UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

A study of modulation spoofing techniques in deep learning based signal classification

Diploma Thesis

Dimitrios Varkatzas

Supervisor: Antonios Argyriou

June 2023

UNIVERSITY OF THESSALY

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

A study of modulation spoofing techniques in deep learning based signal classification

Diploma Thesis

Dimitrios Varkatzas

Supervisor: Antonios Argyriou

June 2023

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

Μελέτη της επίπτωσης τεχνικών απόκρυψης διαμόρφωσης στην ταξινόμηση σημάτων με την χρήση βαθιάς μάθησης

Διπλωματική Εργασία

Δημήτριος Βαρκάτζας

Επιβλέπων: Αντώνιος Αργυρίου

Ιούνιος 2023

Approved by the Examination Committee:

Supervisor **Antonios Argyriou**
Associate Professor, Department of Electrical and Computer Engineering, University of Thessaly

Member **Gerasimos Potamianos**
Associate Professor, Department of Electrical and Computer Engineering, University of Thessaly

Member **Alexander Chroneos**
Professor, Department of Electrical and Computer Engineering, University of Thessaly

Acknowledgements

For Maria, who was there for me in every difficult moment I faced.

DISCLAIMER ON ACADEMIC ETHICS AND INTELLECTUAL PROPERTY RIGHTS

«Being fully aware of the implications of copyright laws, I expressly state that this diploma thesis, as well as the electronic files and source codes developed or modified in the course of this thesis, are solely the product of my personal work and do not infringe any rights of intellectual property, personality and personal data of third parties, do not contain work / contributions of third parties for which the permission of the authors / beneficiaries is required and are not a product of partial or complete plagiarism, while the sources used are limited to the bibliographic references only and meet the rules of scientific citing. The points where I have used ideas, text, files and / or sources of other authors are clearly mentioned in the text with the appropriate citation and the relevant complete reference is included in the bibliographic references section. I also declare that the results of the work have not been used to obtain another degree. I fully, individually and personally undertake all legal and administrative consequences that may arise in the event that it is proven, in the course of time, that this thesis or part of it does not belong to me because it is a product of plagiarism».

The declarant

Dimitrios Varkatzas

Diploma Thesis

A study of modulation spoofing techniques in deep learning based signal classification

Dimitrios Varkatzas

Abstract

This thesis investigates a new wireless signal obfuscating waveform that increases misclassifications of modulated wireless communication signals with Deep Learning (DL). Convolutional neural network (CNN) is a deep learning technique hat was used for modulation classification due to the very good performance. The proposed waveform is embedded in the transmitted signal in such a way that prevents the extraction of clean data for training a DL-based classifier but at the same time it allows a legitimate receiver (LRx) to demodulate the data. The proposed scheme works for both single carrier and multi-carrier orthogonal frequency division multiplexing (OFDM) waveforms and can be implemented as part of frame-based communication protocols. The effect of a variety of spoofing signal parameters values on the accuracy of the classifier will be investigated for different sorts of channels. For AWGN and fading channels, it was determined that a wide range of values for these parameters substantially affects the accuracy of the classifier. As for the fading channel with partial spoofing knowledge, the obfuscating waveform can drop classification performance at the eavesdropper to less than 10%, while there is a slightly performance loss at the LRx.

Keywords:

Deep learning, Convolutional Neural Network, Modulation Classification , OFDM, obfuscation

Διπλωματική Εργασία

Μελέτη της επίπτωσης τεχνικών απόκρυψης διαμόρφωσης στην ταξινόμηση σημάτων με την χρήση βαθιάς μάθησης

Δημήτριος Βαρκάτζας

Περίληψη

Η παρούσα εργασία διερευνά μια νέα κυματομορφή απόκρυψης ασύρματου σήματος που αυξάνει τις λανθασμένες ταξινομήσεις διαμορφωμένων σημάτων ασύρματης επικοινωνίας με Βαθιά Μάθηση (ΒΘ). Το συνελικτικό νευρωνικό δίκτυο (ΣΝΝ) είναι η δομή βαθιάς μάθησης που χρησιμοποιήθηκε λόγω των προηγμένων επιδόσεών της, το οποίο έχει εκπαιδευτεί για την κατηγοριοποίηση των δεδομένων με βάση τη διαμόρφωση των σημάτων. Η προτεινόμενη κυματομορφή ενσωματώνεται στο μεταδιδόμενο σήμα με τέτοιο τρόπο που αποτρέπει την εξαγωγή καθαρών δεδομένων για την εκπαίδευση ενός ταξινομητή που βασίζεται σε ΒΘ, αλλά ταυτόχρονα επιτρέπει σε έναν νόμιμο δέκτη (LRx) να αποδιαμορφώσει τα δεδομένα. Το προτεινόμενο μοντέλο λειτουργεί τόσο για κυματομορφές ενός φορέα όσο και για κυματομορφές Ορθογωνικής Πολυπλεξίας Διαίρεσης Συχνότητας (ΟΠΔΣ) πολλαπλών φορέων και μπορεί να υλοποιηθεί ως μέρος πρωτοκόλλων επικοινωνίας που βασίζονται σε frames. Θα διερευνηθεί η επίδραση ποικίλων τιμών παραμέτρων σήματος υποκλοπής στην ακρίβεια του ταξινομητή για διαφορετικά είδη καναλιών. Για τα κανάλια AWGN και εξασθένισης, διαπιστώθηκε ότι ένα ευρύ φάσμα τιμών για αυτές τις παραμέτρους διαβρώνει σημαντικά την ακρίβεια του ταξινομητή. Όσον αφορά το κανάλι εξασθένισης με μερική γνώση του σήματος απόκρυψης, η κυματομορφή αυτή μπορεί να ρίξει την απόδοση ταξινόμησης στον υποκλοπέα σε λιγότερο από 10%, ενώ υπάρχει μια ελαφριά απώλεια απόδοσης στον LRx.

Λέξεις-κλειδιά:

Βαθιά Μάθηση, Συνελικτικά Νευρωνικά Δίκτυα, Ταξινόμηση σημάτων, Ορθογωνική Πολυπλεξία Διαίρεσης Συχνότητας, Σήμα απόκρυψης

Table of contents

List of figures

List of tables

Abbreviations

DL	Deep Learning
CNN	Convolutional Neural Network
NNs	Neural Networks
AMC	Automatic Modulation Classification
OFDM	Orthogonal Frequency-Division Multiplexing
LRx	Legitimate receiver
AWGN	Additive White Gaussian Noise
SNR	Signal to Noise Ratio
DSA	Dynamic Spectrum Access
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent
ISI	Intersymbol Interference
PSK	Phase-Shift Keying
QAM	Quadrature Amplitude Modulation
PAM	Pulse-Amplitude Modulation
GFSK	Gaussian frequency-shift keying
B-FM	Broadcat Frequency Modulation
DSB-AM	Double-Sideband Amplitude Modulation
SSB-AM	Single-Sideband Amplitude Modulation

Chapter 1

Introduction

Automatic Modulation Classification (AMC) of communication signals determines the type of modulation used in the received signal and sets the stage for further signal processing. It is significant for both military or civil applications. As communication technology advances quickly, communication systems frequently use a variety of modulation schemes, making it harder and harder to use AMC. Due to its superior performance, Deep Learning (DL) has been effectively used in many different domains, including computer vision, natural language processing, speech recognition, and many more. Neural Networks (NNs) are the foundation of DL technology, which is based on the idea of how the neural system operates. In addition to training the neural network, we must test it in order to determine its performance. There could be multiple malevolent users on the network attempting to intercept the data we transmit. As a result, data protection is of utmost importance and as we will see it has a significant impact on the performance of the classifier.

1.1 Subject of thesis

The goal of this thesis is to investigate in detail how an obfuscating signal can be implemented in a practical bursty communication protocol like 802.11, how it affects the spectrograms, the classification performance, and finally Bit Error Rate (BER) for various channels. We will employ two datasets in the classifier, with the first one referring to single-carrier systems and the second one dealing with multi-carrier systems including signals like OFDM-based ones. Each of the aforementioned systems is known to be impacted uniquely by the channels under consideration.

1.1.1 Contribution

The advantages and contributions of our research in this thesis are stated as follows:

1. We propose the first method for self-jamming (or signal obfuscation) that can be implemented as part of frame-based wireless communication and prevents AMC with Deep Learning at an arbitrary receiver, but is robust to bit decoding at a legitimate receiver. The method is independent of the used modulation, i.e. it masks features in the resulting spectrogram.

2. We present and open-source a new dataset that contains obfuscated modulated signals and can be used by the community for evaluating newer DL-based signal classification algorithms.

1.2 Organization

The remainder of the thesis is organized as follows: In Chapter 2, a background is given to the reader so they may more readily understand the issue of the thesis. Chapter 3 lists the references that will be mentioned in the article and provides a brief explanation of how to cite them. The signal models, NN architecture, and the unique hyper-parameter settings that we used to obtain high accuracy rates of the classifier are all covered in depth in Chapter 4, which is described in one brief paragraph how each source will be referenced in the document that will be utilized. The results for three distinct channels are presented in Chapter 5 along with an effort to address the thesis issue. Finally, in Chapter 6 we summarise the results of the thesis and describe the conclusions drawn, both positive and negative. Some ideas for the extension of the diploma are also given.

Chapter 2

Background

2.1 Introduction

There has been lots of interest in Deep Learning (DL) in recent years which has been used in many applications. It has successfully resolved very challenging issues in a great variety of scientific domains. In the early phases of DL, researchers have started using it for image processing. Self-driving cars are one of the most well-known primary examples of its' application. To be more precise, these sophisticated vehicles are equipped with video cameras and they are able to analyze the live video feed from those cameras, determining road lines, observing any nearby vehicles or pedestrians on the road and eventually being driven by themselves. However, DL also applies to signal processing, where it has been demonstrated in audio applications for instance. In particular, some mobile phones have the Siri app installed and by saying "Hey Siri", they can recognize the voice of user, process it, decipher what the user is requesting, and then fulfill any request. More recently, we have also started to see wireless communications applications that incorporate DL, in which numerous studies have been conducted.

2.2 Modulation Classification

One of the primary uses of DL in wireless communications is Automatic Modulation Classification (AMC). The issue is described as follows: we have a receiver for which we don't know what type of modulation uses and a transmitter that it is attempting to determine what kind of modulation receives with the aid of DL. In essence, AMC is defined as the pro-

cess of blindly identifying the modulation scheme of a received signal. AMC is critical in this field of study, because it has a vital role for monitoring, managing, and controlling communications systems in order to address technical difficulties such as spectrum awareness, adaptive transmissions, and interference avoidance.

Taking an in-depth look at the applications of AMC, one obvious implementation is cognitive radio. More specifically, the major objective of cognitive radio is to detect activity in the RF band and make the most of it. By using AMC we can determine which modulation types are present in this spectrum as well as various properties for network users. Spectrum surveillance and management is also another basic application for AMC. To be more precise, not all spectra support the same types of modulation. A few might be blocked by the network. So, to accomplish network security, it is necessary to identify and impede any unwanted transmitters as soon as they are discovered. Last implementation is signals intelligence. It is important to gather data from different modulated signals and analyze the data they contain. We can identify transmitters that emit specific modulation types and get a clear picture of what it is going on in this network with the aid of AMC.

It has to be stated that if everything was ideal we wouldn't need any kind of intelligent system to tell us what kind of modulation type is airing on the network. But things aren't perfect in the real world. Consequently, there are a number of unidentified impairments in the channels where AMC is used.

2.3 Deep Learning

Because of its' superiority in feature extraction and classification accuracy, DL based AMC has lately gained prominence. One major issue in DL based AMC is accurately preprocessing and representing a received signal before feeding it into Deep Neural Networks (DNNs).

Before we get into the architecture of a DL network, let's first define it. On the substance, DL is a machine learning technique that trains computers to gain knowledge by example. A computer model learns to execute classification tasks by analyzing images, text, or sound in DL. DL models can attain cutting-edge accuracy, sometimes outperforming humans. Models are trained utilizing a huge quantity of labeled data and multi-layered Neural Network (NN) architectures. This explains its significance, given the high accuracy values it achieves. First

and foremost, significant quantities of labeled data are required. For instance, developing a driverless car necessitates millions of photos and thousands of hours of video. Secondly, DL demands a significant amount of processing power. High-performance GPUs have an efficient parallel architecture for DL. When paired with clusters or cloud computing, this allows development teams to cut DNN training time from weeks to hours or less.

Furthermore, the following question rises: how does a DNN function? Because most DL approaches employ NN structures, DL models are frequently referred to as DNNs. The term ”deep” is commonly used to refer to the amount of hidden layers in a NN. Traditional NNs have only 2-3 hidden layers, whereas DNNs can have up to 150. DL models are taught utilizing massive amounts of labeled data and NN layouts that learn features directly from the data, eliminating the requirement for manual feature extraction.

We shall utilize Convolutional Neural Network (CNN) as the DNN in this thesis, while it is one of the most well-known in this industry. A CNN convolves learned features with input data, making this architecture well suited to processing 2D data such as photos. CNNs reduce the requirement for manual feature extraction, so one doesn’t have to identify image-classification features. The CNN extracts features from pictures directly. The essential characteristics are not pre-trained; instead, they are learned as the network trains on a set of images. DL models are highly accurate for computer vision applications such as object classification thanks to automatic feature extraction. CNNs learn to recognize various visual aspects by utilizing tens or hundreds of hidden layers. Every hidden layer adds to the complexity of the visual features that have been learned. For example, the first hidden layer could learn to detect edges, while the final learns to detect more complicated forms that are specific to the shape of the object we’re attempting to recognize.

These basic procedures are followed each time to effectively create the DNN. To begin, it is self-evident that the network need data access in order to function properly. This information could be stored in a database, on the hard drive of a computer , or even through sensors spread around the network. After the data is acquired, it is pre-processed and an attempt is made to extract features. Data can be pre-processed in a variety of ways. Denoising and smoothing the data, for example, is one method. In addition, we may need to minimize the data or convert it to a format that is more suitable for conducting the experiment. In order to extract features, a new and smaller set of features is created, which captures most of the useful information of raw data. Of course, this increases the training of classifier and

inference speed. Then, predictive models are developed. This model has parameters, and it requires extensive testing with various system settings before it is validated. The network is ready for deployment once the performance is satisfactory. There are several methods for implementing it, like creating a desktop application or embedding generated code into the system.

Finally, wireless communications are fitted differently in this procedure. We now have an advantage, because synthetic wireless signals are man-made and they can be simulated. This suggests that data collection is less difficult and time-consuming. For validation, software-defined radio (SDR) is implemented.

Chapter 3

Related Work

3.1 Related Work on Modulation Classification

AMC (Automatic Modulation Classification) is a fundamental approach in noncooperative communication systems [1],[2],[3],[4],[5],[6]. It has been widely studied in the past 20 years and finds applications in various commercial and military areas. Deep Learning (DL) has been applied in many fields such as image classification and natural language processing. The recent success of DL algorithms is associated with applications that suffer from inaccuracies in existing mathematical models and enjoy the availability of large data sets. Recently, the idea of DL has been introduced for AMC using a variety of different Neural Networks (NNs) [7], [8]. Also, extensive research has been done on different types of channels, such as fading channels [6],[9] and even channels that are equalized [4],[10].

This article will employ two datasets: the well-known RadioML2016.10a and a variation of the first, which contains QAM signals and being subjected to OFDM. We do not want to limit the analysis to single-carrier systems, but the classifier is not advanced enough to simultaneously distinguish between OFDM signals and single-carrier modulations, hence the separation of the two groups of data. The article [1] describes an advanced DL-based AMC system that uses fast Fourier transformation Window Bank (FWB) and in-phase and quadrature-phase signals to classify OFDM symbol length and single-carrier modulation methods at the same time. Future work may examine the effect of malicious signals on the accuracy of this classifier. It has to be addressed that the receiver performs OFDM demodulation for the second dataset in order for the classifier to be able to distinguish various QAM modulations.

The works in [2] and [7] implies that Convolutional Neural Networks (CNN) has been

recently identified as a powerful tool in image classification and voice signal processing. Simulation results show that CNN not only demonstrates better accuracy results, but also provides more flexibility to learn features across a wide range of tasks and demonstrates improved classification accuracy against current day approaches. It is also compared to other DL architectures.

Furthermore, AMC can be implemented with various methods that provide a variety of benefits [3],[5],[6],[8],[9]. Important information on the layers of the CNN to be used in this research can be found in the works of [3] and [5]. The many types of layers are thoroughly examined, along with the functions that each one serves and the connections that exist between them. The NN structure is expected to play a significant effect in the accuracy of the classifier, which is over 96% for both datasets. These thesiss consider DL architectures for AMC, where [5] mainly focuses on CNNs. The work in [6], a hierarchical polynomial classifier is used for AMC, which uses Higher Order Cumulants (HOCs). The work in [8] implies that CNNs are capable in identifying emitters to estimate the IQ imbalance parameters of each emitter, using only raw IQ data as input. AMC is an integral part of spectrum enforcement, which is one of the main goals of Electrosence. As for [9], it presents results for CNN-based SEI using seven RF-DNA fingerprinting scenarios using OFDM-based waveforms that underwent Rayleigh fading followed by waveform-level channel estimation and equalization.

It has to be pointed out that in this thesis will not take into consideration only AWGN channels, but it will be also expanded into fading channels, such as flat-fading or frequency-selective channels. The study [6] provides helpful information regarding fading channels that will be utilized to help the reader understand their impact on both single-carrier systems and multi-carrier systems. Additionally, the work of [9] uses a Rayleigh Fading and proved that it is quite useful for multi-carrier systems, because OFDM provides the benefit of high spectrum utilization and anti-multipath fading and can be widely used in various commnunications systems. So, it is feasible to construct effective systems for distinguishing the modulation patterns of OFDM signal subcarriers.

Also, channel equalization has a significant impact on the accuracy of the classifier as [4],[9] ,[10]. The main type of equalization to be used in this thesis is Maximum Likelihood Sequence Estimation (MLSE), which is implemented using Viterbi Algorithm. It has to be pointed out that not only channel equalization compensates basic imperfections of the receiver, but also a portion of the spoofing signal. We suppose that the system is equipped with

an estimator that can identify a percentage of the obfuscated signal, which will be considered by the receiver as part of the channel.

3.2 Related Work on Signal Obfuscation

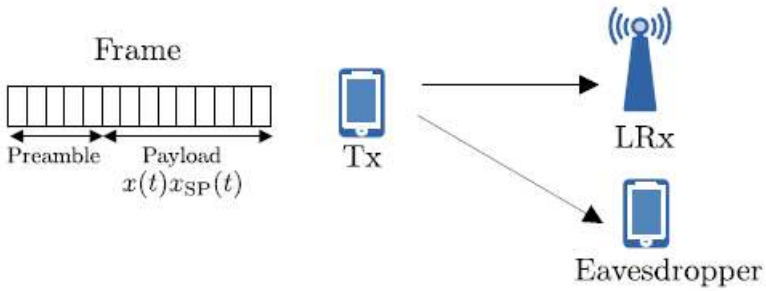

Figure 3.1: Representative topology of a legitimate receiver and an eavesdropper that tries to classify the received communication signals. The transmitter emits only the symbols $x(t)$ of the payload mixed with an obfuscating waveform $x_{sp}(t)$.

Recent work on adversarial deep learning focuses on understanding how the learning process can be affected by providing malicious input to the deep network [11]. In this thesis our aim is precisely that but for the problem of modulation classification. In particular we seek ways to prevent the effective operation of a DL-based AMC algorithm at a wireless receiver. The reason we would want to do that is that this particular receiver might be an eavesdropper. Figure 3.1 illustrates the scenario we consider: A legitimate receiver (LRx) of the transmitted wireless communication signal should receive the modulated symbols $x(t)$ and should not experience significant performance loss. The eavesdropper should not be able to learn the used modulation in $x(t)$. We propose to insert a spoofing signal $x_{sp}(t)$ only in the payload of the transmitted frame so that when the eavesdropper receives it cannot estimate $x_{sp}(t)$ from the preamble (by treating it as a channel impairment). Regarding the LRx we explore two cases, one that it does not know $x_{sp}(t)$ (and so it experiences some demodulation performance loss) and one case where it knows it (e.g. with header encryption as we discuss below and also proposed in [12]). This scenario is useful when the transmitter does not desire AMC by an arbitrary eavesdropper of the signal. One example is out of privacy concerns, or to avoid fingerprinting (something invaluable in military communications).

Hiding the modulation, also called (modulation spoofing or obfuscation) has been looked at the literature. In [12],[13],[14],[15] the authors presented methods to obfuscate quadrature

amplitude modulated (QAM) symbols at the transmitter. The authors of these works may flip the symbol location in the constellation diagram [13], embed symbols from a lower order modulation to a higher order constellation [12], or reorder the symbols in time [14]. Then they detect the symbols of the modified constellation which means that the symbol detection probability is dictated by the distance between the constellation points in the new mapping. However, these methods produce QAM symbols that can still be classified by an AMC system. Header encryption methods are proposed in [12] so that metadata are not readily available to an arbitrary receiver. Other works like [16], [17] randomize the value of a received modulated signal by allowing intentional RF jamming of the complete transmitted signal from another source. But this systems require a synchronized third party.

Recent work on adversarial deep learning focuses on understanding how the learning process can be affected by providing malicious input to the deep network [11]. DL-based AMC has also been investigated in the context of adversarial learning but all the works focus on deriving the theoretical perturbation necessary to yield a lower classification performance [18], [19], without considering how to deliver this perturbation to the eavesdropper (e.g. with a protocol) when a legitimate receiver is also present. In this work we are interested in designing a practical signal for a realistic scenario where both a legitimate receiver and an eavesdropper will receive the desired signal due to the broadcast medium but with the added perturbation. Hence, we will focus on signals that are easily recoverable at the LRx. This class of signals should not include any amplitude variations (e.g. act as a wireless fade) but only time-dependent phase variations. This class of signals was proposed in [20] with the purpose of smearing the resulting spectrogram of a wireless digital communication signal. However, that work was not concerned with modulation obfuscation and the performance of the communication subsystem.

Chapter 4

Proposed Idea

4.1 Baseline Dataset for Training

The data generation for the neural network will be the starting point for the analysis of the system built for this research. This intelligent network will be employed for each channel evaluated in this project. Two separate datasets will be used for this project, the first of which is RadioML2016.10a and contains 11 different types of modulations—eight digital and three analog—in total. These include B-FM, DSB-AM, and SSB-AM for analog modulations, as well as BPSK, QPSK, 8PSK, QAM16, QAM64, BFSK, CPFSK, and PAM4 for digital modulations. With a normalized average transmit power of 0 dB, data is modulated at a rate of about 8 samples per symbol. The full dataset was created using GNU radio as a 1024-sample complex time-domain vector. Through 1024-sample rectangular windowing processing, samples for both datasets are separated into training, validation, and testing. The training examples - each consisting of 1024 samples - are fed into the neural network in 2*1024 vectors with real and imaginary parts separated in complex time samples. The second dataset is actually a variation of the first one. It contains QAM modulations, which are then subjected to OFDM. It should be emphasized that demodulation is performed on these signals by the receiver because the system is not yet sufficiently developed to distinguish the modulation of OFDM-based signals. [1] provides a full description of how this difficulty is overcome. These two datasets are obviously required since the various channels and spoofing signals under consideration have distinct effects on these signals.

4.2 Problem, Basic Idea and Eavesdropper Assumptions

We want to prevent powerful DL techniques from classifying the modulation type used in a frame-based wireless communication signals. We propose to do that by polluting the data used for classification in a random way that has not been previously seen during training and it also does not allow the data to exhibit unique features across different modulations. We also propose to do that in as subset of the transmitted frame as discussed next. In this way the DL algorithm will not be able to efficiently classify the signals.

With the proposed system the transmitter emits an altered waveform that contains an obfuscating signal. In particular it does two things: 1) the obfuscating signal $x_{sp}(t)$ does not smear the preamble of the PHY frame that the emitter transmits but only the payload. The reason is that the preamble is typically a known bitstream even at the eavesdropper. So it is used for channel estimation. If it is altered with the spoofing waveform then the eavesdropper will be able to estimate from it the aggregate effect of the channel fading plus obfuscation with $x_{sp}(t)$, and so it can remove it from the data part of the frame. By not doing so, the eavesdropper will try to classify the data part of the frame that is polluted with $x_{sp}(t)$. 2) We can employ a header encryption method like the one in [12] so that parameters of $x_{sp}(t)$ are not readily available to an arbitrary eavesdropper but only to the LRx.

The eavesdropper is assumed to operate in a way that is consistent with its capabilities as an eavesdropper, that is it is not part of the communication network. However, it is reasonable to assume that it knows for example the carrier frequency so that it downconverts any passband signal. This information is trivial to acquire, since it only requires knowledge of the PHY protocol in use (e.g. WiFi in 2.4GHz or 5GHz, etc). Hence, AMC that we discuss later will take place from baseband samples, consistent with the literature we discussed.

4.3 Obfuscated Signal Generation for the Proposed Dataset

Now we are going to have a look on how these signals are generated [2]. One of the most important sensing tasks in Dynamic Spectrum Access (DSA) is alerting users to adjacent emitters in order to prevent radio interference and optimize spectrum allocation. This distinguishing and recognizing sources of potential radio interference in the area, such as radar users, local and wide area data and voice radios, broadcast radios, and other sources, each of which has unique behaviors and needs. In order to determine the type of communica-

tions scheme and emitter present, modulation recognition is the process of categorizing the modulation type of a received radio signal. Given that the received signal is represented as a complex base-band time series, this can be thought of as an N-class decision problem. To get a 1xN complex valued vector, we sample the in-phase and quadrature components of a radio signal at discrete time steps using an analog to digital converter with a carrier frequency that is generally centered on the carrier of interest. This is traditionally expressed as in the following equation, where s(t) is a time series signal of either a continuous signal or a sequence of discrete bits modulated onto a sinusoid with either varying frequency, phase, amplitude, trajectory, or some permutation of multiple thereof. c is some path loss or constant gain term on the signal, n(t) is an additive Gaussian white noise process reflecting thermal noise, and $x_{sp}(t)$ is the spoofed signal [2].

$$r(t) = (x_{sp}(t) \cdot s(t)) * c + n(t)$$

Although the following equation depicts the real-world relationship considerably more accurately, it is often utilized analytically in the formulation of expert features and decision statistics.

$$r(t) = \int_{\tau=0}^{\tau_0} e^{j\frac{\delta f}{f_m} sin(2\pi f_m(t-\tau)} s(t-\tau)h(\tau)d\tau + n_{Add}(t)$$

This takes into account a number of real-world effects that are difficult to predict, such as convolution with a time varying rotating non-constant amplitude channel impulse response h(t − τ), the addition of noise which may not be white, $n_{Add}(t)$, and f_m, δf representing the parameters of the spoofing signal. Each presents an unknown time varying source of error.

4.3.1 Obfuscating Signal

The transmitted signal x(t) contains an oscillating sinewave with a frequency of a few tens of Hz and was added by [20], which distorts the spectrograms of the generated signals. $x_{sp}(t)$ in the time domain is defined to be an frequency-modulated (FM) waveform with maximum instantaneous frequency shift δf, and frequency f_m, making thus the instantaneous frequency equal $f_i(t) = f_c + \delta f cos(2\pi f_m t)$. As a result, the signal generated by this waveform oscillates in the frequency domain between the maximum instantaneous frequencies -δf and +δf at a rate of f_m Hz. Therefore, the fluctuation of the frequency f_m should be greater than this

amount, meaning that even a few tens of Hz for f_m will greatly smear the spectrogram of the received signal.

4.3.2 OFDM Dataset

The signals for the second dataset are first subjected to QAM modulation and then to OFDM modulation. With N subcarriers that are spaced relative to the carrier fc at locations $f_k = k\Delta f$ Hz that can contain data, pilot symbols, or a combination of both, the desired baseband OFDM symbol in continuous time based in [20] is:

$$s(t) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} S[k] e^{j2\pi k \Delta f t},\ 0 \leq t \leq T_N$$

S[k] is the complex QAM symbol modulated onto subcarrier k, and $T_N = N\Delta f$ is the OFDM symbol duration.

4.4 Spoofed Dataset Characteristics

Our main method for learning is a Convolutional Neural Network (CNN), which receives a windowed input of the unprocessed radio time series r(t). We employ r(t) as a set of 2xN vectors into a narrow 2D Convolutional Network, where the orthogonal synchronously sampled In-Phase and Quadrature (I & Q) samples make up this 2-wide dimension, and we treat the complex valued input as an input dimension of 2 real valued inputs.

Radio communications provide a particular instance even though simulation and the use of synthetic datasets for learning are occasionally frowned upon in machine learning. Radio communication signals are actually created artificially, and we do it deterministically in a way that is identical to a real system, adding modulation, pulse shaping, carried data, and other well-characterized transmit parameters that are identical to a real-world signal. To verify that the bits in digital modulation are equiprobable, we whiten the data using a block randomizer. Impacts on radio channels can be represented fairly well. We employ trustworthy models for signals such as spoofing, time-varying multipath fading of the channel impulse response, [6], [9], additive Gaussian white noise, and channel equalization. The first dataset was created in GNU Radio using features from the Communications Toolbox (Matlab), and each time series signal was then divided into a training, validation, and test set using a rectangle windowing process with 1024 samples. The entire dataset consists of roughly 1.25 GB of frames, which

are Matlab files holding intricate bit floating point samples. The second dataset is actually a variation of the first one, in which only QAM symbols are used and endure OFDM modulation. Then, the signals are once reprocessed and the receiver performs OFDM demodulation on them.

We can visually distinguish a number of similarities and differences between modulations by examining a single sample from each class of modulation in the time and frequency domain, but a human expert would not be able to easily differentiate them all due to pulse shaping, distortion, and other channel effects.

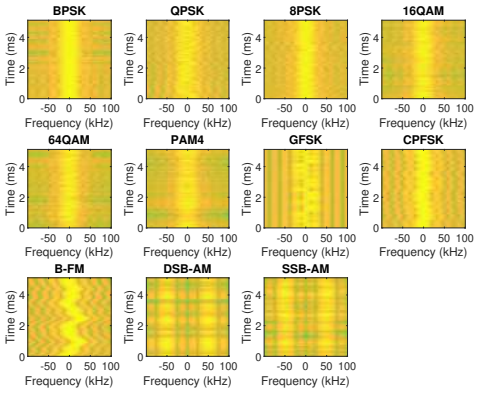

Figure 4.1: Spectrograms

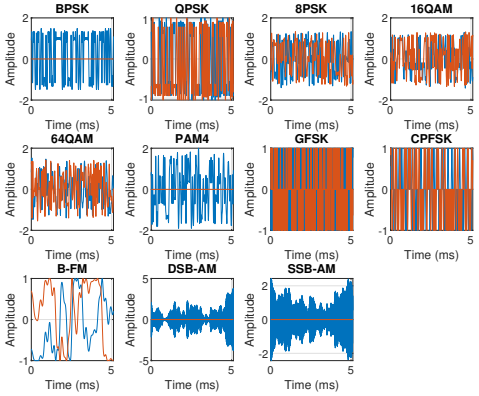

Figure 4.2: I & Q parts of example frames

In radio communications, signals are typically comprised of a number of modulated data bits on well defined and understood basis functions into discrete modes formed by these bases. Complex baseband representation of a signal decomposes a radio voltage level time-series into its projections onto the sine and cosine functions at a carrier frequency. Data bits are then modulated into this space through discrete and separable modes for each distinct symbol period in time in the case of digital modulation, or continuous location in the case of analog modulation, by adjusting the frequency, amplitude, phase, or sum thereof. The equation below illustrates this phase-mapping for the QPSK scenario.

$$y(t_i) = e^{j2\pi f_c t + \pi \frac{2c_i+1}{4}}, c_i \in 0, 1, 2, 3$$

Pulse shaping filters such as root-raised cosine are then typically applied to band-limit the signal in frequency and remove sharp wide-band transients between these distinct modes, resulting in mixing of adjacent symbols' bases at the transmitter in a deterministic and invertible. In our simulated data set we use a root-raised cosine pulse shaping filter with an excess bandwidth of 0.35 for each digital signal.

In contrast, channel effects in a communications system are not fully invertible and deterministic. The sent signal is subjected to a variety of influences in real systems, which makes its recovery and representation difficult. At the receiver, thermal noise produces relatively flat white Gaussian noise that establishes the noise floor, sensitivity level, and signal-to-noise ratio. These effects cause the received signal to be spatially scaled, linearly mixed/rotated between channels, and spun as a result of unidentified time-varying processes. Last but not least, random filtering is applied to real channels based on the received signal modes that arrive at the receiver with changing amplitude, phase, Doppler, and latency. This phenomenon, often referred to as multi-path fading or frequency selective fading, can happen in any environment where signals may be reflected off of objects like cars, buildings, or other reflectors. This is normally eliminated at the receiver by deconvolution of the received signal and estimate of the instantaneous value of the time-varying channel response.

In a radio communication system, one class of receiver which is commonly considered is a ”matched-filter” receiver. As the correct symbol slides across the correct symbol time in the received signal, expertly constructed filters that are matched with each sent symbol representation are convolved with the incoming time signal and produce peaks. In an effort to improve the signal to noise ratio, we convolve to average out the impulsive noise in the receiver. Typically, before this convolutional stage, symbol timing and carrier frequency is recovered using an expert envelope or moment based estimators derived analytically for a specific modulation and channel model. Therefore, the justification for using a convolutional neural network in this application is that they will learn to form matched filters for a variety of temporal features, each of which will have some filter gain to operate at lower SNR, and which when combined can form a reliable basis for classification.

The invariance to linear mixing, rotation, time shifting, scaling, and convolution through random filters (with well-characterized probabilistic envelopes and coherence periods) can be applied to many of these recovery processes in radio communications systems. These are comparable to similar learning invariance, which is a major focus in the learning of the vision domain, where matched filters for particular items or features in the image may experience scaling, shifting, rotation, occlusion, lighting variation, and other types of noise. In order to build matched filters that can naively distinguish symbol encoding features without expert comprehension or estimation of the underlying waveform, we aim to take advantage of the shift-invariant capabilities of the convolutional neural network.

4.5 Deep Learning For AMC At The Receiver

Now the question is how the Legitimate Receiver (LRx) can classify the received modulated waveforms with state-of-the-art DL tools? We created a state-of-art classifier so as to increase the performance based on the core ideas that came from previous works.

4.5.1 Deep-Learning Signal Classifier

Based on [2] and [7], a 7-layer CCN will be constructed, which consists of six convolution layers and one fully connected layer. Each convolution layer except the last is followed by a batch normalization layer, rectified linear unit (ReLU) activation layer, and max pooling layer. In the last convolution layer, the max pooling layer is replaced with an average pooling layer. The output layer has softmax activation.

The publications [3] and [5] examined each layer and the tasks of the neural network has been utilised in great detail. Convolution layer is typically the dot product of the weights (hyper-parameters in CNNs) of a given kernel and the constituents of an input map within a receptive field identified by a spatial size (height & width), where the depth size of the kernel is usually the same as the amount of channels in the input map. Convolution at a coordinate (i, j) can be calculated mathematically as follows:

$$conv\ (W\ , I)_{(i,j)}\ = \sum_{m,n,k}^{M,N,K} W_{(m,n,k)} I_{(i+m,j+n,k)} + b,$$

where $W \in \mathbb{R}^{M \times N \times K}$ and I denote the kernel and the input, respectively, and b is the scalar bias.

All layers, excluding the output layer, utilize a Rectified Linear Unit (ReLU) activation function [8]. The ReLU function is a popular activation function in the literature, as it has been shown to be robust to saturation (when output is near zero or one), which usually causes learning to slow:

$$f(x) = ReLU(x) = \begin{cases} x, & \text{if } x \geq 0. \\ 0, & \text{if } x < 0. \end{cases}$$

In order to condense the spatial size of feature maps, pooling layers are frequently placed between succeeding convolutional layers. A max pooling layer conducts downsampling by determining the maximum of each region, whereas an average pooling layer determines the

mean value of each region. The input is divided into rectangular pooling sections (determined by pool size). Pooling layers not only reduce the amount of parameters that need to be learned in the subsequent convolutional layers, but they additionally mitigate the problem of over-fitting. By computing the height and width dimensions along the channel dimension of the input, a global average pooling layer performs down-sampling in a number of classification networks. This layer can be used before the final fully connected layer to significantly reduce the network size without negatively impacting performance.

In numerous typical CNNs, connecting all neurons in a fully connected layer to all neurons in the previous layer allows the network to summarize the locally informative features learned in many former layers to classify the input. The number of neurons in the last fully connected layer is identical to the number of classes in a given dataset.

For general multi-class classification tasks, a softmax layer, which usually follows the last fully connected layer, applies a softmax function to calculate a probability for every possible class from the activation outcomes of the previous layer. The classification output (i.e., the output of the entire network) is implied by the class with the highest probability. CNNs often determine the cross entropy loss for multi-class classification as follows during the training process:

$$\mathfrak{L} = -\sum_{i=1}^{N_s}\sum_{j=1}^{n_c} u_{ij} ln(v_{ij})$$

where N_S is the number of training samples, u_{ij} denotes the ground-truth of the i-th sample associated with the j-th class, and v_{ij} denotes the output inferred by the network for the class j of the sample i.

In addition to the layers indicated above, more layers can be used to increase the learning effectiveness of CNNs. These extra layers include the batch normalization (bn) layer, element-wise addition layer, and concatenation layer (which includes both depth-wise and spatial concatenation). By lessening the internal co-variate shift (i.e., the change in output distribution brought on by activation functions during the training process), batch normalization is typically employed in CNNs to speed up the network training convergence and decrease the sensitivity to network initialization. In the proposed network, the batch normalization layer is deployed between the convolutional layer and the activation layer.

4.5.2 Settings of The Classifier

Then, we optimized the hyperparameters of the classifier. The hyper-parameters are chosen according to the basic observation that we make from simple CNN architectures that having larger filters close to the input layer followed by smaller filters close to the output layer leads to significant accuracy improvement. 2,000 frames produced for each modulation for both datasets, where 80% is used for training, 10% for validation and the remaining 10% for testing. There is no general rule of thumb in choosing percentages, because it often depends on the SNR of the data and the size of the training data. However, it is accepted that the more frames I get, the more accurate the accuracy of the classifier will be. SGD with a Mini-batch size of 256 was used as the solver for CNN training. Regarding the learning rate of the classifier, it controls the rate or speed at which the classifier learns. Given a perfectly configured learning rate, the model will learn to best approximate the function given available resources in a given number of training epochs. We should not use a learning rate that is too large or too small. Nevertheless, we must configure the model in such a way that on average a "good enough" set of weights is found to approximate the mapping problem as represented by the training dataset. After several attempts to find the ideal learning rate for the model, we concluded that it is around 0.02. However, surprisingly it seemed that after about 10 epochs the learning rate of our model needed to be reduced because it was quite high for that epoch. To be able to solve this problem, we added the learn rate schedule to the classifier. More specifically, after surpassing the 9th epoch, the learning rate will be reduced by one tenth, and simultaneously overfitting is reduced. It should be stressed that the learning rate interacts with other hyper-parameters, such as the Mini-Batch size and epochs. In this experiment, this interaction may make it hard to isolate the effect of batch size alone on model quality. However, the choice of the mini-batch size influences the training time until convergence, the training time per epoch and the resulting model quality. We made multiple tries to determine the appropriate mini-batch size for the model and ultimately came to the conclusion that it must be set at 256. Using a larger or a smaller batch there is a degradation in the quality of the model.

Chapter 5

Numerical results and discussion

The objective of the simulator is to thoroughly investigate how obfuscated signals affect the accuracy of classifier and the demodulation performance at the nominal receiver. The AWGN, a fading channel (we'll use the Rayleigh channel in this case), and finally an equalization channel with a partial understanding of the spoofing signal will all be used to produce the findings. For the first two channels we will consider that the communication receiver LRx does not employ any equalization of the obfuscating signal. The spectrograms, confusion matrices, BER, and SER, which are created in each case and allow us to precisely understand how they are distorted by the influence of the spoofing techniques, will be used to draw conclusions. However, regardless of the channel to be used, it is worth mentioning some general information regarding how the parameters of the classifier affect its accuracy under the influence of obfuscation techniques. Initially all the system parameters for the two datasets are shown in tables 5.1 and 5.2.

Table 5.1: Description of Variables used in this Study containing Dataset 1

Modulations of dataset 1	**BPSK, QPSK, 8PSK, 16QAM, 64QAM, PAM4, GFSK, B-FM, DSB-AM, SSB-AM**
Samples per symbol	8
Samples per frame	1024
SNR tested	-10 to 20 db
Number of training samples	17600
Number of validation samples	2200
Number of test samples	2200
Sample rate	200 MHz
Transmission Delay	50 ms

Clearly, in order for the classification for various modulations to be coherent, the system must be trained with a large number of samples. An insufficient number of samples in a neural network results in a decrease in the accuracy of the classifier. Also, with a huge number of samples the computational complexity would increase exponentially. So, by taking 2000 samples for each type of modulation, it turned out to be sufficient enough to train the neural network system [10]. Another important parameter of the system is the size of the frame. According to [8], the network observes the signal for a longer period of time with larger input samples, giving it more knowledge about the signal to use in its estimation. In order to get a complete view of what is happening in each channel, it should be remembered that the range of SNR values vary from -10 to 20 db.

Table 5.2: Description of Variables used in this Study for Dataset 2

Modulations of dataset 2	**BPSK-OFDM, QPSK-OFDM, 8PSK-OFDM, 16QAM-OFDM, 64QAM-OFDM**
Samples per symbol	8
Samples per frame	1024
SNR range	-10 to 20 db
Number of training samples	8000
Number of validation samples	1000
Number of test samples	1000
Sample rate	200 MHz
Number of OFDM subcarriers	1024
OFDM cyclic prefix length	16
Transmission Delay	50 ms

5.1 AWGN Channel

We'll concentrate on the AWGN channel in the following section. First, we will examine the outcomes for the first dataset. Figure 5.1 show how the accuracy of the classifier varies for different values of the spoofing signal parameters. For the construction of this graph, δf values were taken which are varied accordingly by multiples of f_m. It was shown that the range of values of $f_m \in [10\ 60]$, satisfactorily distorts the transmitted signal and there was no need to extend to larger values because they showed similar results. The results showed that not all values change in the same way. As the δf values tended towards 0, the less effect the spoofing signal had on the accuracy of the classifier. This makes sense because the spoofing

signal has a smaller and smaller frequency domain, which means that the signal becomes less and less corrupted.

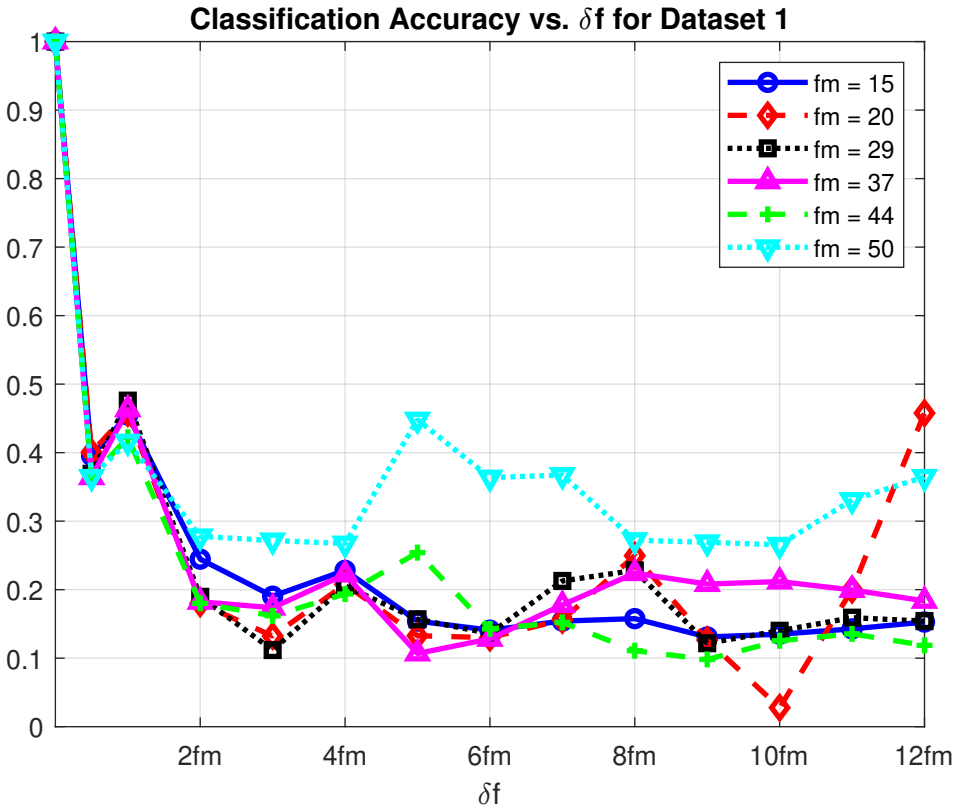

Figure 5.1: Accuracy vs. δf for AWGN channel

A special case where f_m is equal to δf was also recalled. It is expected that for all values of f_m the accuracy of the classifier increases minimally, because the alteration of the transmitted signal is only hit by the parameter f_m. For this reason, the frequency domain remains unchanged and the accuracy of the classifier was estimated at about 45% for all the above mentioned f_m values. Then, for δf values greater than f_m, three sets of results will be presented. These include:

1. f_m values divisible by 10.
2. f_m values divisible by 25.
3. All remaining f_m values.

The values chosen to be simulated represent an aforementioned case. In the first case, it was demonstrated that for δf values greater than $8f_m$, the accuracy of the classifier changes

significantly more abruptly than in the third case. In the second case, there is a milder alteration of the spoofing signal on the accuracy of the classifier. On average, it shows an increase in accuracy of 20% compared to the third case. Then, some spectrograms and confusion matrices will be examined in order to study the results found previously more extensively.

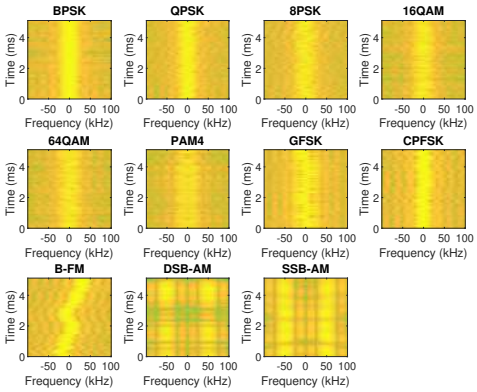

Figure 5.2: Spectrograms at 20db SNR with no spoofing

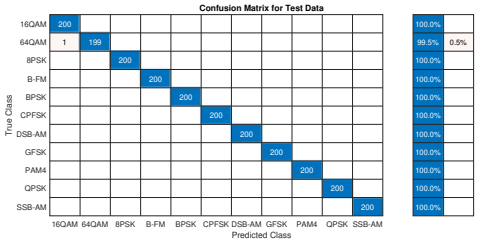

Figure 5.3: Confusion Matrix at 20db SNR with no spoofing

It is clear from Figures 5.2 and 5.3 that the CNN model works effectively in the absence of signal spoofing. It is clear that the network confuses 16-QAM and 64-QAM frames. As each frame only contains 1024 symbols and 16-QAM is a subset of 64-QAM, this issue is expected. Note that this issue appears in both datasets. Regarding the other modulations, the classifier easily separates them, which accounts for its extremely high accuracy rates. An investigation into the impact of signal spoofing is about to be conducted for the three cases mentioned previously.

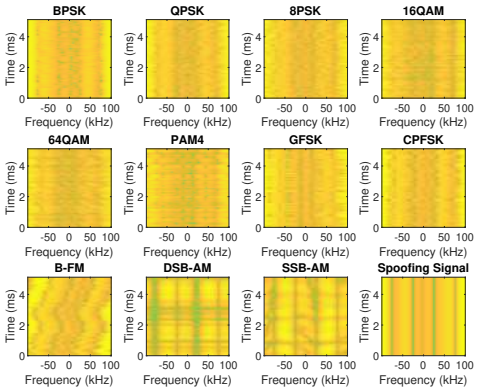

Figure 5.4: Spectrograms at 20db SNR with spoofing parameters $\delta f = 500$ and $f_m = 50$

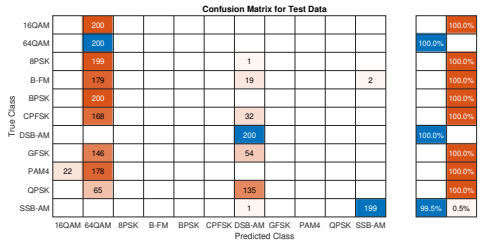

Figure 5.5: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 500$ and $f_m = 50$

Figures 5.4 and 5.5 relate to the first case. For all configurations except SSB-AM and DSB-AM, the aliasing is so significant that the classifier cannot separate any modulation.

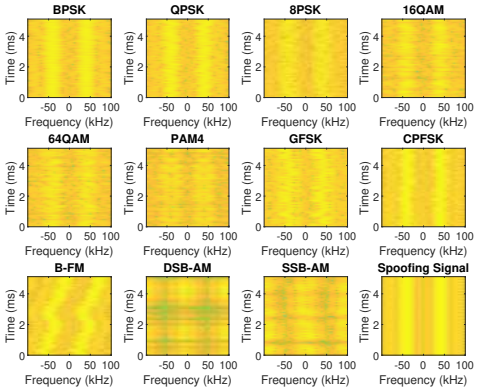

Figure 5.6: Spectrograms at 20db SNR with spoofing parameters $\delta f = 200$ and $f_m = 20$

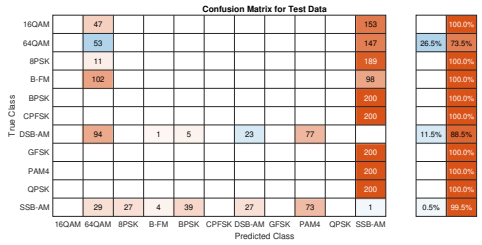

Figure 5.7: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 200$ and $f_m = 20$

Figures 5.6 and 5.7 relate to the second case. The spoofing signal essentially converts all signals in the DSB-AM configuration and at the same time confuses the SSB-AM with other configurations. This explains the tremendously low performance of the classifier which is at 3%.

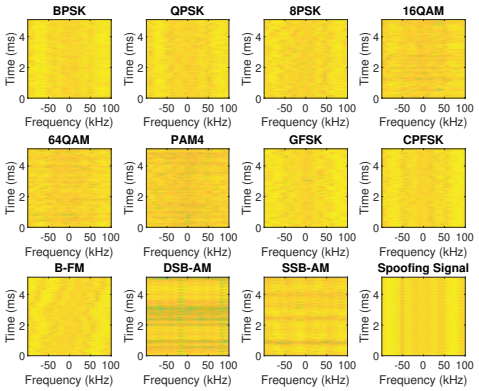

Figure 5.8: Spectrograms at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 15$

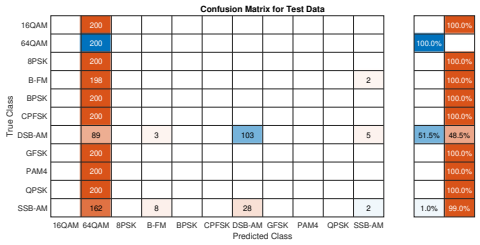

Figure 5.9: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 15$

The obfuscation is seen to alter the modulations to a fairly large extent of every transmitted signal. Since all of them appear to be the same, the classifier cannot operate. The least corrupted modulation is DSB-AM and is therefore the only one that can be detected. It is worth noting that the errors made by the classifier are due to the fact that it has been trained in advance to recognise specific modulations of transmitted signals. Therefore, it is expected that with the existence of spoofing signal, depending on the distortion it achieves on these signals, its performance is drastically reduced. For the other datasets and channels that will be considered shortly, it will not be necessary to go into as much detail on how each modulation is corrupted. The rationale for how the errors occur will remain the same. In this point we will analyze the effect of spoofing signals under various values of channel noise.

Figure 5.10 shows the curve by which the accuracy of the classifier is varied for different values of the channel noise power while keeping the values of the spoofing signal parameters constant. The values of f_m equal to 20 and 50 were used because they showed steeper changes in accuracy of the classifier. It was observed that as the channel noise increases, the effect of signal spoofing decreases. This explains why for higher noise powers a slight increase in the accuracy of the classifier occurs for certain values of δf.

The multi-carrier system was examined afterwards. For almost all values of f_m, the accuracy of the classifier collimated the curve of f_m equal to 15 of the figure 5.11. That is, for values where δf was greater than f_m, the accuracy of the classifier remained constant at 20%. In contrast, for values δf less than f_m, as they tended towards 0, the accuracy gradually increased. However, a special case was discovered, which refers to f_m values divisible 50.

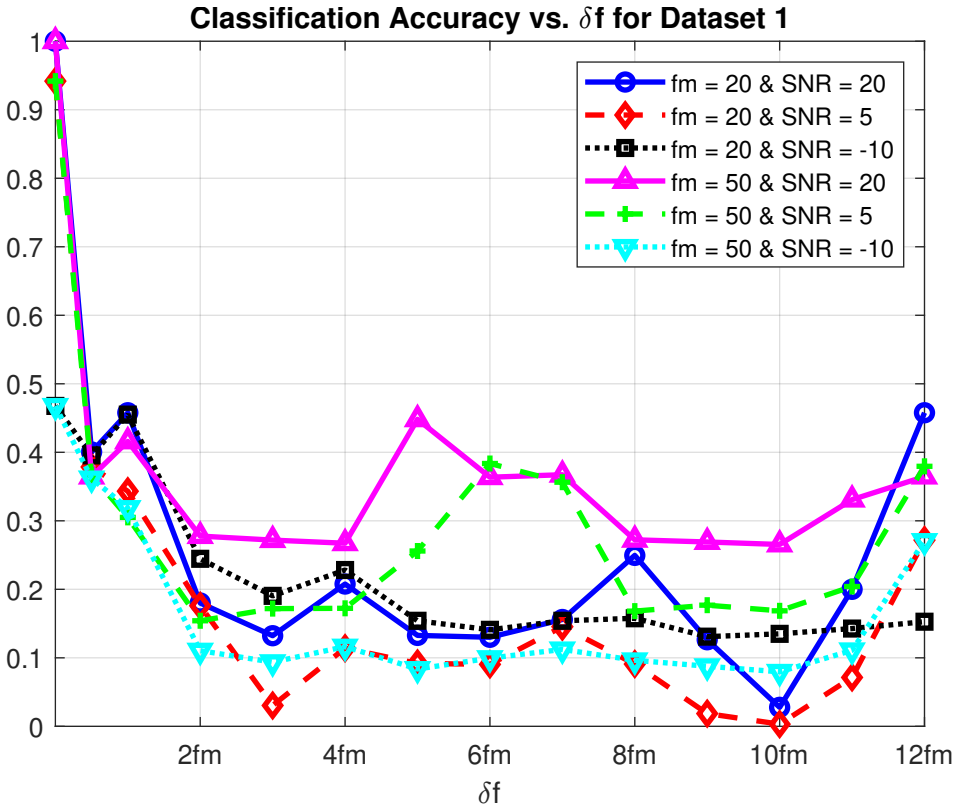

Figure 5.10: Accuracy vs. δf for AWGN Channel for Dataset 1

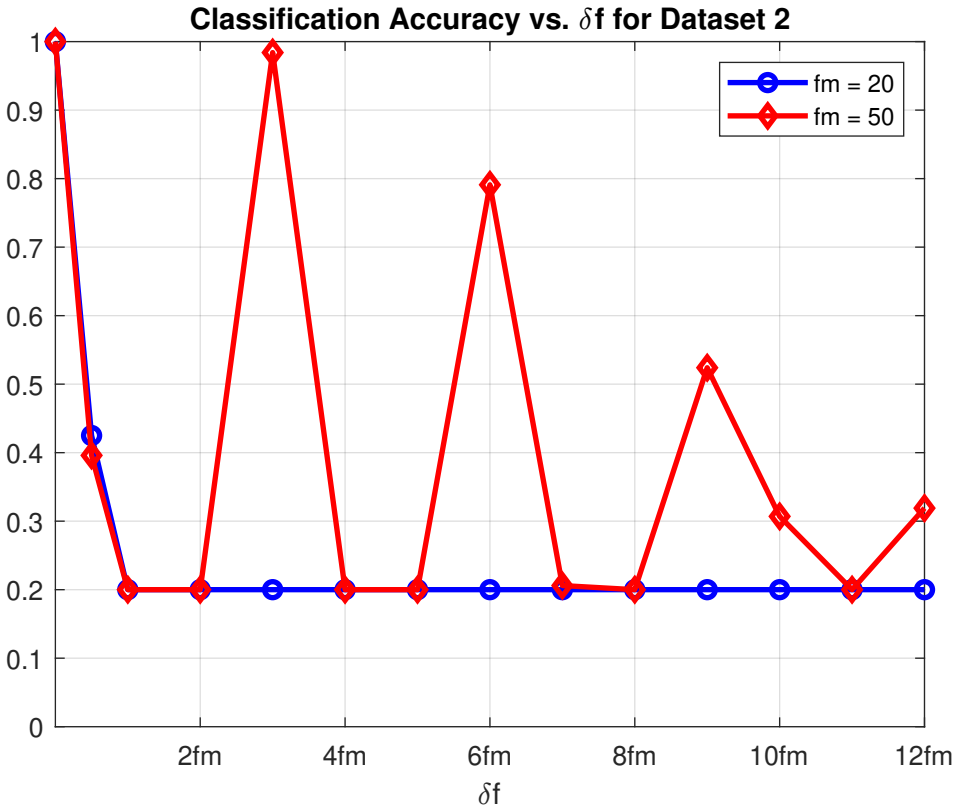

Figure 5.11: Accuracy vs. δf for AWGN Channel for Dataset 2

The following spectrograms indicate that the signals have not undergone any change for values of δf equal to $3f_m$. The accuracy of the classifier appears to behave like a damped oscillation. The influence of channel noise in the presence of signal spoofing was also examined.

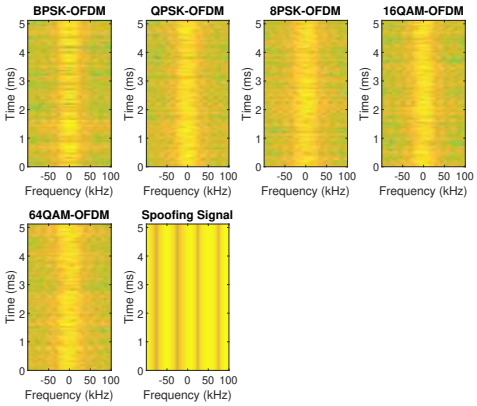

Figure 5.12: Spectrograms at 20db SNR with spoofing parameters $\delta f = 230$ and $f_m = 50$

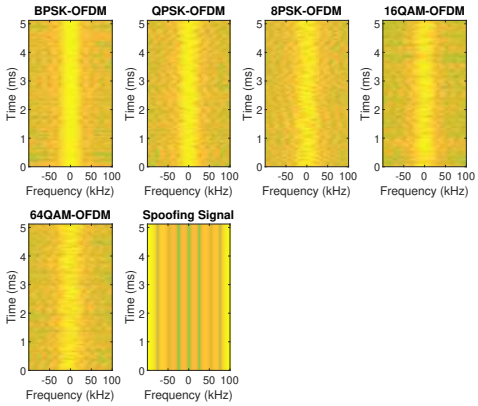

Figure 5.13: Spectrograms at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 50$

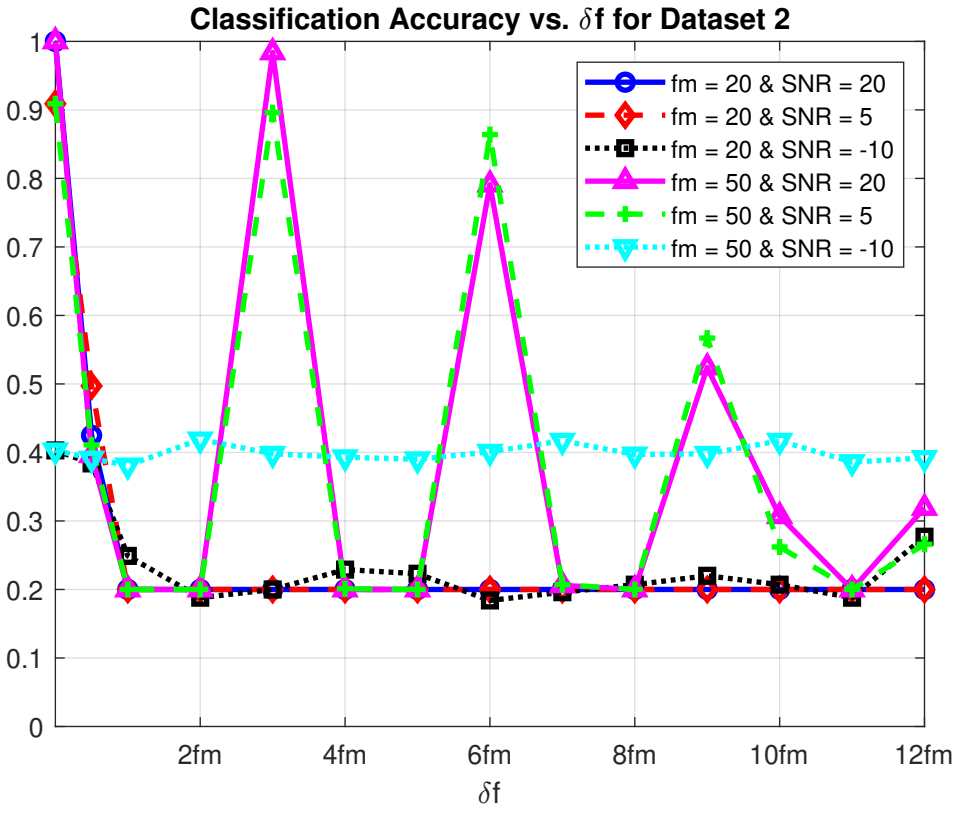

Figure 5.14: Accuracy vs. δf for AWGN Channel for Dataset 2

From Figure 5.14, we draw the same conclusions as in the case of the single carrier system. This means that increasing the power of the channel noise results in a reduction in the effect of obfuscation.

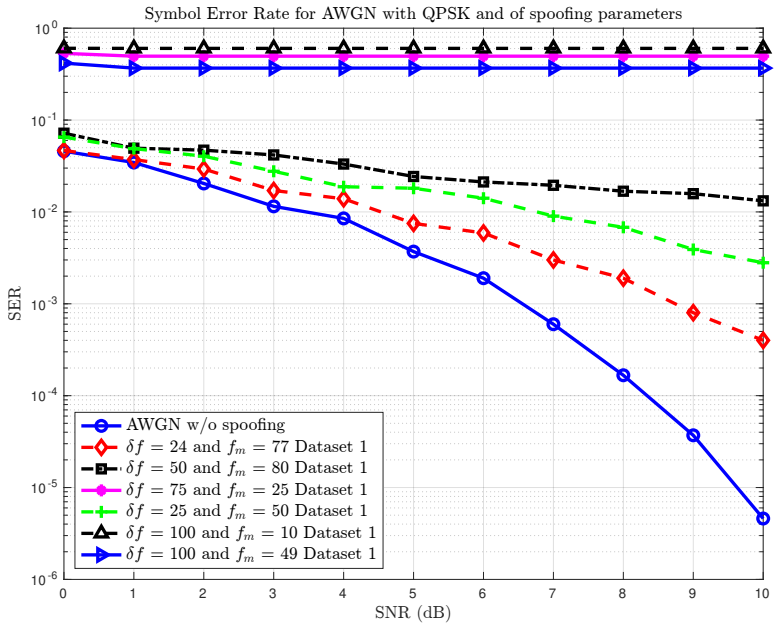

Figure 5.15: SER vs. SNR for AWGN channel w/ and w/o spoofing

The impact of signal spoofing on the SER of the AWGN channel will be discussed in this section. It is important to know that the effects of QPSK modulation for both single-carrier and multi-carrier systems are investigated. It appears that while the spoofing signal significantly affects the accuracy of the classifier, it has little impact on the SER for the AWGN channel. However, for both single-carrier and multi-carrier systems, this statement is particularly true for f_m values greater than δf. For the other parameters of the spoofing signal it seems that they largely alter the SER.

5.2 Fading Channel

In this section we will focus on the influence of the spoofing signal on a fading channel. First, we have to be sure that the properties of the channel have suitable values for the situation we want to model. The parameters we will analyze are Path Delays, Average Path Gains and Maximum Doppler Shifts.

In essence, the presence of reflectors in the environment surrounding a transmitter and receiver create multiple paths that a transmitted signal can traverse. As a result, the receiver

sees the superposition of multiple copies of the transmitted signal, each traversing a different path. So the period of time it takes for the reflected signal to reach the receiver is the path delay. For indoor environments, path delays after the first are typically between 1e-9 seconds and 1e-7 seconds. For outdoor environments, path delays after the first are typically between 1e-7 seconds and 1e-5 seconds. Large delays in this range might correspond, for example, to an area surrounded by mountains. Therefore, we suppose that the system is used in an outdoor environment with the following values:

$$\tau_k = \begin{bmatrix} 0 & 9 & 17 \end{bmatrix} \cdot 10^{-6}$$

The average path gains in the fading channel indicate the average power gain of each fading path. In practice, an average path gain value is a large negative dB value. However, computer models typically use average path gains in the range of [-20, 0] dB.

$$\pi_k = \begin{bmatrix} 0 & -2 & -10 \end{bmatrix}$$

Doppler shifts are specified in terms of the relative speed between a transmitter and a receiver. The maximum Doppler shift in Hertz, $f_d = v \cdot f/c$ In the formula, v is the relative speed in m/s, f is the transmission carrier frequency in Hertz, and c is the speed of light $(3 \cdot 10^8 m/s)$. The relative speed is a magnitude with no directional information. Apply the maximum Doppler shift formula assuming a transmission carrier frequency of 900 MHz, a car moving at freeway speed, and a walking pedestrian. A signal transmitted from a car moving at freeway speed to a stationary receiver would experience a maximum Doppler shift of approximately 80 Hz. A signal transmitted from a mobile held by a walking pedestrian to a stationary receiver would experience a maximum Doppler shift of approximately 4 Hz, where this will be considered for the system.

Table 5.3: Fading parameters and accuracy of the network

Fading parameters			Accuracy	
Path Delays	**Path Gains**	**Maximum Doppler Shift**	**Dataset 1**	**Dataset 2**
$\tau_k = \begin{bmatrix} 0 & 9 & 17 \end{bmatrix} \cdot 10^{-6}$	$\pi_k = \begin{bmatrix} 0 & -2 & -10 \end{bmatrix}$	4	91.68 %	86.1%
$\tau_k = \begin{bmatrix} 0 & 9 & 17 & 28 & 36 & 42 \end{bmatrix} \cdot 10^{-6}$	$\pi_k = \begin{bmatrix} 0 & -2 & -4 & -6 & -8 & -10 \end{bmatrix}$	4	79.77%	84.2%
$\tau_k = \begin{bmatrix} 0 & 9 & 17 \end{bmatrix} \cdot 10^{-6}$	$\pi_k = \begin{bmatrix} 0 & -2 & -10 \end{bmatrix}$	100	90.05 %	63.5%

In table 5.3 some values of the parameters were taken to decide which ones are suitable

for the case where we examine the effect of spoofing signals on the accuracy of the classifier. Since the spoofing signals, as we saw in the previous section, already diminish accuracy to a sufficient degree, we do not want the fading channel variables to further reduce it significantly. Therefore, we need fading channel characteristics that marginally alter the accuracy of the classifier. For example, we can see that for the first dataset, the doppler shift may have zero effect, but for the second one it drops the accuracy dramatically. Also, the number of signal reflections affects the respective datasets in an inverse way. We get to the conclusion that the parameters of the system will be chosen from the first row of the table.

The accuracy of the classifier changes for various values of δf, as seen in figure 5.16. These values were assumed to increase depending on the value of f_m. In general, several values of the f_m parameter were examined and it was found that most of them follow the curve formed for f_m equal to 15 Hz. However, two special cases were found where a curve was formed that was distinct from the others. In the case of f_m equal to 25 Hz (and its multiples) it was found that the obfuscated signal has less effect on the channel signal. For the case of f_m equal to 40 (and its multiples), it did not need to be assumed further since it has similar behaviour to f_m equal to 25. As a result, in the section that follows, we will take into account for f_m equal to 25 Hz the influence of the obfuscated signal for various parameters of the fading channel, including its noise, as in the other circumstances, the spoofing signal already has a considerable impact on the accuracy.

The analysis will start with the single carrier system. Figure 5.17 shows a particular behaviour of the curve with respect to the channel noise. While it was expected that with increasing noise the accuracy of the classifier would decrease linearly, surprisingly there is a slight increase in the accuracy of the classifier in the range of values from $5f_m$ to $11f_m$. To explain what is actually occurring, some spectrograms and confusion matrices will be presented in figures 5.18 to 5.21.

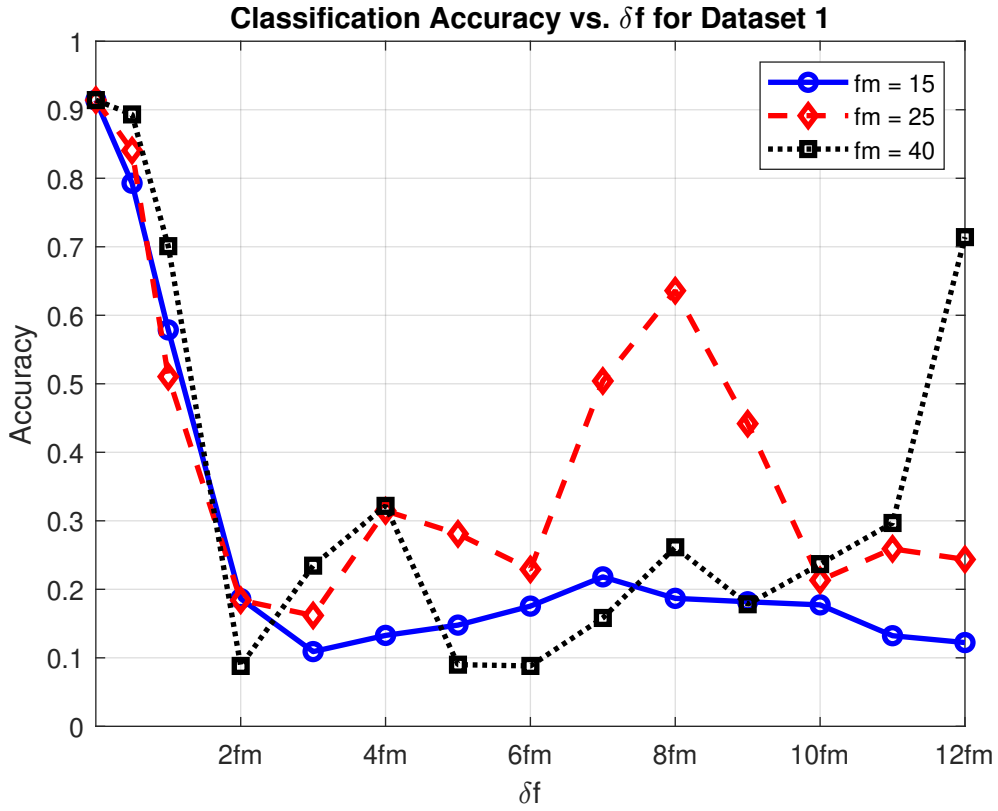

Figure 5.16: Accuracy vs. δf for Fading channel

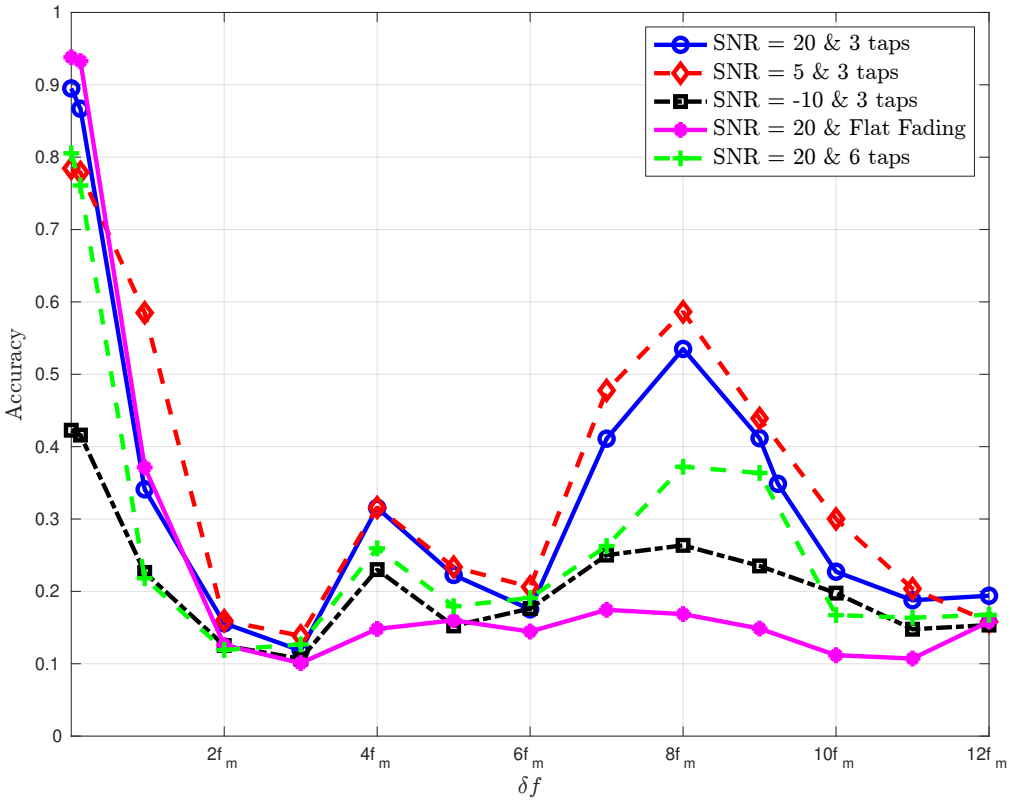

Figure 5.17: Accuracy vs. δf for various values of Fading Channel Parameters

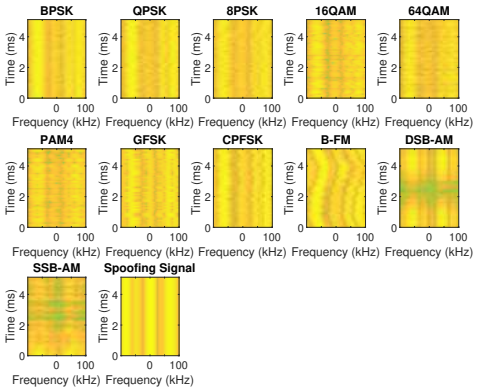

Figure 5.18: Spectrograms at 20db SNR with spoofing parameters $\delta f = 200$ and $f_m = 25$ for Fading Channel

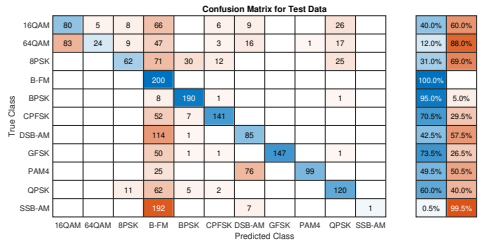

Figure 5.19: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 200$ and $f_m = 25$ for Fading Channel

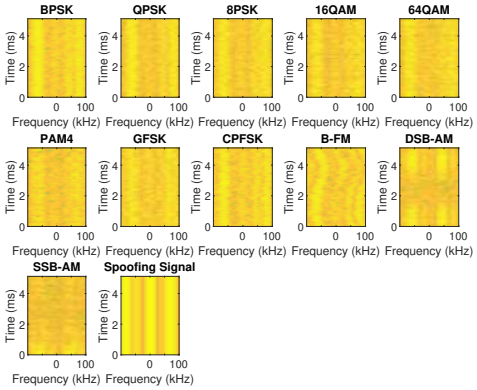

Figure 5.20: Spectrograms at 5db SNR with spoofing parameters $\delta f = 200$ and $f_m = 25$ for Fading Channel

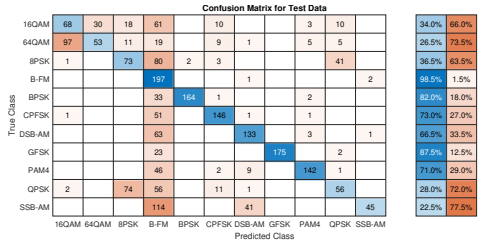

Figure 5.21: Confusion Matrix at 5db SNR with spoofing parameters $\delta f = 200$ and $f_m = 25$ for Fading Channel

It is observed that with increasing noise power the effect of signal spoofing decreases more and more. The result for the 5db was therefore judged by noise, rather than by the spoofing signal. Also for -10db, there is expected to be a decrease in accuracy. The number of paths was another semantic component of the fading channel taken into consideration. It is common knowledge that a channel is flat-fading if there is just one fading path. So, we looked at two further scenarios, one of which has fewer fading paths and the other of which has more. For the flat-fading channel, it appears that the curve has a similar behavior as in the AWGN case. This is anticipated as the spoofing signal is multiplied by the whole fading channel and does not become entangled with the convolution procedure as in the other circumstances. We

then increased the fading paths from 3 to 6 and noticed that the accuracy of the classifier decreased significantly as expected.

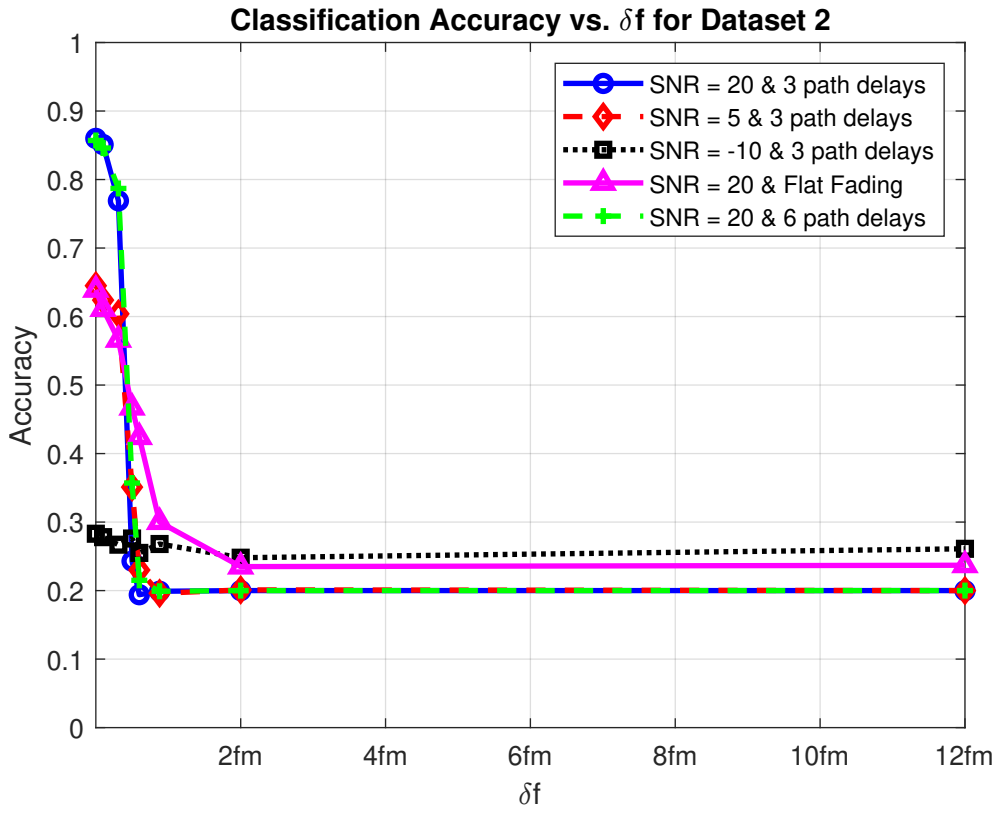

Figure 5.22: Accuracy vs. δf for Fading channel

Now we will consider the effect of spoofing signals on multi-carrier systems. It is evident that the obfuscation significantly lowers the accuracy of the classifier in each situation. Since the classifier is aware that this dataset only comprises OFDM signals, it cannot mistake them for other signals like SSB-AM or DSB-AM. Additionally, the curves were looked at for each value of f_m and fading channel parameters, and behaved similarly to the AWGN channel. The effect of obfuscation on BER was also examined. Various values of the spoofing signal parameters were obtained for a Flat Fading Channel. It has been demonstrated to have a similar impact on classification. That is to say, the Bit Error Rate had a high error rate for values when low classification accuracy was evident. Note that this was demonstrated for any fading channel for which MLSE equalizer was used to calculate the SER.

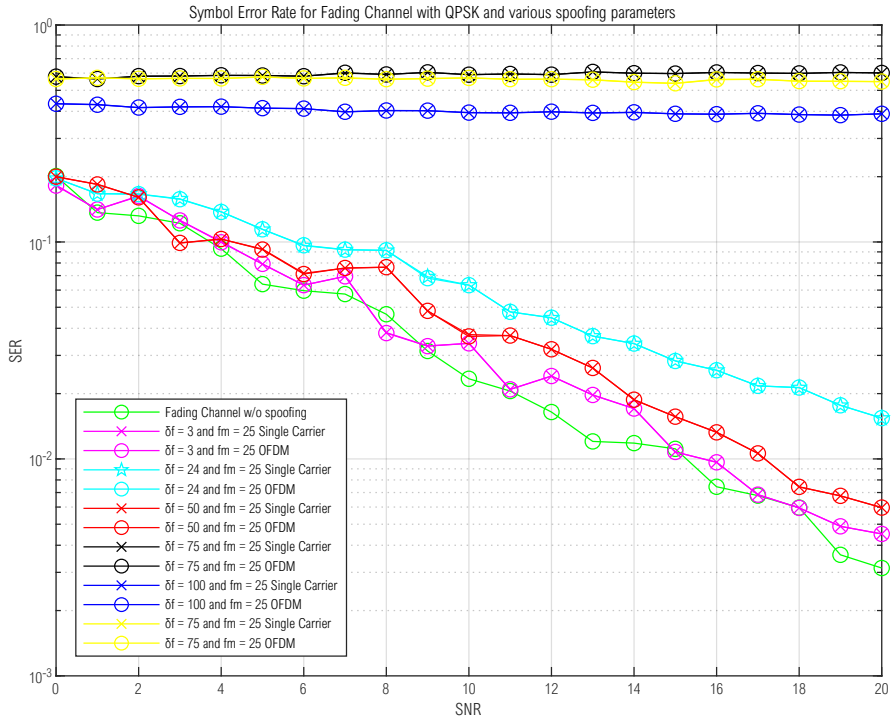

Figure 5.23: SER vs. SNR for flat Fading channel w/ and w/o spoofing

5.3 Fading Channel with partial Spoofing Knowledge

Our key objective is to be able to employ obfuscation techniques to conceal the signals we generate from eavesdroppers as much as we can. Therefore, we assume that the legitimate receiver LRx employ equalization of the obfuscating signal. It is well known that all receivers, including hostile ones, may apply equalization to the channel we are transmitting on. Therefore, a crucial scenario that needs to be taken into account is when the receiver equalizes the channel while being oblivious of the spoofing signal. The single carrier system will be the starting point for our investigation. Note that MLSE algorithm will be used in the equalization method.

As in the case of the fading channel, several values of f_m were considered and again most seemed to form the curve of f_m equal to 15 Hz. The obfuscated signal was observed to affect drastically even the equalization channel. The accuracy of the classifier only slightly increases (between 2% and 3%) for certain spoofing signal parameter levels, according to what has been identified. For these f_m values, although the receiver is aware of the spoofing signal, there is no effect of channel equalization and it doesn't benefit the LRx receiver after all.

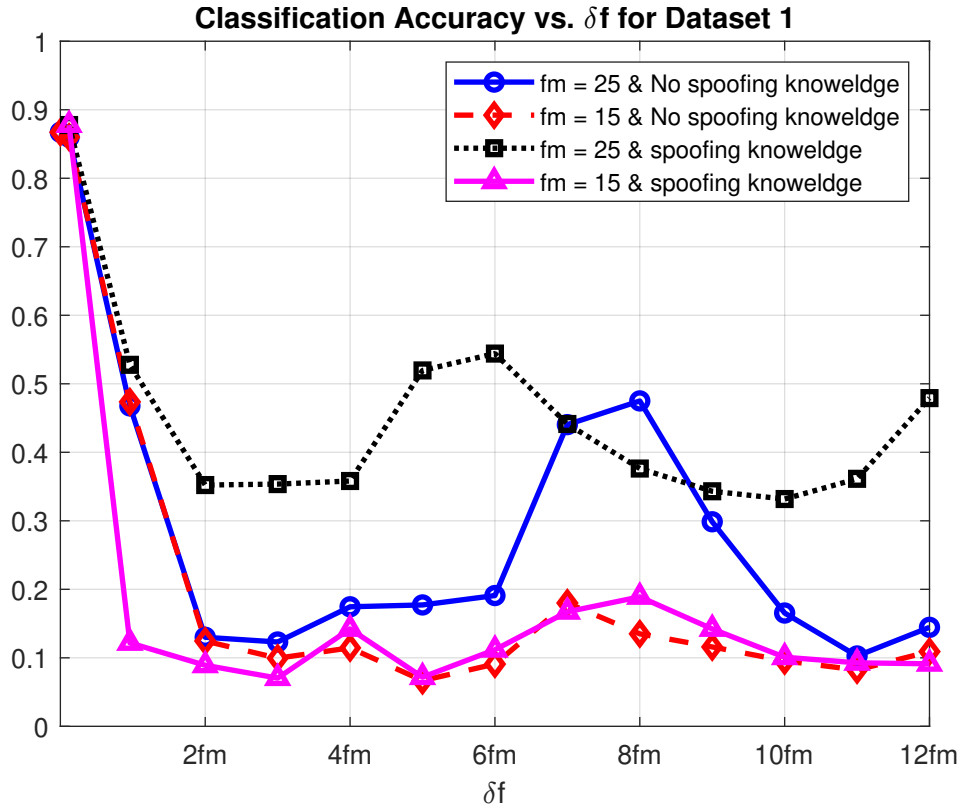

Figure 5.24: Accuracy vs. δf for equalization channel

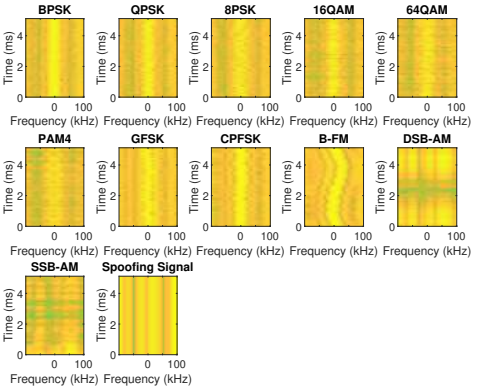

Figure 5.25: Spectrograms at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 25$ for Fading Channel with partial Spoofing Knowledge

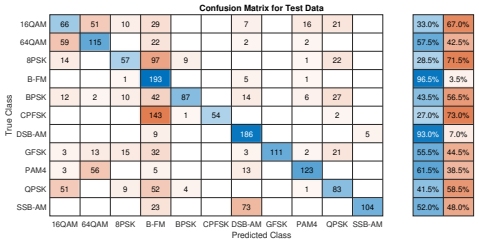

Figure 5.26: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 25$ for Fading Channel with partial Spoofing Knowledge

However, with respect to the value of the f_m parameter located at 25 Hz, the above graph suggests that for values of δf between $5 f_m$ and $6 f_m$, it is quite beneficial for the LRx receiver to be aware of the spoofing signal. This is explained by the fact that the spoofing signal

significantly conceals signal modulations and by equalizing the channel, the classifier can achieve an accuracy of just under 60%.

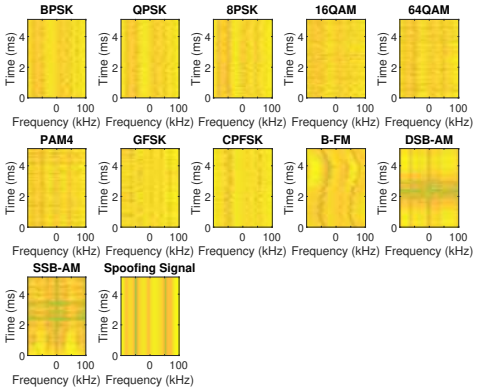

Figure 5.27: Spectrograms at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 25$ for Fading Channel without Spoofing Knowledge

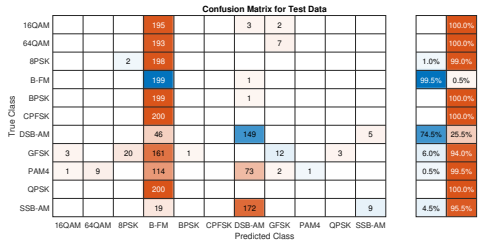

Figure 5.28: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 150$ and $f_m = 25$ for Fading Channel without Spoofing Knowledge

Even from the spectrograms it appears that the effect of signal spoofing has been corrected to a satisfactory degree.

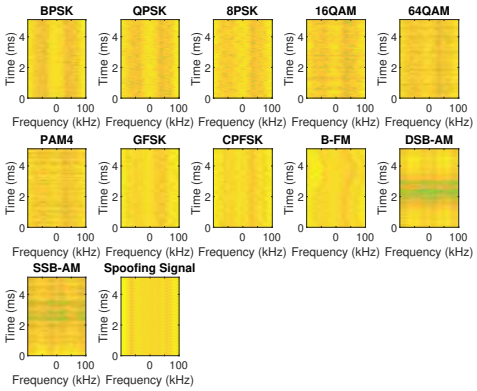

Figure 5.29: Spectrograms at 20db SNR with spoofing parameters $\delta f = 120$ and $f_m = 15$ for Fading Channel without Spoofing Knowledge

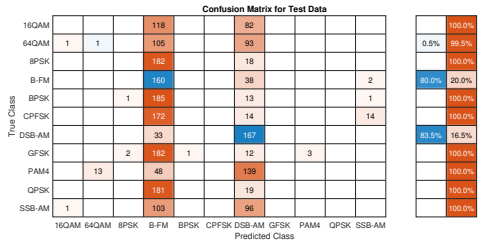

Figure 5.30: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 120$ and $f_m = 15$ for Fading Channel without Spoofing Knowledge

The confusion matrices imply that most of the errors are made by the classifier by confusing almost all modulated signals with B-FM signals. For the remaining fm values, from figures 5.27 to 5.30, although LRx is aware of the spoofing signal, it is not able to eventually

identify the modulations of the received signals. So these values of spoofing parameters do not serve either the receiver because the signals are heavily distorted by the obfuscation.

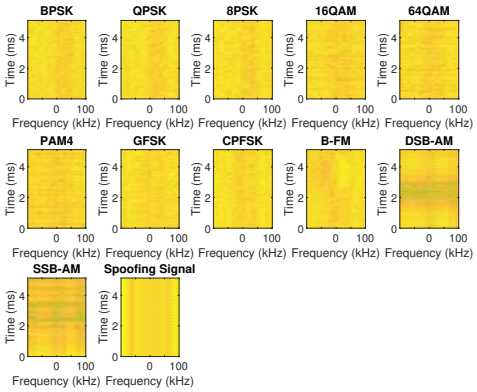

Figure 5.31: Spectrograms at 20db SNR with spoofing parameters $\delta f = 120$ and $f_m = 15$ for Fading Channel with partial Spoofing Knowledge

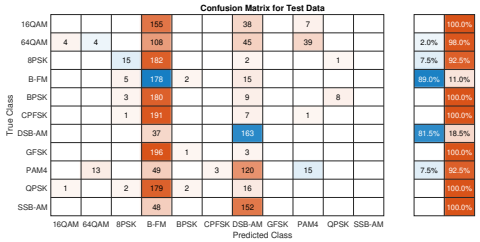

Figure 5.32: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 120$ and $f_m = 15$ for Fading Channel with partial Spoofing Knowledge

We can infer that the transmitter must perform obfuscation by spoofing signals that do not significantly alter the signal it wants to transmit, so that LRx to be able to equalize the channel and receive the actual information sent by the transmitter. At the same time, however, the information must be corrupted as much as necessary to protect it from eavesdroppers.

Subsequently, now we will focus on the impact of spoofing signals in multi carrier systems. It should be noted that the analysis must be performed on a different set of values than those acquired from the dataset related to single carrier systems. The spoofing signal severely affects the modulated signal for δf values larger than $1.5 f_m$, hence the accuracy of the classifier remained constant at 20% in both scenarios of the awareness of the spoofing signal. For this reason, a few submultiples of f_m were taken for the δf values. Several values of f_m were examined and it was found that in the range 25 to 30 and their multiples, LRx has the ability using channel equalization, since it knows the obfuscated signal, to recognize a very large percentage of the signal sent by the transmitter and at the same time the eavesdropper cannot separate any modulation because it does not have knowledge of the spoofing signal. In contrast, for all other values of f_m, the correct receiver, although aware of the spoofing signal, this technique further reduces the accuracy of the classifier.

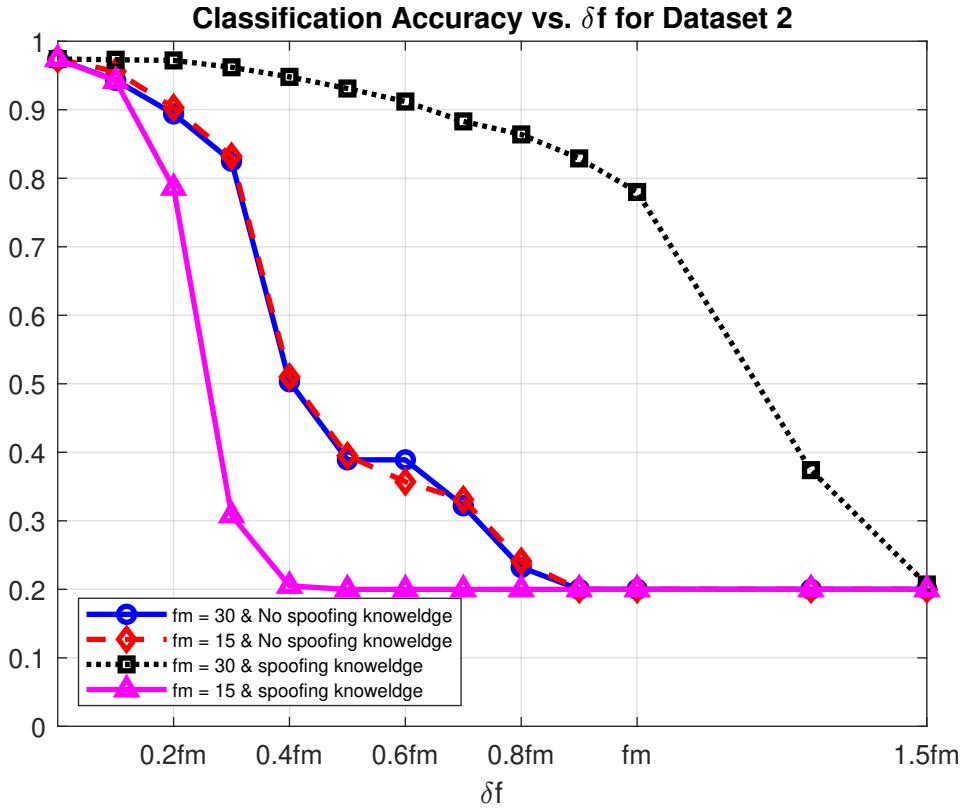

Figure 5.33: Accuracy vs. δf for equalization channel

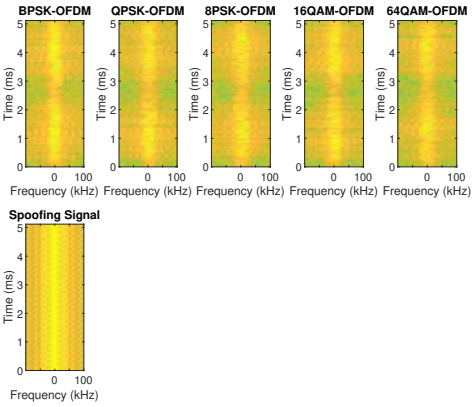

Figure 5.34: Spectrograms at 20db SNR with spoofing parameters $\delta f = 24$ and $f_m = 30$ for Fading Channel without Spoofing Knowledge

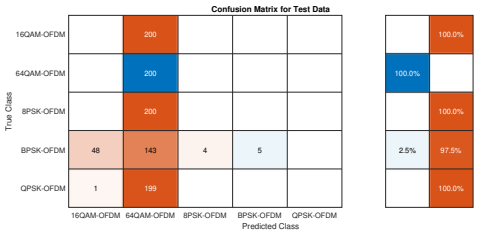

Figure 5.35: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 24$ and $f_m = 30$ for Fading Channel without Spoofing Knowledge

So for the transmission of OFDM signals it is suggested to avoid these f_m values. Some spectrograms and confusion matrices will then be presented (figures 5.34 - 5.41). There are no obvious conclusions from these. In any case, the spoofing signal visually has very little

effect on the alteration of the spectrograms, but the classifier, because it is trained and more experienced than a human, notices even the smallest differences formed due to the effect of spoofing signals. The selection of the parameters f_m and δf is considered particularly important for the multi-carrier system and must be done cautiously.

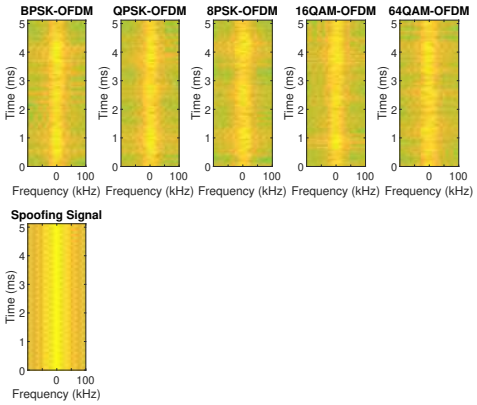

Figure 5.36: Spectrograms at 20db SNR with spoofing parameters $\delta f = 24$ and $f_m = 30$ for Fading Channel with partial Spoofing Knowledge

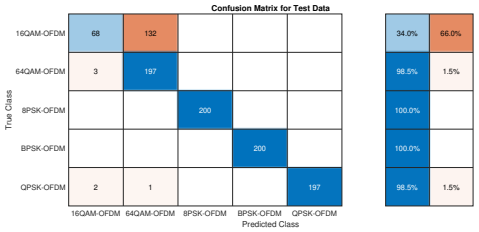

Figure 5.37: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 24$ and $f_m = 30$ for Fading Channel with partial Spoofing Knowledge

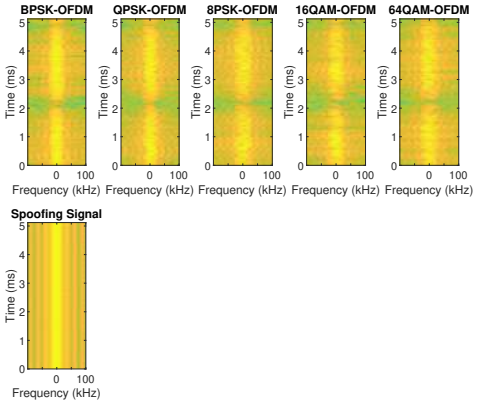

Figure 5.38: Spectrograms at 20db SNR with spoofing parameters $\delta f = 6$ and $f_m = 15$ for Fading Channel without Spoofing Knowledge

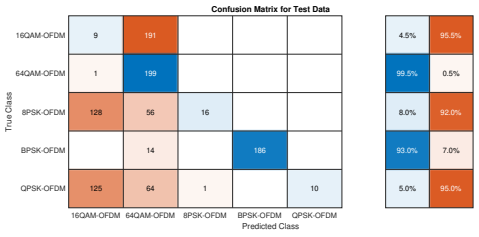

Figure 5.39: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 6$ and $f_m = 15$ for Fading Channel without Spoofing Knowledge

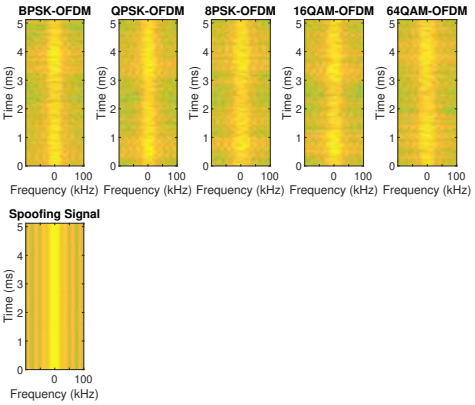

Figure 5.40: Spectrograms at 20db SNR with spoofing parameters $\delta f = 6$ and $f_m = 15$ for Fading Channel with partial Spoofing Knowledge

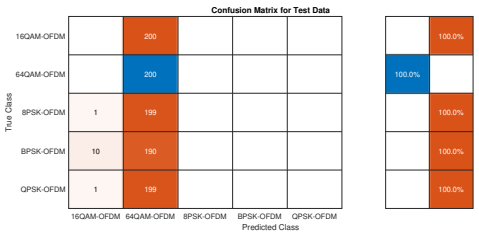

Figure 5.41: Confusion Matrix at 20db SNR with spoofing parameters $\delta f = 6$ and $f_m = 15$ for Fading Channel with partial Spoofing Knowledge

We therefore conclude that the system for different values of the spoofing signal parameters becomes precarious. Figure 5.42 depicts the SER for the equalization channel with various spoofing signal parameters values. As expected, the understanding of LRx receiver for the spoofing signal corrects the Symbol Error Rate, allowing it to receive the correct information.

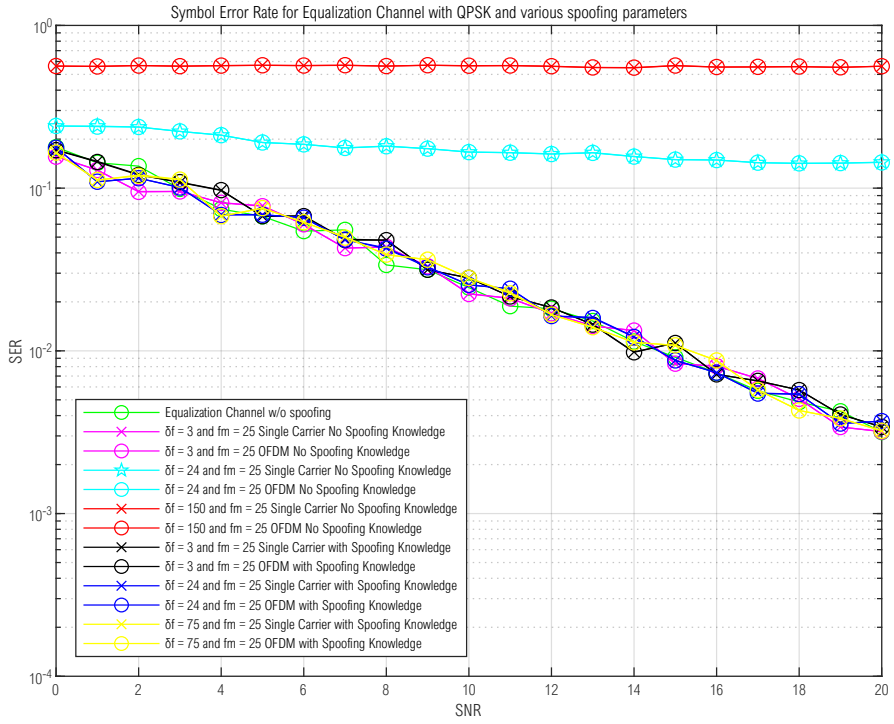

Figure 5.42: SER vs. SNR for Equalization channel w/ and w/o spoofing

Chapter 6

Conclusions

6.1 Summary

In this thesis we explored the use of a new wireless signal obfuscation waveform for preventing modulation classification with Deep Learning (DL) techniques. Their impact on different channel types was examined and several interesting results were presented. The proposed waveform is coupled with its careful planting in the transmitted signal so as to prevent the extraction of clean data for training the DL-based classifier. A way for the transmitter to be able to protect the data it sends from malicious network receivers was also proposed for both single-carrier and multi-carrier systems. Assuming that the spoofing signal is known to the LRx receiver, through the channel equalization technique, it is able to recover a percentage of the data. However, this way is quite restrictive and only works for a certain range of spoofing signal parameters.

6.2 Future work

The analysis would not be complete without emphasizing the limitations of the DL-based models used for this work. First, this neural network cannot recognize single-carrier and multi-carrier signals simultaneously. According to [1], it essentially joins the two datasets used in this work and also, the receiver does not need to demodulate the received OFDM signals for the classification to work properly. In short, according to the size of the OFDM symbol, it has the ability to perform classification. In this way, we believe that not only space but also time is saved for training the neural network. Second, channel equalization technique

contributes significantly to the performance of the classification in the correct receiver. In general, we believe that there is potential to build a more sophisticated equalizer to cancel as much as possible the effect of obfuscation while hiding from any malicious user the modulation of the transmitted signal. It was shown to work for a limited range of values of the spoofing signal parameters, without completely cancelling its effect. The classifier showed at most a 50% increase in classifier accuracy.

Bibliography

[1] Myung Chul Park and Dong Seog Han. Deep learning-based automatic modulation classification with blind ofdm parameter estimation. *IEEE Access*, 9:108305–108317, 2021.

[2] Timothy J O'Shea, Johnathan Corgan, and T. Charles Clancy. Convolutional radio modulation recognition networks, 2016.

[3] Thien Huynh-The, Quoc-Viet Pham, Toan-Van Nguyen, Thanh Thi Nguyen, Rukhsana Ruby, Ming Zeng, and Dong-Seong Kim. Automatic modulation classification: A deep architecture survey. *IEEE Access*, 9:142950–142971, 2021.

[4] Yu Wang, Jie Gui, Yue Yin, Juan Wang, Jinlong Sun, Guan Gui, Haris Gacanin, Hikmet Sari, and Fumiyuki Adachi. Automatic modulation classification for mimo systems via deep learning and zero-forcing equalization. *IEEE Transactions on Vehicular Technology*, 69(5):5688–5692, 2020.

[5] Siyang Zhou, Zhendong Yin, Zhilu Wu, Yunfei Chen, Nan Zhao, and Zhutian Yang. A robust modulation classification method using convolutional neural networks. *EURASIP Journal on Advances in Signal Processing*, 2019(1):21, Mar 2019.

[6] Ameen Abdelmutalab, Khaled Assaleh, and Mohamed El-Tarhuni. Automatic modulation classification based on high order cumulants and hierarchical polynomial classifiers. *Physical Communication*, 21:10–18, 2016.

[7] Xiaoyu Liu, Diyu Yang, and Aly El Gamal. Deep neural network architectures for modulation classification, 2018.

[8] Lauren J. Wong, William C. Headley, and Alan J. Michaels. Emitter identification using cnn iq imbalance estimators, 2018.

[9] Mohamed K. M. Fadul, Donald R. Reising, and Mina Sartipi. Identification of ofdm-based radios under rayleigh fading using rf-dna and deep learning. *IEEE Access*, 9:17100–17113, 2021.

[10] Sreeraj Rajendran, Wannes Meert, Domenico Giustiniano, Vincent Lenders, and Sofie Pollin. Deep learning models for wireless signal classification with distributed low-cost spectrum sensors. *IEEE Transactions on Cognitive Communications and Networking*, 4(3):433–445, 2018.

[11] Nicolas Papernot, Patrick McDaniel, Somesh Jha, Matt Fredrikson, Z. Berkay Celik, and Ananthram Swami. The limitations of deep learning in adversarial settings. In *2016 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 372–387, 2016.

[12] Hanif Rahbari and Marwan Krunz. Friendly cryptojam: A mechanism for securing physical-layer attributes. In *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks*, WiSec '14, page 129–140, New York, NY, USA, 2014. Association for Computing Machinery.

[13] Christina Pöpper, Nils Ole Tippenhauer, Boris Danev, and Srdjan Capkun. Investigation of signal and message manipulations on the wireless channel. In Vijay Atluri and Claudia Diaz, editors, *Computer Security – ESORICS 2011*, pages 40–59, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.

[14] Mohammad Iftekhar Husain, Suyash Mahant, and Ramalingam Sridhar. Cd-phy: Physical layer security in wireless networks through constellation diversity. In *MILCOM 2012 - 2012 IEEE Military Communications Conference*, pages 1–9, 2012.

[15] Tao Xiong, Wei Lou, Jin Zhang, and Hailun Tan. Mio: Enhancing wireless communications security through physical layer multiple inter-symbol obfuscation. *IEEE Transactions on Information Forensics and Security*, 10(8):1678–1691, 2015.

[16] Shyamnath Gollakota and Dina Katabi. Physical layer wireless security made fast and channel independent. In *2011 Proceedings IEEE INFOCOM*, pages 1125–1133, 2011.

[17] Haitham Hassanieh, Jue Wang, Dina Katabi, and Tadayoshi Kohno. Securing RFIDs by randomizing the modulation and channel. In *12th USENIX Symposium on Networked*

Systems Design and Implementation (NSDI 15), pages 235–249, Oakland, CA, May 2015. USENIX Association.

[18] Muhammad Usama, Muhammad Asim, Junaid Qadir, Ala Al-Fuqaha, and Muhammad Ali Imran. Adversarial machine learning attack on modulation classification. In *2019 UK/ China Emerging Technologies (UCET)*, pages 1–4, 2019.

[19] Chao Wang, Xianglin Wei, Jianhua Fan, Yongyang Hu, and Long Yu. Universal attack against automatic modulation classification dnns under frequency and data constraints. *IEEE Internet of Things Journal*, pages 1–1, 2023.

[20] Antonios Argyriou. Obfuscation of human micro-doppler signatures in passive wireless radar. *IEEE Access*, 11:40121–40127, 2023.