only manifests in the context of a specific transmitter-receiver pair.

Following the exploration of these impairments, we present a sophisticated method that facilitates the deliberate introduction of controlled impairments. To achieve this, we utilize the GNU Radio UHD API at the receiver end. The main goal of this methodology is to deliberately create variations within the IQ samples in a controlled manner, thereby enhancing the uniqueness and identifiability of each device's RF signature.

Subsequent to the discussion on controlled impairments, we provide a comprehensive explanation of the experimental testbed setup that we employed for collecting trace data. This setup is pivotal to our research as it serves as the platform for studying the impacts of the aforementioned RF hardware impairments on IQ samples, and for testing the effectiveness of our proposed method for introducing controlled impairments. Through this experimental setup, we collect and analyze a significant amount of trace data, paving the way for a deeper understanding of RF fingerprinting in the context of IQ sample variations. Our findings from these experiments further enrich our insights into the distinctive signatures of individual devices, providing a firm basis for advanced classification and identification schemes in radio communications.

4.2 RF Impairments

Leveraging the MATLAB Communications System Toolbox, we emulate a typical wireless communications processing chain, as depicted in Figure 4.1. This figure illustrates the shifts observable in the received complex-valued IQ samples. By modifying the ideal operational blocks within this simulated processing chain, we are able to intentionally introduce RF impairments that commonly occur in real- world hardware implementations. This simulation approach facilitates an in-depth, individualized study of various RF impairments, including IQ imbalance, DC offset, phase noise, carrier frequency offset, and nonlinear distortions typically introduced by power amplifiers.

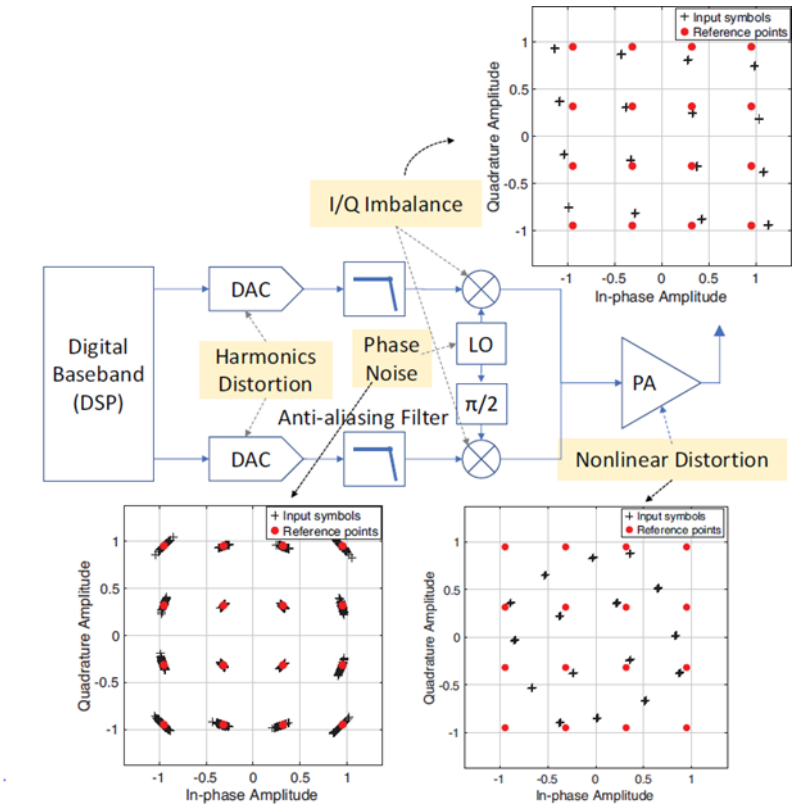

Figure 4.1: Typical transceiver chain with various sources of RF impairments.

The impact of drifting instantaneous carrier frequency offset (CFO) on system stability and misclassification is discussed in the study [32]. To counteract this, a hybrid classifier is created to modify CNN outputs based on the estimated CFO, and CFO compensation is implemented. For device authentication in LoRa systems, the suggested spectrogram-based RFFI approach performs better than other current IQ-based and FFT-based schemes.

The simulation investigation [40] of redesigned wireless communication processing blocks to show RF impairments is one of the articles' significant contributions. Using an over-the-air dataset, it suggests an improved deep convolutional neural network (CNN) architecture and assesses how well it performs in comparison to existing methods like logistic regression and support vector machines.

However, in the scope of this paper, we primarily concentrate on two particular impairments: IQ imbalance and DC offset. This specific focus is due to the limitation of space for our discussion, yet it's important to mention that our methodology can be effortlessly expanded to encompass the investigation of the remaining RF impairments [10], [41]. By using the MATLAB Communications System Toolbox for our simulation, we effectively create an insightful, controlled environment that permits us to study the impacts of these impairments on the behavior and performance of a typical wireless communications processing chain. Through this approach, we aim to deepen our understanding of the characteristic variations in IQ samples that can serve as distinctive identifiers or 'signatures' for individual devices in a wireless network.

4.2.1 IQ imbalance

Quadrature mixers, fundamental components in the radio frequency (RF) communication chain, are frequently subject to gain and phase discrepancies between the parallel segments that handle the in-phase (I) and quadrature (Q) signal paths. These discrepancies engender an imbalance in the IQ components of the signal. The differential gain in the I and Q paths results in amplitude imbalance, while any deviation from the ideal 90-degree phase difference between the I and Q components triggers phase imbalance. Interestingly, the degree of IQ imbalance fluctuates solely with frequency due to the presence of frequency-dependent low pass filters in the signal chain [42]. Therefore, it effectively provides a unique, frequency-specific signature of a transmitter, marking each device with an identifiable characteristic in its transmitted signals.

4.2.2 DC Offset

Another impairment observed within the domain of quadrature mixers is DC offset. It primarily originates due to the limited isolation between the Local Oscillator (LO) and RF ports within a mixer. Given this imperfect isolation, the LO signal often experiences a

direct feedthrough which gets coupled to the output[43]. This leakage of the LO signal into the output path results in a constant offset, or DC offset, in the signal. This DC offset can act as another distinctive feature or 'signature' of a device's transmitter, further enriching the RF fingerprinting process.

4.3 Software-based Control of Impairments

The process of transmitter chain calibration is a critical component in our exploration, particularly in relation to the setting of IQ imbalance and DC offset. To facilitate this, we leverage the self-calibration utilities offered by Ettus Research, a leading provider of Software Defined Radio (SDR) solutions. Their calibration utilities aid in managing and mitigating the discrepancies introduced by hardware characteristics in the transmitter chain, thereby enhancing the precision and reliability of our CNN model's performance.

The Ettus self-calibration utilities are integrated within the GNU Radio software, which is a free, open-source platform for designing, simulating, and deploying radio systems. By utilizing the utilities within this software framework, we're able to adjust the IQ imbalance and DC offset directly. GNU Radio functions are used as part of this calibration process, contributing to the practicality and effectiveness of the procedure.

The IQ imbalance refers to the mismatch in amplitude and phase between the in-phase (I) and quadrature (Q) components of a signal. If left uncalibrated, this imbalance can introduce significant distortions in the signal, thereby undermining the model's accuracy in classifying radios [44]. Similarly, the DC offset, which is an average amplitude displacement from zero, if not corrected, can lead to a systemic error in signal interpretation.

By employing the Ettus self-calibration utilities within GNU Radio, we can correct for these potential distortions, thereby ensuring that the signals input to our model are as accurate and representative as possible. This calibration process, therefore, represents a crucial preparatory step in our methodology, laying the groundwork for the reliable functioning and performance evaluation of our CNN model [43].

4.3.1 IQ Imbalance Compensation

Understanding the effects of IQ imbalance in the transmitter chain is critical for the accurate classification of devices in a wireless communication system. To elaborate further on the model described, we must first examine the transmitted baseband complex signal $s(t) \in C$, which is the ideal form of the signal before any form of distortion. The distorted signal, termed $s_d(t)$, in the time domain is described by the equation:

$$s_d(t) = \mu_t s(t) + v_t s^*(t)$$

This equation demonstrates how the transmitted signal s(t) is distorted by the IQ imbalance parameters μ_t and v_t. These parameters encapsulate the discrepancies between the amplitude and phase in the I (In-phase) and Q (Quadrature) signal paths within the quadrature mixer, which is a fundamental component of the transmitter chain in most wireless devices.

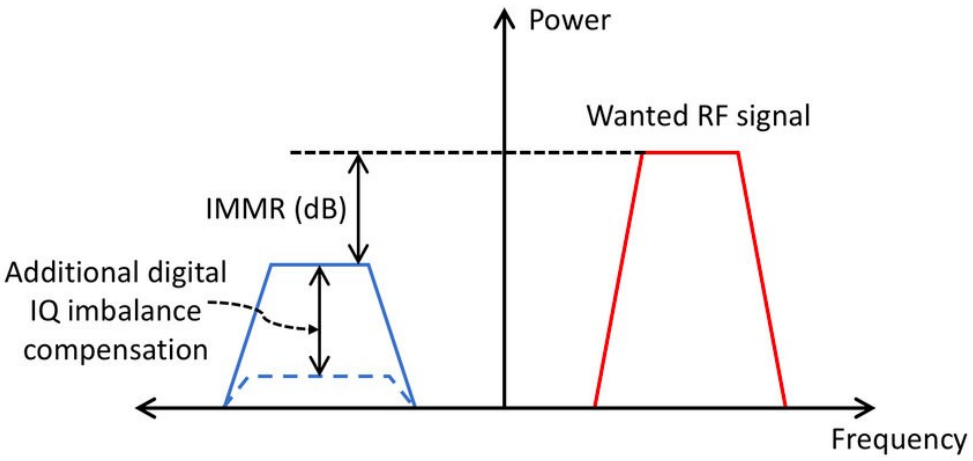

Figure 4.2: Effect of IQ imbalance quantified through IMRR

To simplify these distortion parameters, we can express μ_t as $\cos(\theta_t/2) + j\alpha_t \sin(\theta_t/2)$ and v_t as $\alpha_t \cos(\theta_t/2) - j \sin(\theta_t/2)$ where α_t represents the amplitude imbalance and θ_t the phase imbalance between the I and Q paths at the transmitter. Specifically, the phase imbalance is defined as any deviation from the ideal 90° phase difference between I and Q paths, while the amplitude imbalance is measured as $\frac{\alpha_I - \alpha_Q}{\alpha_I + \alpha_Q}$, where α_I and α_Q represent the gain amplitudes on the I and Q paths, respectively.

The effects of IQ imbalance include interference with the original signal, creating a mirror image of the signal at an opposing frequency. This effect can be measured and quantified by the Image Rejection Ratio (IMRR), which calculates the power of the mirror image (or undesired signal) in relation to the desired signal. The equation for IMRR, considering amplitude imbalance αt and phase difference θ_t,

$$IMRR = \frac{\gamma_t^2 + 1 - 2\gamma_t \cos\theta_t}{\gamma_t^2 + 1 + 2\gamma_t \cos\theta_t}$$

where γ_t is defined as $\alpha_t + 1$.

To mitigate the effects of IQ imbalance, one can use several theoretical time and frequency domain methods. However, for this study, we decided to utilize the UHD calibration utility named '*uhd_cal_tx_iq_balance*' provided by Ettus Research. This utility allows for a calibration sweep over a range of frequencies, effectively checking the leakage of the transmission path signal into the receive path, which is a common consequence of IQ imbalance.

Once a calibration sweep is completed, the UHD software automatically applies a corrective factor to the transmit chain of the RF daughterboard. This correction factor is usually a single complex number that is applied to the transmission signal. For each applied correction factor, a single frequency tone is transmitted, and the power of the desired and mirror frequency tones are measured to calculate the IMRR.

Using only IQ samples at the physical layer, ORACLE is a method for identifying a specific radio equipment from a wide group of similar devices. Convolutional neural networks (CNNs) are used by ORACLE to attain a 99% classification accuracy rate while balancing computational time and accuracy. The authors look into the aspects of the transmitter chain that are hardware-centric and contribute to variances in IQ samples, and they suggest a CNN architecture that simply needs raw IQ samples and doesn't require any prior knowledge of the communication protocol [45].

We have made modifications to this calibration utility to record both the applied correction factors and the corresponding IMRR for each. A snapshot of these recorded IMMR levels for the USRP X310 radio operating at a center frequency of 2.45 GHz is provided in Table II. This comprehensive record allows for better evaluation of our model's resilience against IQ imbalance in the transmitter chain, ultimately improving the performance of our wireless device classification system.

4.3.2 DC Offset Compensation

DC offset, as aforementioned, engenders a substantial spike at the center of the signal spectrum. The power of this dominant tone at the Direct Current (DC) frequency can be measured to determine the magnitude of DC offset present. In an effort to correct this DC offset level, a UHD calibration utility, namely '*uhd_cal_tx_dc_offset*', employs a single complex factor. It meticulously searches for the optimal correction factor that minimizes the power of the DC tone. To gather data on the DC offset levels, we have adapted this utility to record the levels corresponding to the correction factor.

In the domain of Software Defined Radio (SDR) operation, we utilize the open-source GNU Radio Companion (GRC), which facilitates the transmission of standard-compliant IEEE 802.11a Wi-Fi packets through the SDR. By leveraging the '*set_iq_balance*' and '*set_dc_offset*' functions in GRC, we can assign two distinct complex correction factors. These factors are instrumental in deliberately introducing the necessary levels of impairments into the radio, providing us with a means to experimentally emulate these hardware distortions. This method of intentionally introducing impairments enables us to better understand their impact on the radio signal's behavior and potentially manipulate them to enhance our radio fingerprinting capability.

4.4 Experimental Framework for Gathering Trace Data

The performance of the Convolutional Neural Network (CNN) is examined utilizing In-phase and Quadrature (IQ) samples gathered from an experimental configuration consisting of Universal Software Radio Peripheral (USRP) Software Defined Radios (SDRs) [46], [47]. This setup, as depicted in Figure 03, employs a fixed USRP B210 as the receiver. All transmitting radios in this configuration are bit- similar USRP X310 devices that broadcast IEEE 802.11a standards-compliant frames, which are generated through the MATLAB WLAN System Toolbox.

The data frames produced contain a randomized payload but maintain consistent address fields. These frames are subsequently streamed to the selected SDR for over-the-air wireless transmission. The receiver SDR processes the incoming signals by sampling at a rate of 5 million samples per second (MS/s) with a central frequency set at 2.45 GHz, a typical frequency for Wi-Fi communication.

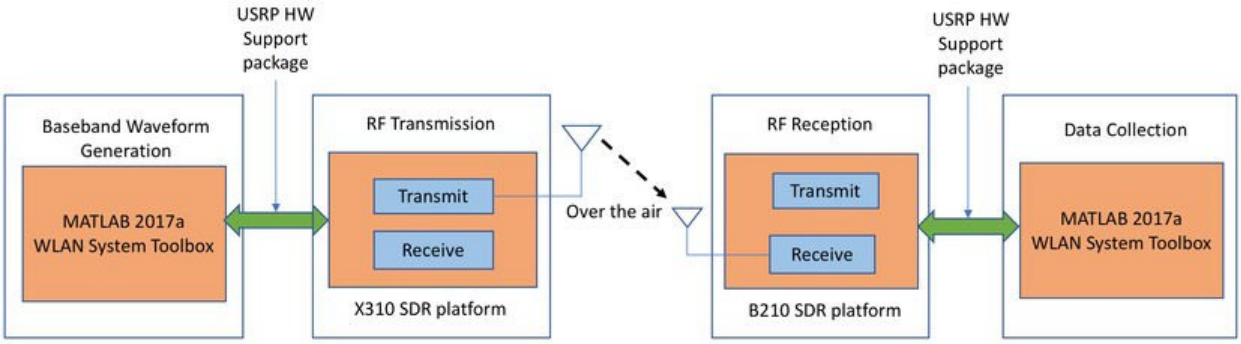

Figure 4.3: Experimental setup for data collection using SDR

The received complex IQ samples are segmented into subsequences to facilitate analysis. In this study, we establish a fixed subsequence length of 128. This means that each contiguous block of 128 samples is processed as a unit during training and classification. In total, we gather over 20 million samples for each radio. This colossal dataset is then subdivided into distinct sets for training, validation, and testing, ensuring a robust and comprehensive evaluation of our CNN's performance in the context of radio fingerprinting. The data frames produced contain a randomized payload but maintain consistent address fields. These frames are subsequently streamed to the selected SDR for over-the-air wireless transmission. The receiver SDR processes the incoming signals by sampling at a rate of 5 million samples per second (MS/s) with a central frequency set at 2.45 GHz, a typical frequency for Wi-Fi communication.

The received complex IQ samples are segmented into subsequences to facilitate analysis. In this study, we establish a fixed subsequence length of 128. This means that each contiguous block of 128 samples is processed as a unit during training and classification. In total, we gather over 20 million samples for each radio. This colossal dataset is then subdivided into distinct sets for training, validation, and testing, ensuring a robust and comprehensive evaluation of our CNN's performance in the context of radio fingerprinting [47].

4.5 CNN ARCHITECTURE FOR STATIC CHANNEL

4.5.1 Classifier Architecture

In our research, we construct a Convolutional Neural Network (CNN) model tailored to work with static channels in a wireless communication context. The data feeding this model consists of raw time-series In-phase and Quadrature (IQ) samples. These samples are generated from a 16-node USRP X310 Software Defined Radio (SDR) testbed, and our model is further enriched by an external database that comprises 140 Commercial Off-The-Shelf (COTS) Wi-Fi devices. Such a database is a rich source of real-world variability, enhancing the robustness and generalizability of our model.

Our CNN design is, in part, inspired by AlexNet [13], a ground-breaking deep CNN model which was originally designed for image classification tasks. AlexNet made significant contributions in the field of computer vision by classifying a substantial volume of high-resolution images, precisely 1.2 million, from the ImageNet dataset into 1000 diverse classes.

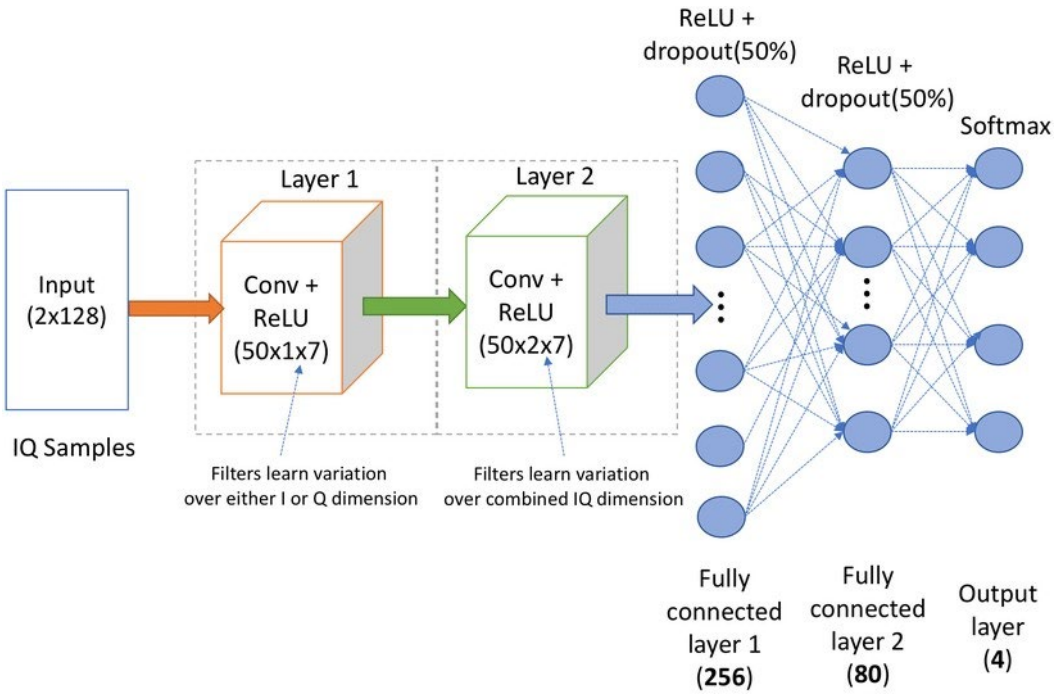

Figure 4.4: Proposed CNN Architecture

Unlike AlexNet, which is composed of eight layers (five convolution and three fully connected), our proposed CNN architecture consists of four layers, two of which are

convolution layers and two are fully connected (or dense) layers. This difference in structure is due to the distinct nature of our problem compared to the image classification task that AlexNet was designed to solve.

The input to our CNN is a windowed sequence of raw IQ samples, where each sequence comprises 128 samples. We employ a sliding window technique to partition the training samples, which enhances the shift-invariance of the features learned by the CNN. This approach provides a level of robustness to time delays and channel shifts that are inherent in wireless communications.

In the realm of wireless communication, complex values are represented as two-dimensional real values, where In-phase and Quadrature components are treated as two distinct real-value streams. This conversion amplifies the dimension of our input data to 2×128. These values are then fed into the initial convolution layer of the network.

Each convolution layer in our architecture houses a set of spatial filters or kernels. These kernels perform a convolution operation over the input data to extract critical features. The first convolution layer is composed of 50 filters, each of size 1×7. Every filter is tasked with learning a 7-sample variation over time within the *I* or *Q* dimension independently. This procedure yields 50 distinct feature maps spanning the entire input sample.

On similar lines, our second convolution layer comprises 50 filters, each of size 2×7. These filters learn variations of 7 activation values across both *I* and *Q* dimensions of the 50-dimensional activation volume, obtained post the initial convolution layer. Each convolution layer is succeeded by a Rectified Linear Unit (ReLU) activation function, which applies a predetermined non-linear transformation to each element of the convolved output.

The output of the second convolution layer is passed to the first fully connected layer, which contains 256 neurons. A secondary fully connected layer with 80 neurons is added to discern higher-level non-linear combinations of the features, which are extracted from previous layers. These features are then passed onto a classifier layer.

In the final layer, we employ a Softmax classifier. This type of classifier outputs the probabilities of each sample belonging to each class. In essence, it provides a measure of certainty about the model's predictions, which is valuable for decision- making and model interpretation.

The selection of hyperparameters such as filter size, number of filters in the convolution layers, and the depth of the CNN, are crucial factors in shaping the performance of our CNN model. Through rigorous cross-validation, we carefully fine-tune these parameters to ensure that our model generalizes well to unseen data. To prevent overfitting, a common pitfall in deep learning models, we set the dropout rate to 50% in the dense layers.

In addition to the dropout mechanism, we also employ l_2 regularization with a parameter λ=0.0001. Regularization introduces a penalty on the magnitude of the parameters, preventing the model from relying too heavily on any single feature, thereby enhancing its generalization ability. The weights of the network are trained using the Adam optimizer with a learning rate of $l_r = 0.0001$.

We utilize back-propagation, a staple in neural network training, to minimize the prediction error. For this purpose, we use categorical cross-entropy as a loss function, calculated on the classifier output. Cross-entropy loss provides a suitable measure of dissimilarity between the model's predictions and the actual labels.

Our CNN architecture is implemented in Keras, a user-friendly neural network library, running on top of TensorFlow—an open-source platform renowned for its flexible and extensive machine learning capabilities. The implementation of this model, along with rigorous evaluation and validation, constitutes a significant step forward in applying deep learning techniques to wireless communication systems.

Chapter 5

Result

5.1 Overview of the Result

The initial stage of our evaluation is designed with two main goals. Firstly, we want to substantiate the accuracy of the Convolutional Neural Network (CNN) architecture in accurately identifying and classifying radios in static conditions. Secondly, this stage sets the stage for understanding the need for receiver-feedback driven modifications that are crucial for dynamic channels.

5.2 Analyzing Accuracy in Static Channel Condition

In this segment, we embark on a thorough investigation of our proposed CNN's performance in classifying Commercial Off-The-Shelf (COTS) Wi-Fi devices. The devices used for this evaluation are sourced from an expansive external database. This database comprises of meticulously labelled In-phase and Quadrature (IQ) samples that have been gathered from a diverse array of 140 devices. The collection includes mobile phones, tablets, laptops, and drones, and covers a vast spectrum of 122 different manufacturers.

This extensive collection of devices offers a unique advantage by subjecting our model to a wide range of real-world variations. This approach ensures that the model's robustness and adaptability are adequately tested.

The training regimen for each device involves using 4,500 windowed examples, while 1,000 examples are set aside for testing. This segregation is based on the samples available for each device in the database. In addition to the training and testing samples, we reserve a validation set with 300 examples for each device. This validation set plays a pivotal role during each training epoch as it allows us to monitor the model's performance on data it has not previously encountered.

Our training protocol is designed for efficiency and effectiveness. If the validation accuracy does not increase for 10 consecutive training epochs, the training process is halted.

This mechanism prevents wasteful computations and helps optimize the model's learning curve.

The training phase for this experiment, which includes all 140 devices, takes approximately 15 minutes. A graphical representation of our model's performance, shown in Figure 5.1, provides detailed insights. It presents an exhaustive report including the minimum accuracy, first quartile, median, third quartile, and maximum accuracy for each dataset. The graph has the number of randomly chosen devices on the X-axis, while the Y-axis denotes the classification accuracy.

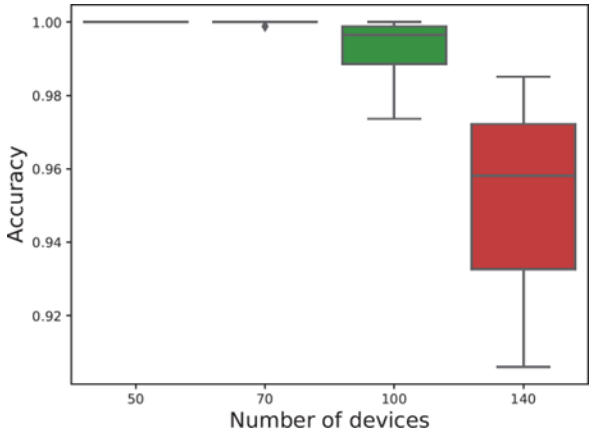

Figure 5.1: Box plot for the classification of Wi-Fi devices using CNN

Remarkably, our model exhibits a robust performance with a median accuracy of 97% for up to 100 different devices. As the number of devices increases to 140, there is a minor decrease in accuracy to 96%. However, this small dip is expected and acceptable considering the complexity and diversity of devices being classified.

Despite the large number of radios, it is crucial to remember that these devices are not identical in their bit characteristics. This variability presents a unique challenge and an opportunity to 'stress-test' our model.

To push our classifier to its limits, we gathered IQ samples from 16 high-end X310 USRP SDRs. These radios, though of superior quality, exhibit a narrower range of impairments. For these trials, we use the same B210 radio as a receiver. The training set for this experiment, per radio, consists of a massive 200,000 windowed training examples and 10,000 examples for validation. We also have an additional set of 50,000 examples for each device reserved for testing the performance of our trained model.

Even under these stringent conditions, our model takes roughly 50 minutes to train for 16 radios. The results from this setup are promising and exhibit a test accuracy of 93.6%. This high accuracy further validates the efficiency and effectiveness of our CNN architecture under static channel conditions.

5.3 Limitations of Raw IQ Samples in Dynamic Channels

In the realm of real-world wireless communications, various factors such as multipath reflections and fading pose significant challenges, particularly affecting the In-phase and Quadrature (IQ) samples received by a radio. These elements can potentially distort the IQ samples to the extent that the classifier, which is designed to accurately identify the radios based on these samples, fails to perform its task effectively.

In a typical wireless communication setup, these channel-induced distortions are mitigated using various techniques such as channel estimation and equalization. The primary aim of these methods is to accurately extract the over-the-air transmitted data, thus compensating for the adversarial effects of the channel. However, in the context of our CNN-based classifier, we observed a notable decrease in performance under two primary conditions: (i) when the classifiers, trained on raw IQ samples from a specific channel environment, are tested on IQ samples acquired under different channel conditions; and (ii) when the transmitters undergo very similar channel conditions, hence generating similarly affected IQ samples.

Our results illustrate a nearly flawless classification accuracy across 16 X310 radios when tested in the same channel conditions they were trained in. But, replicating the same experiment under a different channel condition yields a significantly different outcome. There are considerable deviations in the confusion matrix, with radio pairs like (5,15) and (10, 14) producing some noteworthy outliers. This discrepancy arises due to the transmitters experiencing similar channel conditions that overshadow the subtle but crucial hardware variations which aid in differentiating the radios.

Let's consider a set of R radios. For each radio $r_i \in R$ within this set, $H_i(k)$ represents the average channel gain on the k^{th} subcarrier, estimated based on Wi-Fi packets from the training dataset.

Figures 5.2 (a) and 5.2 (b) elucidate an interesting observation. The received samples from transmitters which exhibit minimal differences in their channel estimation are more likely to be misclassified during testing. It implies that the state of the wireless

channel has a substantial impact on the distribution of complex symbols that are captured by the receiver. As such, the wireless channel state itself emerges as a discriminating factor when the classifier is trained using raw IQ samples.

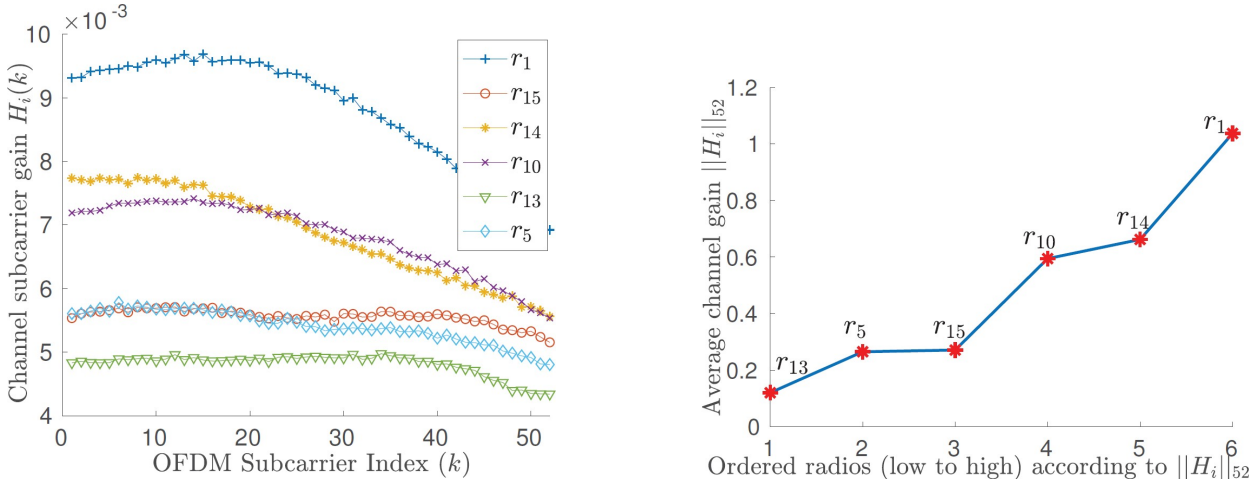

Figure 5.2: (a) Estimated channel gain $H_i(k)$ for k^{th} subcarrier for each radio $r_i \in R$ (b) Magnitude of estimated channel $\|H_i\|_{52}$ for all radios $r_i \in R$ (ordered from lower to higher).

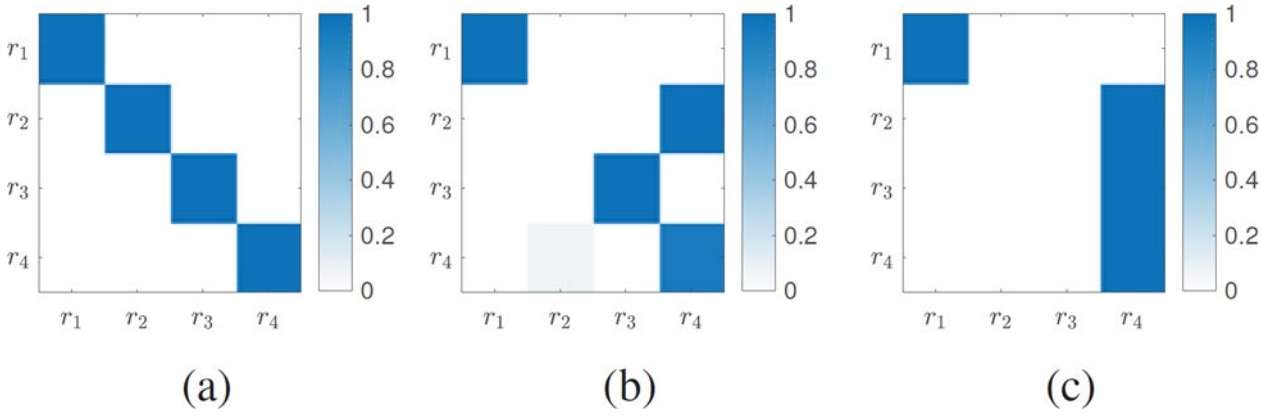

Figure 5.3: (a) Classification accuracy for 4 devices tested at time t_1 and location l_1 (b) time t_2 and same location l_1 (c) time t_3 and different location l_3

The influence of the channel state is further underscored when we employ a pre-trained model for classifying samples. If the pre-trained model is used for classifying samples collected from the same devices but at different times or locations, the

classification results become unpredictable. This underscores the significance of the time and location dependence of the trained classifier, as depicted in Figures 5.3 (a), 5.3 (b), and 5.3 (c), which showcase how the classification results change depending on the time and location at which the samples were collected.

In summary, although our proposed CNN architecture showcases impressive results under static channel conditions, it is clear that the dynamic and often unpredictable nature of real-world wireless channels requires more adaptive solutions. Therefore, a primary focus moving forward will be on developing robust techniques that can adapt to these dynamic channel conditions while maintaining high classification accuracy. This will help to ensure the reliability of the CNN- based classifier, and thus improve the overall performance and efficacy of the system in diverse real-world scenarios.

Chapter 6

Discussion

6.1 Performance Analysis and Evaluation

The Performance Evaluation chapter represents a critical segment of our research, where the functionality, robustness, and adaptability of our developed convolutional neural network (CNN) model are methodically examined and interpreted. The primary aim of this chapter is to assess the model's efficacy in the realm of radio classification based on raw in-phase and quadrature (IQ) samples.

In our experiments, we consider a range of conditions, including static channels and varying degrees of complexity in radio classification tasks. By doing so, we scrutinize the model's performance under various realistic scenarios, thereby facilitating a comprehensive evaluation of its capabilities. Each experimental condition offers distinctive insights that allow us to understand not just the model's strengths, but also its weaknesses and areas that need improvement.

The process of performance evaluation is an iterative one. Our analysis isn't confined to the initial performance of the model. Instead, it is an ongoing process, repeated after every significant modification to the model, to ensure continuous improvement in terms of precision and robustness. The metrics that we gather from the model's current performance form the benchmark against which the effects of subsequent changes are measured.

Further, the evaluation extends beyond mere quantitative analysis. Each result is contextualized and interpreted against the backdrop of our objectives and the specific constraints of the task at hand. Such detailed evaluation helps us understand the nuances of the model's operation and draws attention to any unanticipated behavior or patterns.

In essence, this chapter provides the evidence for the effectiveness of the current model while also offering a roadmap for future improvements. The insights gleaned from this comprehensive and systematic evaluation play a crucial role in shaping our future efforts, directing us towards areas that hold promise for further refinement and exploration.

This makes the performance evaluation not just an analysis of the present state, but also a guide for future research directions.

6.2 Classifier Accuracy with Different Channel Conditions

In our continued efforts to ascertain the capabilities of the Convolutional Neural Network (CNN) model that we meticulously trained; we undertook a series of experiments utilizing 16 X310 radios. The initial acquisition of samples from these radios was conducted via cable, with each radio having been configured with a distinct impairment, drawn from an enumerated set (denoted as set 'S').

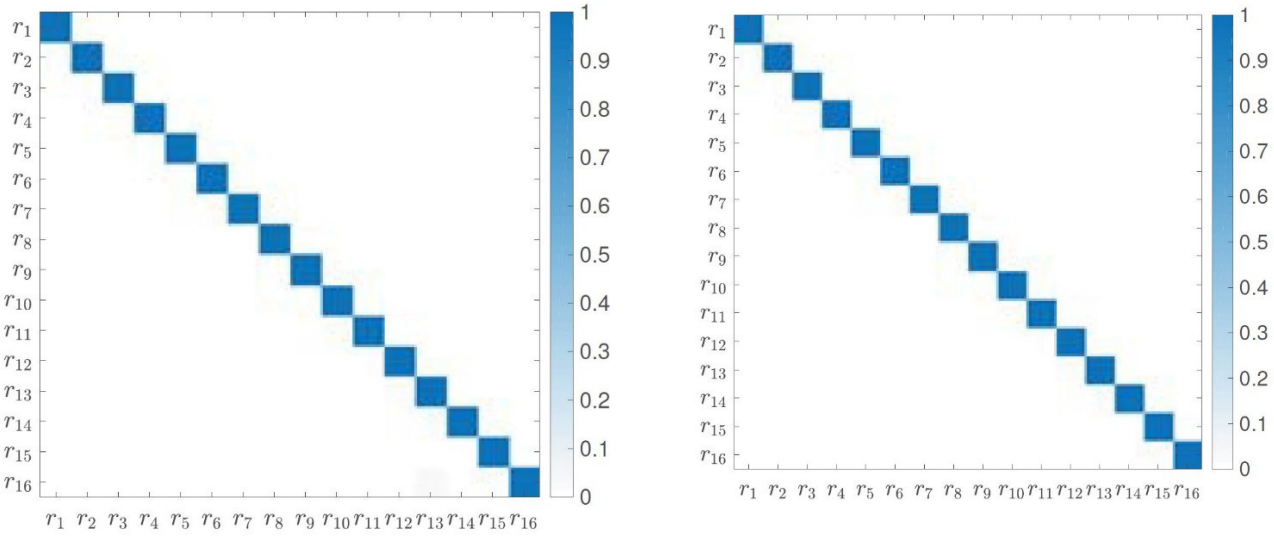

Figure 6.1: Classification Accuracy (a) via cable; (b) over the air

The results of these tests offer compelling evidence of our model's proficiency. Even in circumstances where bit-similar radios were intentionally subjected to selected impairments, the model exhibited outstanding discerning capabilities, thereby achieving a remarkable classification accuracy of 93.76%. Such a high level of accuracy emphasizes the exceptional capability of our pre-trained CNN classifier in accurately identifying and differentiating between bit-similar radios.

To push the boundaries of our CNN model, we progressed from cable-based tests to an assessment of its performance with data collated over wireless channels. The primary goal of this shift was to demonstrate the resilience of our classifier amidst variations in channel conditions. In pursuit of this, we performed the experiments across two unique

locations: firstly, within our laboratory premises, representing a typical indoor environment, and secondly, in a more open and less congested recreational area characterized by reduced reflections.

The subsequent classification accuracy, visualized via confusion matrices in Figure 08b, proved impressive in both environments. Our CNN model consistently maintained an exceptional accuracy exceeding 94.5%. This outcome stands as a testament to the model's robustness in detecting the unique patterns manifested by the impairments, even in the face of random noise that is typical of wireless communication.

For the sake of comparative evaluation, we carried out an additional training exercise with the same CNN classifier, using these 16 X310 devices. This time, however, no artificial hardware impairments were introduced. The result of this experiment was far from satisfactory. The classification accuracy for these bit- similar radios stood at a mere 35.96%, a stark contrast to the outcomes from the previous experiments. This divergence in results underscores the pivotal role and benefits of a judicious impairment allocation process, further highlighting the remarkable performance of our CNN model when these impairments are present.

To encapsulate, these findings provide solid evidence of the robust performance of our CNN model across a range of settings and under different conditions. While there is always room for further enhancements, especially in dealing with bit- similar radios that have not been subjected to artificial impairments, our CNN model has demonstrated its potential and viability. The results strongly indicate the promising prospects of applying CNNs for the complex task of classifying radios in a multitude of real-world scenarios.

6.3 Reduced BER with Heuristic Impairments Selection

In our study, we utilized the metric of average total sum of Bit Error Rate (BER) of all the transmitters to rigorously evaluate the performance of our convolutional neural network (CNN) classifier. This approach allowed us to generate a robust benchmark against which we compared two distinctive impairment allocation strategies: i) random allocation, and ii) a greedy heuristic algorithm.

The structure of the experimental setup considered radios (R) with counts of 4, 8, 12, and 16. Each radio was randomly assigned average Signal-to-Noise Ratio (SNR) values from the set {20,25,30} dB. We only introduced a single type of impairment, which was IQ imbalance, bound by an Image Rejection Ratio (IMMR) value of -13.5 dB. Despite this,

we took into account a spectrum of 16 available impairment levels that ranged from an IMMR of -13.5 dB to -21 dB, with each successive level being separated by a margin of 0.5 dB. A crucial aspect of each selection step was the assurance that our CNN classifier could accurately classify these impairment levels with an impressive accuracy of over 95%.

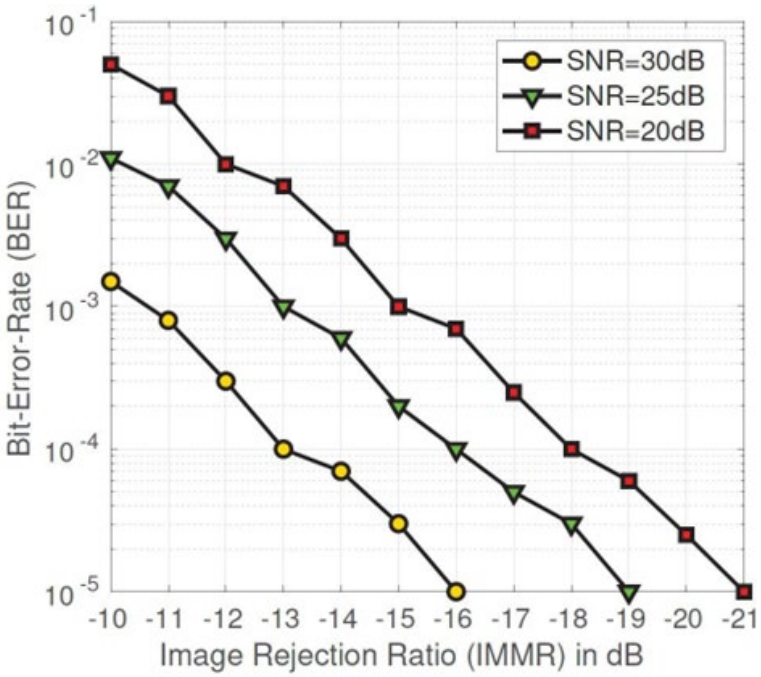

Figure 6.2: BER vs. IMMR value of IQ imbalance

For the random allocation approach, each of the R radios was randomly allocated one out of the 16 selected impairment levels. In contrast, our proposed greedy heuristic algorithm was designed to iteratively assign the lowest available impairment level to the radio that demonstrated the least average SNR level. The BER value for each radio was computed across a variety of SNR levels, as visualized in Figure 6.2.

The experimental process was executed over 1000 iterations. In each iteration, a random SNR level was assigned to each radio. Simultaneously, a unique impairment level was randomly allocated to each radio using the random allocation strategy, and this procedure was repeated for a total of 500 times. This facilitated the computation of the total sum of BER for all radios, averaged over the 500 iterations for the given SNR assignment. The obtained values were then averaged once more over 1000 SNR assignments.

In parallel, we applied a similar process to compute the total sum of BER for all radios, using the greedy heuristic algorithm, averaged over 1000 SNR assignments. The

results convincingly demonstrate that our method of allocating impairments consistently outperforms the random allocation approach in terms of the BER of all radios.

These results underscore the efficacy of our proposed method, highlighting the fact that a systematic and thoughtfully designed allocation of impairments can significantly enhance the performance of the CNN classifier in accurately identifying various radios. This consequently bolsters the robustness of our model, underscoring its potential utility in a variety of real-world applications.

6.4 Fulfillment of the Objective and Expected Outcome

Our research endeavors focused on the development and implementation of a CNN model specifically designed to detect intruders in Wi-Fi networks. Through rigorous experimentation and analysis, we have achieved remarkable results, demonstrating the effectiveness and accuracy of our model in identifying security breaches within the network.

By successfully detecting intruders with a high level of accuracy, our model significantly contributes to achieving our first objective of promptly recognizing when breaches are occurring. This capability allows network administrators to swiftly respond to potential threats, mitigating the risk of unauthorized access and potential data breaches.

Moreover, the provision of detailed information about the detected intruders plays a crucial role in reinforcing network security. With this valuable insight, organizations can strengthen their defense mechanisms, develop countermeasures, and enhance their overall security posture. The difficulty in altering RF traces, which serve as unique identifiers, poses a significant challenge for potential attackers, making it substantially more difficult for them to manipulate or tamper with the system. This achievement aligns with our second objective of increasing the difficulty and complexity of security breaches, thereby reducing the risk of data theft and unauthorized access.

Furthermore, the successful detection and subsequent blocking of intruders have a profound impact on the overall security of the network. By proactively identifying and eliminating potential threats, our model helps create a secure environment where sensitive data remains protected and private. This accomplishment addresses our third objective of enhancing the overall security of the network, ensuring the confidentiality, integrity, and availability of critical information.

In summary, the comprehensive fulfillment of our research objectives highlights the effectiveness and significance of our CNN model in detecting intruders, fortifying network security, and safeguarding data. By promptly recognizing breaches, making them more challenging to execute, and strengthening overall security measures, our research contributes to creating a resilient and secure Wi-Fi network environment.

Chapter 7

Conclusion

In this work, we have meticulously developed and evaluated a robust CNN architecture specifically tailored for the task of classifying radios under static conditions. This architecture has been constructed by adapting key principles from AlexNet, but with crucial modifications to ensure optimal performance for the task at hand. Our CNN model has been trained on raw time-series IQ samples generated from a 16-node USRP X310 Software Defined Radio (SDR) testbed and an external database of 140 Commercial Off-The-Shelf (COTS) Wi-Fi devices. Through our evaluation, we have demonstrated that our CNN architecture can successfully identify devices with impressive accuracy in static channel conditions.

However, we also observed that the performance of the classifier could be severely degraded under the influence of changing channel conditions due to phenomena such as multipath reflection and fading. We have shown how this channel state impacts the distribution of complex symbols captured by the receiver, influencing the classifier's ability to correctly identify devices. As such, we've stressed the importance of considering these dynamic conditions when training and testing the CNN classifier.

We further scrutinized our classifier's performance by experimenting with a variety of unique impairments to test the resilience of our model. A rigorous evaluation was performed where we intentionally introduced various types of impairments to challenge our CNN's ability to accurately identify bit-similar radios. These experiments provided evidence of our classifier's robustness, maintaining a high accuracy even under such challenging conditions.

Beyond the performance of the classifier, we also explored different strategies of allocating impairments to the radios in order to optimize the performance of our model. We compared a random allocation approach to a more calculated greedy heuristic algorithm that assigns impairments based on each radio's average SNR level. Our results demonstrated that our proposed allocation method significantly outperformed the random allocation strategy in terms of the Bit Error Rate (BER) of all radios.

In conclusion, our research presents a promising direction for the use of CNN architectures in the task of classifying radios under varying conditions. Despite the challenges presented by dynamic channel conditions and device similarities, our method of careful allocation of impairments proved successful in enhancing the performance of the classifier. While this work represents a significant step forward, it also reveals potential areas of future investigation, such as the development of more sophisticated allocation strategies and the adaptation of the CNN architecture for more complex radio environments. Our research findings underscore the potential of machine learning techniques in enhancing our ability to accurately identify devices, thereby opening the door for more secure and efficient communication networks in the future.

7.1 Limitation of the Research

While conducting our research, we encountered certain limitations that are important to acknowledge. Firstly, it is crucial to note that our study was conducted within a simulated environment. We collected, processed, and analyzed all data using software simulations, which may introduce some degree of discrepancy when compared to real-world scenarios.

Additionally, a challenge we faced was the unavailability of the necessary hardware devices, specifically the Ettus USRP SDR devices, in our country. These devices, which are essential for transmitting radio signals, proved to be quite expensive and inaccessible for our research purposes. As a result, we relied solely on software simulations for our experiments.

Furthermore, we faced difficulties in obtaining the required license from the telecom regulatory authority to transmit wireless signals. This regulatory requirement posed a significant obstacle due to the complex procedures and strict regulations involved. Despite our best efforts, we were unable to secure the necessary license given our circumstances.

It is essential to acknowledge these limitations as they may impact the generalizability and practical applicability of our research findings. However, we mitigated these challenges to the best of our abilities, utilizing software simulations as a reliable alternative and drawing insights from the available resources and data.

7.2 Future Prospects of Our Work

In this study, we have explored and demonstrated the potential of utilizing Convolutional Neural Networks (CNNs) to classify radios in a static environment based on raw IQ samples. We have highlighted the effectiveness of our model under specific configurations and settings. However, there remains an extensive scope for further exploration and improvement in this research area.

As we embark on future work, our primary objective will be to enhance the robustness and adaptability of the model to a variety of channel conditions and RF impairments. Furthermore, we aim to explore the possibilities of optimizing the neural network structure to better suit the specific requirements of this task. In the following sections, we delve into the specific aspects we intend to focus on in our future investigations.

7.2.1 Under Different Channel and RF Impairments

7.2.1.1 Multipath Profile

Our future work can investigate the effect of varying multipath profiles on the performance of the classifier. The multipath profile, including parameters such as path delays and average path gains, can significantly impact the received IQ samples and hence the classifier's accuracy. Exploring this aspect will further our understanding of how to design a robust classifier that can maintain high accuracy despite these variations in the multipath profile.

7.2.1.2 Channel Noise Level

Another dimension to explore is the impact of different channel noise levels on the classification accuracy. The Signal-to-Noise Ratio (SNR) significantly influences the quality of the received signals and can vary across different channels. A thorough examination of how the model performs under varying SNR levels would be crucial for further improving its resilience to such changes in the channel environment.

7.2.1.3 RF Impairments

Lastly, varying RF impairments could be introduced to test and further enhance the robustness of our model. This can include changes in the phase noise range, frequency offset range, and DC offset range. Observing how these impairments affect the performance of our model will help in improving its capability to adapt to such conditions.

7.2.2 Modify the Neural Network Structure

7.2.2.1 Convolutional Layer Parameters

The design of convolutional layers significantly impacts the performance of a CNN model. Future work could involve experimenting with different convolutional layer parameters such as the filter size, the number of filters, and padding. This would provide insights into how these parameters influence the performance of the CNN and could potentially lead to an optimized network structure.

7.2.2.2 Number of Fully Connected Layers

The number of fully connected layers in a CNN model can also be a subject for future exploration. Adding more fully connected layers can help the model learn more complex features, but it also comes with the risk of overfitting. By experimenting with different numbers of fully connected layers, we can determine the optimal architecture that maximizes the model's classification accuracy while avoiding overfitting.

7.2.2.3 Number of Convolutional Layers

Finally, altering the number of convolutional layers in the network can also impact the model's performance. Additional convolutional layers could allow the model to extract more intricate features from the data. Future studies can look into determining the optimal number of convolutional layers for this specific task, balancing the need for a deeper network with the computational resources available.

References

[1] A. Kumar Dalai, K. Kumar, and S. Kumar Jena, 'Wireless device authentication using fingerprinting technique', *Advances in Intelligent Systems and Computing*, vol. 707, pp. 163–172, 2019, doi: 10.1007/978-981-10-8639-7_17/COVER.

[2] B. W. Ramsey, T. D. Stubbs, B. E. Mullins, M. A. Temple, and M. A. Buckner, 'Wireless infrastructure protection using low-cost radio frequency fingerprinting receivers', *International Journal of Critical Infrastructure Protection*, vol. 8, pp. 27–39, Jan. 2015, doi: 10.1016/J.IJCIP.2014.11.002.

[3] B. Sieka, 'Using radio device fingerprinting for the detection of impersonation and Sybil attacks in wireless networks', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4357 LNCS, pp. 179–192, 2006, doi: 10.1007/11964254_16.

[4] A. K. Dalai and B. Sahoo, 'A Device Fingerprinting Technique to Authenticate End-user Devices in Wireless Networks', pp. 1–6, Feb. 2023, doi: 10.1109/ISSSC56467.2022.10051406.

[5] Y. Qin, B. Li, M. Yang, and Z. Yan, 'Attack Detection for Wireless Enterprise Network: A Machine Learning Approach', *2018 IEEE International Conference on Signal Processing, Communications and Computing, ICSPCC 2018*, Dec. 2018, doi: 10.1109/ICSPCC.2018.8567797.

[6] B. W. Ramsey, B. E. Mullins, M. A. Temple, and M. R. Grimaila, 'Wireless Intrusion Detection and Device Fingerprinting through Preamble Manipulation', *IEEE Trans Dependable Secure Comput*, vol. 12, no. 5, pp. 585–596, Sep. 2015, doi: 10.1109/TDSC.2014.2366455.

[7] A. Elmaghbub and B. Hamdaoui, 'Comprehensive RF Dataset Collection and Release: A Deep Learning-Based Device Fingerprinting Use Case', *2021 IEEE Globecom Workshops, GC Wkshps 2021 - Proceedings*, 2021, doi: 10.1109/GCWKSHPS52748.2021.9682024.

[8] A. C. Jose, R. Malekian, and N. Ye, 'Improving Home Automation Security; Integrating Device Fingerprinting into Smart Home', *IEEE Access*, vol. 4, pp. 5776–5787, 2016, doi: 10.1109/ACCESS.2016.2606478.

[9] S. Yin, *Q*. Li, and O. Gnawali, 'Interconnecting Wi-Fi devices with IEEE 802.15.4 devices without using a gateway', *Proceedings - IEEE International Conference on*

Distributed Computing in Sensor Systems, DCOSS 2015, pp. 127–136, Jul. 2015, doi: 10.1109/DCOSS.2015.42.

[10] T. Adame, A. Bel, B. Bellalta, J. Barcelo, and M. Oliver, 'IEEE 802.11AH: The Wi-Fi approach for M2M communications', *IEEE Wirel Commun*, vol. 21, no. 6, pp. 144–152, Dec. 2014, doi: 10.1109/MWC.2014.7000982.

[11] J. Hua, H. Sun, Z. Shen, Z. Qian, and S. Zhong, 'Accurate and Efficient Wireless Device Fingerprinting Using Channel State Information', *Proceedings - IEEE INFOCOM*, vol. 2018-April, pp. 1700–1708, Oct. 2018, doi: 10.1109/INFOCOM.2018.8485917.

[12] C. Gentner, M. Ulmschneider, *I.* Kuehner, and A. Dammann, 'Wi-Fi-RTT Indoor Positioning', *2020 IEEE/ION Position, Location and Navigation Symposium, PLANS 2020*, pp. 1029–1035, Apr. 2020, doi: 10.1109/PLANS46316.2020.9110232.

[13] *I.* Martin-Escalona and E. Zola, 'Ranging estimation error in Wi-Fi devices running IEEE 802.11mc', *2020 IEEE Global Communications Conference, GLOBECOM 2020 - Proceedings*, vol. 2020-January, Dec. 2020, doi: 10.1109/GLOBECOM42002.2020.9347973.

[14] M. Ayyash *et al.*, 'Coexistence of Wi-Fi and LiFi toward 5G: Concepts, opportunities, and challenges', *IEEE Communications Magazine*, vol. 54, no. 2, pp. 64–71, Feb. 2016, doi: 10.1109/MCOM.2016.7402263.

[15] J. Baranda, P. Henarejos, and C. G. Gavrincea, 'An SDR implementation of a visible light communication system based on the IEEE 802.15.7 standard', *2013 20th International Conference on Telecommunications, ICT 2013*, 2013, doi: 10.1109/ICTEL.2013.6632076.

[16] B. Bloessl, M. Segata, C. Sommer, and F. Dressler, 'Performance Assessment of IEEE 802.11p with an Open Source SDR-Based Prototype', *IEEE Trans Mob Comput*, vol. 17, no. 5, pp. 1162–1175, May 2018, doi: 10.1109/TMC.2017.2751474.

[17] M. Mishra, A. Potnis, P. Dwivedy, and S. K. Meena, 'Notice of Removal: Software defined radio based receivers using RTL - SDR: A review', *International Conference on Recent Innovations in Signal Processing and Embedded Systems, RISE 2017*, vol. 2018-January, pp. 62–65, Jun. 2018, doi: 10.1109/RISE.2017.8378125.

[18] J. Kim, S. Hyeon, and S. Choi, ‘Implementation of an SDR system using graphics processing unit’, *IEEE Communications Magazine*, vol. 48, no. 3, pp. 156–162, Mar. 2010, doi: 10.1109/MCOM.2010.5434388.

[19] F. D. Vaca and *Q*. Niyaz, ‘An ensemble learning based Wi-Fi network intrusion detection system (WNIDS)’, *NCA 2018 - 2018 IEEE 17th International Symposium on Network Computing and Applications*, Nov. 2018, doi: 10.1109/NCA.2018.8548315.

[20] H. Li, K. Gupta, C. Wang, N. Ghose, and B. Wang, ‘RadioNet: Robust Deep-Learning Based Radio Fingerprinting’, *2022 IEEE Conference on Communications and Network Security, CNS 2022*, pp. 190–198, 2022, doi: 10.1109/CNS56114.2022.9947255.

[21] D. Takahashi, Y. Xiao, Y. Zhang, P. Chatzimisios, and H. H. Chen, ‘IEEE 802.11 user fingerprinting and its applications for intrusion detection’, *Computers & Mathematics with Applications*, vol. 60, no. 2, pp. 307–318, Jul. 2010, doi: 10.1016/J.CAMWA.2010.01.002.

[22] H. C. Shin *et al.*, ‘Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning’, *IEEE Trans Med Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.

[23] M. E. Aminanto, R. Choi, H. C. Tanuwidjaja, P. D. Yoo, and K. Kim, ‘Deep abstraction and weighted feature selection for Wi-Fi impersonation detection’, *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 3, pp. 621–636, Oct. 2017, doi: 10.1109/TIFS.2017.2762828.

[24] W. C. Suski, M. A. Temple, M. J. Mendenhall, and R. F. Mills, ‘Radio frequency fingerprinting commercial communication devices to enhance electronic security’, *International Journal of Electronic Security and Digital Forensics*, vol. 1, no. 3, pp. 301–322, 2008, doi: 10.1504/IJESDF.2008.020946.

[25] K. Merchant, S. Revay, G. Stantchev, and B. Nousain, ‘Deep Learning for RF Device Fingerprinting in Cognitive Communication Networks’, *IEEE Journal on Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 160–167, Feb. 2018, doi: 10.1109/JSTSP.2018.2796446.

[26] V. L. L. Thing, ‘IEEE 802.11 network anomaly detection and attack classification: A deep learning approach’, *IEEE Wireless Communications and Networking Conference, WCNC*, May 2017, doi: 10.1109/WCNC.2017.7925567.

[27] Y. Yang, A. Hu, and J. Yu, ‘A practical radio frequency fingerprinting scheme for mobile phones identification’, *Physical Communication*, vol. 55, p. 101876, Dec. 2022, doi: 10.1016/J.PHYCOM.2022.101876.

[28] B. Li and E. Cetin, ‘Design and Evaluation of a Graphical Deep Learning Approach for RF Fingerprinting’, *IEEE Sens J*, vol. 21, no. 17, pp. 19462–19468, Sep. 2021, doi: 10.1109/JSEN.2021.3088137.

[29] S. Wang, L. Peng, H. Fu, A. Hu, and X. Zhou, ‘A convolutional neural network-based rf fingerprinting identification scheme for mobile phones’, *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2020*, pp. 115–120, Jul. 2020, doi: 10.1109/INFOCOMWKSHPS50562.2020.9163058.

[30] X. Guo, Z. Zhang, and J. Chang, ‘Survey of Mobile Device Authentication Methods Based on RF Fingerprint’, *INFOCOM 2019 - IEEE Conference on Computer Communications Workshops, INFOCOM WKSHPS 2019*, vol. 2019-January, Apr. 2019, doi: 10.1109/INFOCOMWKSHPS47286.2019.9093755.

[31] 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, Apr. 2020, doi: 10.1109/IOTM.0001.1900065.

[32] 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, Aug. 2021, doi: 10.1109/JSAC.2021.3087250.

[33] J. Ran, Y. Ji, and B. Tang, ‘A semi-supervised learning approach to IEEE 802.11 network anomaly detection’, *IEEE Vehicular Technology Conference*, vol. 2019-April, Apr. 2019, doi: 10.1109/VTCSPRING.2019.8746576.

[34] T. Gaber, A. El-Ghamry, and A. E. Hassanien, ‘Injection attack detection using machine learning for smart IoT applications’, *Physical Communication*, vol. 52, p. 101685, Jun. 2022, doi: 10.1016/J.PHYCOM.2022.101685.

[35] M. E. Aminanto and K. Kim, ‘Detecting impersonation attack in Wi-Fi networks using deep learning approach’, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10144 LNCS, pp. 136–147, 2017, doi: 10.1007/978-3-319-56549-1_12/COVER.

[36] K. Sankhe *et al.*, 'No Radio Left Behind: Radio Fingerprinting Through Deep Learning of Physical-Layer Hardware Impairments', *IEEE Trans Cogn Commun Netw*, vol. 6, no. 1, pp. 165–178, Mar. 2020, doi: 10.1109/TCCN.2019.2949308.

[37] A. Al-Shawabka *et al.*, 'Exposing the Fingerprint: Dissecting the Impact of the Wireless Channel on Radio Fingerprinting', *Proceedings - IEEE INFOCOM*, vol. 2020-July, pp. 646–655, Jul. 2020, doi: 10.1109/INFOCOM41043.2020.9155259.

[38] 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, Apr. 2020, doi: 10.1109/IOTM.0001.1900065.

[39] D. Nouichi, M. Abdelsalam, *Q*. Nasir, and S. Abbas, 'IoT Devices Security Using RF Fingerprinting', *2019 Advances in Science and Engineering Technology International Conferences, ASET 2019*, May 2019, doi: 10.1109/ICASET.2019.8714205.

[40] 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, doi: 10.1109/MCOM.2018.1800153.

[41] C. Cordeiro, D. Akhmetov, and M. Park, 'IEEE 802.11ad: Introduction and performance evaluation of the first multi-Gbps Wi-Fi technology', *Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM*, pp. 3–7, 2010, doi: 10.1145/1859964.1859968.

[42] S. Dhawan, 'Analogy of promising wireless technologies on different frequencies: Bluetooth, Wi-Fi, and WiMAX', *The 2nd International Conference on Wireless Broadband and Ultra Wideband Communications, AusWireless 2007*, p. 14, 2007, doi: 10.1109/AUSWIRELESS.2007.27.

[43] H. Jafari, O. Omotere, D. Adesina, H. H. Wu, and L. Qian, 'IoT Devices Fingerprinting Using Deep Learning', *Proceedings - IEEE Military Communications Conference MILCOM*, vol. 2019-October, pp. 901–906, Jan. 2019, doi: 10.1109/MILCOM.2018.8599826.

[44] H. Li, K. Gupta, C. Wang, N. Ghose, and B. Wang, 'RadioNet: Robust Deep-Learning Based Radio Fingerprinting', *2022 IEEE Conference on Communications and Network Security, CNS 2022*, pp. 190–198, 2022, doi: 10.1109/CNS56114.2022.9947255.

[45] K. Sankhe, M. Belgiovine, F. Zhou, S. Riyaz, S. Ioannidis, and K. Chowdhury, 'ORACLE: Optimized Radio clAssification through Convolutional neuraL

nEtworks', *Proceedings - IEEE INFOCOM*, vol. 2019-April, pp. 370–378, Apr. 2019, doi: 10.1109/INFOCOM.2019.8737463.

[46] *Q*. Duan, X. Wei, J. Fan, L. Yu, and Y. Hu, 'CNN-based Intrusion Classification for IEEE 802.11 Wireless Networks', *2020 IEEE 6th International Conference on Computer and Communications, ICCC 2020*, pp. 830–833, Dec. 2020, doi: 10.1109/ICCC51575.2020.9345293.

[47] R. Vishwakarma and R. Vennelakanti, 'CNN Model Tuning for Global Road Damage Detection', *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 5609–5615, Dec. 2020, doi: 10.1109/BIGDATA50022.2020.9377902.