Environment Mapping Using Wireless Channel State Information and Deep Learning

By

Adrian Donarski

A thesis submitted to Macquarie University
for the degree of Master of Research
Faculty of Science and Engineering
25 March 2020

MACQUARIE
University
SYDNEY·AUSTRALIA

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

<div align="center">

_____

Adrian Donarski

</div>

Abstract

Situational awareness is becoming ubiquitous with modern communication technologies. Traditionally, user requirements for location based services have evolved from being driven only by government legislation for improved emergency response, to that of social and market based pressures including social-networking, geo-marketing and augmented reality. Not only do these applications necessitate improved localisation, they must also include awareness of the surrounding environment. Furthermore, recent years have seen a marked explosion in Artificial Intelligence (AI), machine learning and deep learning. Such tools enable inference to be drawn from complex data when traditional mathematical techniques cannot.

This thesis explores the use of AI and wireless Channel State Information (CSI) for obtaining situational awareness through environment mapping. Several deep learning algorithms are developed that permit estimation of interior room dimensions using wireless Channel Impulse Responses (CIRs). Moreover the physical effects that bandwidth and received signal power incur on a CIR, and hence on room dimension prediction, is investigated yielding results with greater than 90% accuracy. This demonstrates that existing consumer wireless systems with similar physical constraints, should be capable of accurately estimating room dimensions given multiple CIR measurements from multiple receiver locations.

Contents

List of Figures

List of Tables

Acronyms

AI Artificial Intelligence. iii, 1, 2, 5, 11, 17, 26, 48

AoA Angle of Arrival. 2, 19

AP Access Point. 23, 24, 27

API Application Programming Interface. 30

BER Bit Error Rate. 8, 48

BP Belief Propagation. 24

BSSID Basic Service Set Identifier. 24

CFR Channel Frequency Response. 7, 20, 21, 26, 48, 50

CIR Channel Impulse Response. iii, v, vii, ix, x, 2–8, 10, 11, 16, 18–20, 22, 23, 26, 29–50, 57, 59–63

CNN Convolutional Neural Network. 28

CPU Central Processing Unit. 31

CR Cognitive Radio. 2

CSI Channel State Information. iii, 2–5, 7, 17, 21–29, 48, 50

CT Computerised Tomography. 12

cuDNN CUDA Deep Neural Network Library. 40

CWT Continuous Wavelet Transform. 21, 22

D2D Device-to-Device. 2

DBN Dynamic Bayesian Network. 24

LBP Loopy Belief Propagation. 24, 25

LoS Line-of-Sight. 1, 2, 6, 11, 18, 19, 25, 33

LSTM Long Short-Term Memory. 15, 16, 40, 60

MDWT Multilevel Discrete Wavelet Transform. 22

MIMO Multiple-In Multiple-Out. 2, 21, 26, 50

ML Machine Learning. 30–32, 39

mmWave Millimeter Wave. 2

MPM Matrix Pencil Method. 3, 18, 19, 42, 49

MRF Markov Random Field. 25

MSE Mean Squared Error. 20, 33, 59, 60

MUSIC MUltiple SIgnal Classification. 3, 18–20, 22, 42, 49

OFDM Orthogonal Frequency Division Multiplexing. 22

PAF Part Affinity Field. 28

PCA Principal Component Analysis. 26–28

PDF Probability Density Function. 24

RBF Radial Basis Function. 27

RBM Restricted Bolzmann Machine. 27

RC Raised Cosine. 43, 60

ReLU Rectified Linear Unit. 33, 59, 60

RF Radio Frequency. 28

RNN Recurrent Neural Network. ix, x, 3–5, 15, 16, 39–46, 49, 62, 63

RRC Root-Raised Cosine. 30

RSS Received Signal Strength. 1, 2, 24

RSSI Received Signal Strength Indicator. v, 17, 18, 22, 24, 25, 28

RT Ray-Tracing. 3–5, 7, 9, 10, 22, 29–31, 45

Rx Receiver. 29, 39, 61

SBR Shooting and Bouncing Rays. 11, 22, 31

SLAM Simultaneous Localisation and Mapping. 1, 4, 23–25, 48

SNR Signal to Noise Ratio. 5, 18, 20, 48, 49

SVC Support Vector Classification. 27, 28

SVD Singular Value Decomposition. 19, 20

SVM Support Vector Machine. 26–28

SVR Support Vector Regression. 27, 28

TDoA Time-Difference-of-Arrival. v, vii, 2, 18, 20

ToA Time-of-Arrival. v, vii, 2, 18–20, 23

ToF Time-of-Flight. 18, 19

Tx Transmitter. 29, 61

UTD Uniform Theory of Diffraction. 10

UWB Ultra-Wideband. 3, 24, 49, 50

VCC Volume Cross-Correlation. 20

VSWR Voltage Standing Wave Ratio. 57

WLAN Wireless Local Area Network. 1, 17

XML eXtensible Markup Language. 31

Chapter **1**

Introduction

Recent advances in communication technologies have been driven by an overwhelming change to user requirements in a modern, social and well-connected society. Today, recent and modern social pressures including the need for assisted navigation, location based services, geo-marketing, social networking, augmented reality, as well as efficient and high throughput systems, are all influencing the development of new communications technologies [1]. Not only do these systems require advanced geolocation techniques designed to provide location estimates in the metre or centimetre ranges far exceeding traditional telecommunication methods, but are also required to provide a degree of situational awareness. Situational awareness is to understand the local environment through not only physical characteristics such as room dimensions and objects present, but also non-physical characteristics such as the network of users both in a wired and wireless medium.

The research presented in this thesis will build upon existing knowledge in environment mapping for localised situational awareness. In particular, the exploitation of wireless communication channels with Artificial Intelligence (AI) techniques are seen to enable accurate prediction of indoor room dimensions, namely a room's width, breadth and height.

1.1 Motivation

With each technological advancement, situational awareness techniques have evolved concurrently to exploit new information produced. Perhaps the most recent and ubiquitous consumer technology found in almost every home is Wi-Fi, short for Wireless Fidelity. Wi-Fi facilitates the construction of Wireless Local Area Network (WLAN) systems, enabling high speed communications between the internet and multiple devices within a limited range, typically 100m or less using a single transmitter.

Classical situational awareness techniques have sought to exploit Wi-Fi for localisation and Simultaneous Localisation and Mapping (SLAM) using Received Signal Strength (RSS) measurements. These measurements exploit the Line-of-Sight (LoS) component of a wireless transmission and when

used in conjunction with traditional localisation techniques such as Time-of-Arrival (ToA), Time-Difference-of-Arrival (TDoA) and Angle of Arrival (AoA), position accuracies within the range of 1m−5m may be obtained [1, 2, 3, 4, 5, 6].

Today however there exists several new areas of research contributing to enhanced situational awareness, all with their own challenges and limitations. Such tools and technologies include Channel State Information (CSI) to enable multipath exploitation in addition to LoS RSS, Massive Multiple-In Multiple-Out (MIMO), Cognitive Radio (CR), new Millimeter Wave (mmWave) waveforms and the emergence of short range highly connected devices such as Device-to-Device (D2D) and Internet of Things (IoT), with each device contributing additional information about the space around them. The use of AI, more specifically machine learning and deep learning, is also becoming mainstream as a predictive and analytical tool to solve problems that previous approaches have been unable to. Such areas include big data and analytics, cloud/edge computing, cyber security and advanced computer vision systems.

The application of modern tools and technologies to existing situational awareness problems motivates this study. Moreover, little research has focused on the use of wireless CSI and AI in the environment mapping space, central to many modern applications such as assisted navigation, location based services and augmented reality systems where a high degree of measurement accuracy of less than 1m is required.

1.2   Research Aims and Significance

This thesis explores the use of AI and deep learning for the analysis of wireless CSI to predict interior room dimensions. The ability to estimate a room's width, breadth and height with a high degree of accuracy using ubiquitous wireless information, will aid the acceleration of situational awareness techniques into modern technologies benefiting all users be it consumer, emergency services, defence or otherwise.

We therefore pose the question; can we use wireless channel information in the representation of a Channel Impulse Response (CIR) with artificial neural networks, to accurately predict indoor room dimensions? New knowledge is created in this thesis as follows:

- It is demonstrated that given only a single wireless CIR unbounded by bandwidth or multipath strength limitations, Feed-Forward Neural Networks (FFNNs) are able to accurately predict room dimensions.
- For a CIR defined by path delay, magnitude and phase components, it is shown that phase component information provides little to no benefit for predicting room dimensions when other

path components are utilised.

- Exploiting CIRs obtained from multiple receiver locations significantly increases prediction performance. This is useful when few multipaths can be resolved due to physical channel constraints such as bandwidth or noise limitation.

- It is demonstrated when using multiple CIRs with a Recurrent Neural Network (RNN) architecture, only the CIR delay component is needed to accurately predict room dimension. This is applicable when using super-resolution algorithms such as MUltiple SIgnal Classification (MUSIC) and Matrix Pencil Method (MPM), whose capability only support extraction of delay path information, or where additional processing of magnitude and/or phase components are expensive [7, 8].

- Limiting wireless CIRs by bandwidth and received signal power to that of practical systems e.g. Wi-Fi 802.11ac, wideband or Ultra-Wideband (UWB) systems, the tools and techniques developed in this thesis demonstrate that a high level of room dimension accuracy can still be obtained using an RNN.

1.3 Summary of Contributions

The primary contributions in this thesis to the fields of situational awareness and environment mapping are summarised as follows:

- Created a system model utilising Ray-Tracing (RT) for the testing, development and analysis of deep learning techniques exploiting wireless CSI for large-scale indoor environment mapping. The CSI form in this case comprises of a CIR and includes multipath.

- Developed a FFNN algorithm capable of predicting empty room dimensions using a single CIR unaffected by bandwidth or path strength limitation with accuracies of 89% and 99% given estimation dimension error margins of 0.5m and 1m respectively. A successful prediction is registered when a room's width, breadth and height are all estimated to within the respective error margin.

- Developed a FFNN algorithm capable of improving room dimension estimation accuracies to 95% and 99% when utilising 2 CIRs gathered from random receiver locations for dimension error margins of 0.5m and 1m respectively.

- Created an RNN algorithm able to include CIR information gathered from an arbitrary number of receiver locations obtained under a random walk. This algorithm further improved room dimension accuracies to 98% and 99% when utilising upwards of 20 CIR receiver location measurements for dimension error margins of 0.5m and 1m respectively.

- Demonstrated that two new RNN algorithms can be used to accurately predict room dimensions given CIRs physically limited by bandwidth or multipath signal strength with accuracies greater

than 90%. This was achieved for waveforms limited to only 160 MHz in bandwidth, such is present in Wi-Fi 802.11ac, or waveforms limited to a multipath magnitude of −50 dBm with infinite bandwidth utilising various receiver random walk lengths and CIR path component information i.e. delay, magnitude and/or phase.

1.4 Thesis Outline

Chapter 2 provides background into three discipline areas relevant to this research. An introduction to wireless CSI is presented and discusses how such information in the form of a CIR describes multipath in a complex physical environment. We then turn attention to how a CIR can be artificially generated using modern EM propagation modelling techniques. Furthermore, to analyse this information, the concepts of machine and deep learning are also detailed.

Chapter 3 surveys current research literature and establishes how CSI can be applied to achieve situational awareness. Localisation techniques are first explored which allows positional information to be determined. SLAM techniques extend this to jointly include environment mapping. We investigate characterisation techniques which obtain situational awareness through object classification and lastly, we explore how machine learning and deep learning have been applied to these fields.

Chapter 4 introduces our deep learning system model to enable room dimension estimation given either a single or multiple CIR measurements. Such a model was necessary to facilitate the collection and analysis of large-scale CIR measurements using RT tools. Two deep learning FFNN algorithms are presented and demonstrate that simultaneous estimation of a room's width, breadth and height is possible with a high degree of accuracy, up to 98% and 99% given dimension error bounds of 0.5m and 1m respectively, exploiting CIRs unaffected by bandwidth and multipath strength limitation.

Chapter 5 extends our deep learning system model presented in Chapter 4. We present several deep RNN algorithms and demonstrate, that room dimension estimation accuracy can be increased using an RNN over a FFNN architecture when exploiting multiple CIR measurements with a reduced quantity of multipath obtained under a random walk receiver. In addition, we also show that when CIRs are either bandwidth or multipath strength limited to values found in practical systems e.g. 802.11ac, accuracies of approximately 90% and above can be achieved given a dimension error margin of 0.5m when utilising multiple CIRs.

Chapter 6 concludes this thesis. A summary is provided and future research directions presented.

Chapter 2

Background

This chapter introduces three key concepts relevant to the research conducted in this thesis. We begin by defining Channel State Information (CSI) and explain how a useful representation of this exists in the form of the Channel Impulse Response (CIR). We also show how a communication system limited by bandwidth will alter this CIR definition.

Next, the technique of Ray-Tracing (RT) is presented which permits the generation of accurate CIR measurements using simulation and 3D modelling tools [9, 10, 11]. Such tools are necessary for large-scale environment mapping projects, whereby many CIR measurements are required from many transmitter and receiver location combinations, in addition to a large number of rooms which may be otherwise experimentally and practically unobtainable.

Lastly, this chapter introduces the concept of Artificial Intelligence (AI), machine learning and deep learning. Two architectures of deep learning used within this thesis are presented, namely the Feed-Forward Neural Network (FFNN) and the Recurrent Neural Network (RNN), with advantages and disadvantages identified for each.

2.1 Channel State Information

CSI is used in a plethora of modern day wireless telecommunications protocols from consumer based technology such as Wi-Fi and WiMAX, to next-generation cellular protocols like 4G LTE and the soon to be released 5G system [9, 12]. CSI allows end-to-end devices to determine short-term characteristics of the channel between each other, allowing improved data rates, increased Signal to Noise Ratio (SNR) and a reduction in data bit errors [13].

CSI describes the changes a signal undergoes as it propagates from a transmitting device to a receiving device. If the waveform is transmitted from an isotropic antenna in free space, these changes can be described according to the Friis transmission equation

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi d)^2}, \tag{2.1}$$

where P_R and P_T represent the power of the receiver and transmitter in Watts, G_T and G_R are the transmitter and receiver antenna gains, λ is the transmitted signal wavelength in metres and d is the distance in metres between the transmitter and receiver [14].

It is important to note that the Friis equation only describes a signal channel under idealised conditions and represents the Line-of-Sight (LoS) path only. It does not include effects from material interactions as the propagating signal undergoes reflection, refraction and/or diffraction.

Consider a scenario with a generalised complex environment and a wireless signal s(t) radiated from an isotropic antenna (Figure 2.1). It can be seen that the received waveform r(t) will be a summation of many multipath propagations of s(t), each with unique amplitude and phase variations due to their distance travelled and interaction with respective objects. In this figure, a transmission refers a multi-refraction process as a waveform propagates through a material.
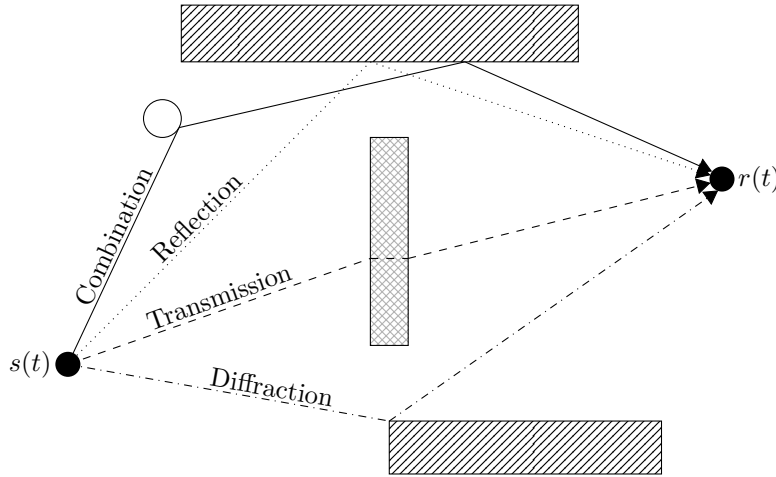
FIGURE 2.1: Multipath in a Generalised Complex Environment

Furthermore, this scenario can be described as a frequency selective multipath fading channel with discrete path gains $\{h(\tau_i : i = 0, 1, \ldots, \Gamma - 1)\}$ where Γ is the number of multipaths including possible paths with zero gains and τ_i is the discrete delay in path i. Assuming perfect transmitter and receiver synchronisation, the received waveform can be written as

$$r(n) = \sum_{i=0}^{\Gamma-1} h(\tau_i)s(n - \tau_i) + w(n), \tag{2.2}$$

where $w(n)$ is often represented as zero-mean additive white Gaussian noise with variance σ^2. Here, $h(\tau_i)$ represents a complex channel and in this form, is also known as the CIR.

Transforming the CIR into the Fourier domain the Channel Frequency Response (CFR) is obtained

$$H(f) = \sum_{i=0}^{\Gamma-1} h(\tau_i) e^{-j2\pi\tau_i f}, \tag{2.3}$$

where $h(\tau_i) = a_i e^{j\phi_i}$ and a_i and ϕ_i represent the complex amplitude gains and phase for each multipath component for each frequency index f. Both the CIR and CFR are representations of CSI as both describe the signal propagation from a transmitter to a receiver.

2.1.1 The effect of bandwidth on a Channel Impulse Response

The following method describes the CIR when a transmitted signal is band-limited to a bandwidth B due to power constraints and spectrum licensing. This was used in Chapter 5 to band-limit RT output CIRs developed in this thesis, to that of practical systems for deep learning analysis.

For a discrete-time channel, according to Nyquist its complex baseband equivalent will be frequency limited to $\pm B/2$ [15]. Following the sampling theorem stating that complex waveforms band-limited to $\pm B/2$ can be expanded in terms of the orthogonal basis $\{\text{sinc}(Bt - n)\}_n$, the complex transmit waveform, sampled at $1/B$, can be represented by

$$s(t) = \sum_n s[n]\,\text{sinc}(Bt - n), \tag{2.4}$$

where $s[n]$ represents $s(n/B)$ and $\text{sinc}(t)$ is the normalised sinc function represented by

$$\text{sinc}(t) = \frac{\sin(\pi t)}{\pi t}. \tag{2.5}$$

Assuming that a received waveform is also sampled at multiples of $1/B$, $r[m] = r(m/B)$, and $r(t) = \sum_{i=0}^{\Gamma-1} h(\tau_i(t)) s(t - \tau_i(t))$ excluding noise, then

$$r[m] = \sum_n s[n] \sum_{i=0}^{\Gamma-1} h(\tau_i(m/B))\,\text{sinc}[m - n - \tau_i(m/B)B]. \tag{2.6}$$

Letting $l = m - n$,

$$r[m] = \sum_l s[m - l] \sum_{i=0}^{\Gamma-1} h(\tau_i(m/B))\,\text{sinc}[m - n - \tau_i(m/B)B]. \tag{2.7}$$

Defining the sampled complex channel with tap l at time m as

$$h_l[m] = \sum_{i=0}^{\Gamma-1} h(\tau_i(m/B)) \operatorname{sinc}[l - \tau_i(m/B)B], \qquad (2.8)$$

the received waveform can be simplified to

$$r[m] = \sum_l h_l[m]s[m-l]. \qquad (2.9)$$

Where the channel gains $h(\tau_i)$ and delays τ_i for each path are time-invariant, the channel simplifies to

$$h_l = \sum_{i=0}^{\Gamma-1} h(\tau_i) \operatorname{sinc}[l - \tau_i B]. \qquad (2.10)$$

This indicates that the wireless channel is equivalent to a low-pass filtered baseband channel response convolved with a sinc filter function. Therefore, as the bandwidth B decreases, the time-domain impulse response of the sinc function will widen resulting in the merging of multipaths whose difference in delays are small relative to the filter pulse width. Furthermore, when the filter bandwidth approaches infinity, the original baseband CIR described by precise time delays can be obtained.

2.2   EM Propagation Modelling Techniques

Electromagnetic (EM) propagation modelling techniques were used in this thesis to generate a multitude of accurate CIRs, representative of wireless channels between a single transmitter and a single receiver for many positions in a room and for many room sizes. Such tools are necessary where extensive physical wireless measurement campaigns are either difficult or practically infeasible.

As stated in Section 2.1, an EM wave propagating in an environment may undergo reflection, scattering through through diffraction, or refraction. The received waveform therefore comprises a summation of many different multipath components, each contributing to signal distortions that lead to a loss of communications link throughput and an increase in Bit Error Rate (BER). The delay, amplitude and phase of each multipath component is a function of the interacted object(s) properties in the environment and the distance travelled.

There are several methods for analytically describing the wireless communications channel; using

full wave techniques described by Maxwell's equations, ray optic techniques such as RT, and statistical methods that describe a more generalist channel when deterministic information is unavailable [10, 16, 17, 18]. These methods provide the foundations for modern day EM propagation modelling due to their reliable results.

It is important to note that all methods previously noted are highly influenced by how the physical environment is defined. This definition is through the specification of dielectric constants unique to each individual material and affect how EM waves are transmitted, reflected, refracted or diffracted.

2.2.1 Dielectric Properties of Materials

EM waves propagate in free space through modulation of alternating electric and magnetic fields. To describe how such waves change as they interact with objects, three important object parameters require definition.

- Permittivity: Measure of a material's ability to resist electric field change. Represents a ratio with reference to a vacuum. Measured in Farads per metre (F/m).
- Permeability: Ability of a material to support the formation of a magnetic field. Measured in Henries per metre (H/m)
- Conductivity: The degree to which a material conducts electricity. It is calculated as the ratio of the current density in the material to the electric field which causes the flow of current. Measured in Siemens per metre (S/m).

Each parameter may vary based on the frequency of the EM wave, in addition to the temperature and humidity of the environment.

2.2.2 Full Wave Techniques

Full wave techniques are the most accurate and complex method for EM modelling and aim to solve Maxwell's equations without using any simplifying assumptions. Maxwell's equations describe the formation of electric and magnetic fields for a self-propagating wavefront.

One such full wave technique is Finite-Difference Time-Domain (FDTD), which uses finite differences to represent the spatial and temporal derivatives in Faraday's and Ampere's laws [16]. Knowing that the electric and magnetic fields of a propagating EM wave are perpendicular, a graphical representation of these fields in space and time can be modelled with arbitrary resolution separating known field strength values. The approximate derivatives are then derived through a Taylor expansion about each field point and an update equation formed to describe future fields as a function of past known field points. The process is repeated in a loop, propagating forward for each time delta and is known as the Yee Algorithm [19].

2.2.3   Ray-Tracing

RT is a method commonly implemented in commercial EM solution software and was used to obtain CIR estimates as part of this thesis in Chapters 4 and 5. In general, it is difficult to solve Maxwell's equations analytically in a complex environment and this typically leads to solutions using asymptotic high frequency techniques such as ray optics. At high frequencies, the electric field, magnetic field and propagation direction are mutually orthogonal with power of the EM wave flowing in a direction perpendicular to the wavefront surface [10]. Assuming that the propagation medium is homogeneous in that its dielectric constants do not change, the ray trajectory will travel in a straight line [10, 17].

The ray optic concept can be related visually to the way a laser beam propagates through air, travelling in a straight tubular fashion until either reflected, diffracted, absorbed or scattered off an object in its path. The techniques to describe these properties include Geometric Optics (GO), for transmissions and reflections, and the Uniform Theory of Diffraction (UTD) for diffractions [20, 21]. Each of these techniques depict the propagating EM wave as a straight line which defines the ray. In reality however ray energy will spread, expanding its cross-sectional wavefront area the further it travels. To model this phenomenon, a single ray is usually represented by a geometric cone called a ray tube. There are a number of different shapes that a ray tube may take based on the ray-tracing model used. Typically, non-overlapping shapes such as triangles or octagons are used to describe the wavefront geometry instead of overlapping shapes such as a circle to avoid counting more than one ray at any point in the ray-tracing model. RT models may start to be generally applied to scenarios where the distance between a transmitter and receiver is large with respect to the wavelength, typically greater than 10λ.

2.2.3.1   Ray-Tracing Algorithms

Several algorithms exist that can be used to trace individual rays through an environment. The first of these methods is the image method. The image method provides the most accurate results for multipath determination and uses a recursive implementation of image trees to obtain a Fermat's principle of least time solution. Fermat's principle states that a ray will take a route that consumes the least time possible travelling from one point to another and applies to all forms of ray propagation mechanisms such as reflection, transmission or diffraction [22].

To obtain the first single-reflection or one-bounce solution, images of the receiver are created by their mirror locations to all visible walls. These mirror locations are often referred to as virtual receivers or transmitters. A line then connects the virtual receiver to the transmitter and where this line intersects the wall, an additional line is drawn to the receiver which forms the complete ray path. A two-bounce solution is found by creating images for each virtual receiver which generates second level branches in the image tree. Transmission through walls and/or other objects are typically taken

into account by simply drawing a straight LoS path between the receiver and transmitter and factoring any loss due to each material encountered.

The Shooting and Bouncing Rays (SBR) method provides an approximation for the total received energy governed by each multipath component [23]. Instead of determining a ray's trajectory between a receiver and transmitter exactly, rays are launched with a predefined angular resolution. This leads to path error as a propagating ray might come close, but not directly intersect with the receiver. This error is usually minimised by creating a receiving sphere of known radius and counting a ray's inbound presence if it intersects this sphere.

Ray angular resolution is generally defined by equally distributed rays. The finest uniformly distributed shape is the regular icosahedron with 20 identical equilateral triangular faces and 12 vertices. For an isotropic antenna pattern, each ray contains equal power. Each ray in addition is represented by a ray tube and can take two forms, lateral ray tube and center ray tube. Lateral ray tubes use the vertices of the regular icosahedron as the edges for the ray tube wavefront. This results in three rays traced for every wavefront as opposed to one. This allows easy determination of wavefront reflection, transmission or diffraction based on six different vertex-to-object face combinations [18]. Center ray tube however launches a single ray from the center of each icosahedron face.

The most commonly used method is a hybrid approach, combining SBR for the lighter processing requirements by quickly trimming erroneous paths, with the image method to maintain accuracy. Once rays are detected by intersection with the receiver sphere, the image method is performed. This traces the ray backwards towards the transmitter, adjusting and fine tuning the ray path recursively.

2.3   Machine Learning

The use of AI or more specifically, machine learning and deep learning, has found its way to dominate new areas of research including big data and analytics, cloud/edge computing, Internet of Things (IoT), cyber security and advanced computer vision systems [24]. It is leveraged in both academia and industry, with some of the largest and most complex systems implemented by Google, Microsoft, Amazon and Apple. It has further been applied to all types of data including visual, audio, numerical text or a combination of these. Examples of mature machine/deep learning applications include but are not limited to those described in Table 2.1. Deep learning was utilised in this thesis to develop environment mapping algorithms for room size estimation given wireless CIR measurements.

There are three main classes of machine learning algorithms; supervised learning, unsupervised learning and reinforcement learning [25, 26]. Supervised learning is used as a predictive mechanism when a known dataset is used to train a predictor. The process of training involves minimising the

difference between known data labels i.e. correct information for each respective data item, and the predicted output. Supervised learning techniques are further broken down into two categories based on the predictive mechanism required, classification and regression. Classification is used to obtain binary or multi-class prediction, regression predicts a continuous-valued quantity.

Unsupervised learning describes techniques that learn where no labelled dataset is available i.e. they only learn from the input data. These techniques are generally employed in such applications as noise and dimensionality reduction (auto-encoders) and clustering (k-means).

Reinforcement learning seeks to learn through cumulative reward by undergoing a series of trial and error events and is largely applied to the field of robotics. Two variations commonly used are Q-Learning and Inverse Optimal Control (IOC). Q-Learning consists of state-action pairs, a reward function for the action taken, and a Q-value function to determine quality of the action taken [27]. In dynamic environments, an expected map is given and Q-learning is used for obstacle avoidance. IOC, also known as inverse reinforcement learning, seeks to recover a reward function given direct demonstration of the optimum action required [28].

Image and Video Recognition	Object and pose detection, compressed video sensing
Audio Processing	Noise reduction, data compression, sound classification
Text Analysis	Translation and personality trait extraction from text
Robotics	Robotic arm movement for manufacturing, navigation, robotic vision
Medical Diagnostic	Detection of diseases, tumours, abnormalities from Computerised Tomography (CT) scans, heart rate and Electroencephalogram (EEG) signal analysis
Computational Biology	Predicting DNA mutation effects and peptide retention time
Physical Sciences	Modelling of cosmological and meteorological phenomena
Economic and Financial	Stock trading, marketing and sales strategies
Cyber Security	Intrusion, Malware, Distributed Denial-of-Service (DDoS) detection

TABLE 2.1: Existing Machine and Deep Learning Applications

2.3.1 Machine Learning and the Neural Network

The basic premise of machine learning is to seek to imitate the learning process of human neurons by creating complex interconnected learning structures. This interconnected structure is referred to as a neural network. Deep learning extends that of machine learning by applying many layers of learning, generally three to one hundred or more, with each layer learning and contextualising new information from data provided by the previous one.

A neural network consists of three main stages; an input layer, hidden layers and an output layer

(Figure 2.2). The input layer reads in a feature vector $\mathbf{x_{N_i}}$ representing data from which to learn at a single time-point [25]. Hidden layers contain the neurons or nodes required to learn structures from the input feature vector. Increasing the number of hidden layers L and the number of nodes within N_l, enables more complex structures to be extracted. The final output layer takes propagated neuron information, and transforms it to a prediction $\mathbf{y_{N_o}}$ suitable for the application context e.g. linear value or probabilities for regression and classification applications respectively.
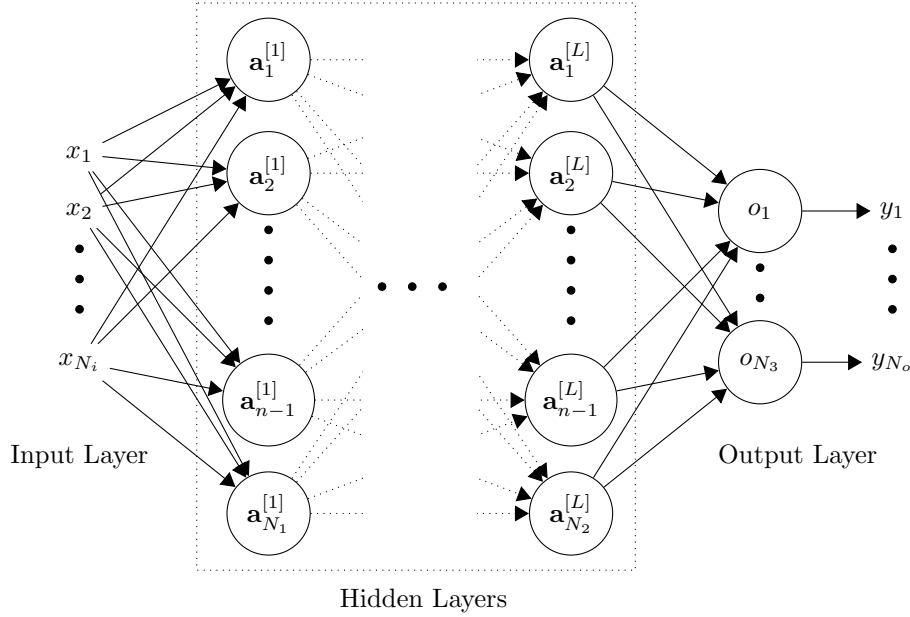
FIGURE 2.2: N-Layer Feed-Forward Neural Network

The method by which a supervised neural network learns is known as backpropagation, derived from backward propagation. To understand how backpropagation works, consider the basic building block for a FFNN, the Perceptron (Figure 2.3). A Perceptron in its forward propagation path consists of a single neuron which takes an input vector $\mathbf{x}$ of any length N_i and processes it through a combination of a weight $\mathbf{w}$ multiplication plus a bias $\mathbf{b}$ addition followed by a non-linear activation function mapping g. Backpropagation therefore is simply the process by which the values of $\mathbf{w}$ and $\mathbf{b}$ are chosen to minimise a cost function $\mathcal{L}(\mathbf{a}, \mathbf{y})$ where $\mathbf{a}$ is the neuron output and $\mathbf{y}$ is the correct label.
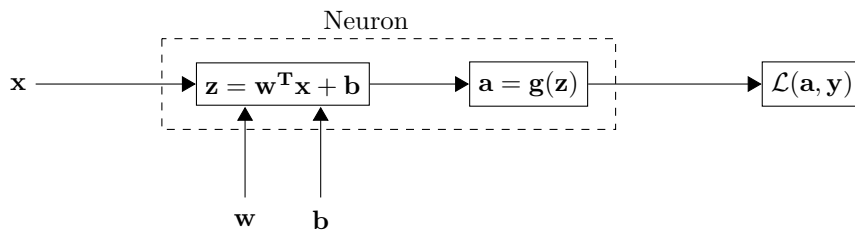
FIGURE 2.3: Single Neuron Perceptron

Backpropagation uses a gradient descent based algorithm which iteratively minimises parameter gradients i.e. their derivatives, in the reverse direction using the chain rule. When the cost function

is convex and its gradient equals to zero, a minima has been reached. Whether it is a local or global minimum depends on several factors, including but not limited to the starting point in the cost landscape (random weight initialisation), rate of change between gradient calculations (learning rate) and the learning optimisation algorithm used (stochastic gradient descent, momentum, Adagrand, Adam etc.).

There are four main equations that govern backpropagation assuming no regularisation. Note that $A \odot B$ is the Hadamard product or element-wise product between vectors A and B.

Error in Output Layer:

$$\frac{\partial \mathcal{L}}{\partial z} = \frac{\partial \mathcal{L}}{\partial a} \odot g'(\mathbf{z}) \tag{2.11}$$

Error in Previous Layers:

$$\frac{\partial \mathcal{L}}{\partial z^{l-1}} = \mathbf{W}^T \frac{\partial \mathcal{L}}{\partial \mathbf{z}^l} \odot g'(\mathbf{z}^{l-1}) \tag{2.12}$$

Error in Layer Weights:

$$\frac{\partial \mathcal{L}}{\partial W^l} = \frac{\partial \mathcal{L}}{\partial z^{[l]}} \odot g(\mathbf{z}^{l-1}) \tag{2.13}$$

Error in Layer Biases:

$$\frac{\partial \mathcal{L}}{\partial b^l} = \frac{\partial \mathcal{L}}{\partial z^l} \tag{2.14}$$

2.3.2 Recurrent Neural Networks

When the input data is described by a sequence, there are several disadvantages to using FFNNs. Namely, (1) the input and output lengths may change, (2) the input layer weight matrix size can grow quite large based on the sequence length, and (3) the neural network model does not share features learnt across different positions in the data.

Consider a neural network designed to input a sentence. One way to circumvent problem 1 would be to pad the input sequence with zeros up to the maximum data width (word length) and sequence length (sentence length). However this is not ideal and will compound problem 2, recalling that the input layer weight matrix size is both a function of the input data dimension and the number of hidden layer nodes. To explain problem 3, we can observe the following two sentences:

- He said, "The bat hung from the tree."
- He said, "The bat hit the ball for six."

From these two sentences it is difficult to determine if the word "bat" directly corresponds to an animal, or an object from the first three words. Information contained at later points of varying delay within the sequence is required which FFNNs are unable utilise.

RNNs are able to solve all three problems and are therefore a popular choice when learning from sequenced information. Figures 2.4 and 2.5 visualise the basic recurrent cell structure. The equations governing forward propagation can therefore be derived as
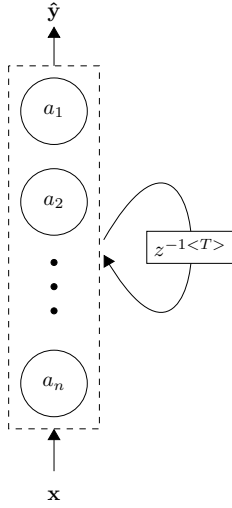
FIGURE 2.4: Recurrent Cell
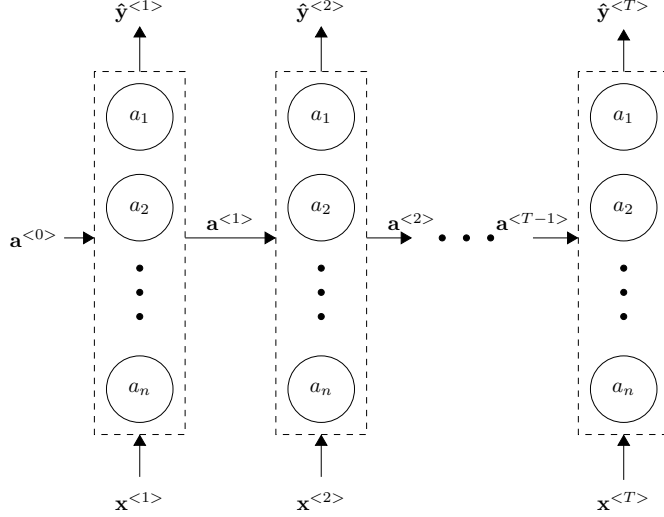
FIGURE 2.5: Recurrent Cell Time Unrolled

$$\mathbf{a}^{<t>} = g(\mathbf{W}_a[\mathbf{a}^{<t-1>}; \mathbf{x}^{<t>}] + \mathbf{b}_a), \tag{2.15}$$

$$\hat{\mathbf{y}}^{<t>} = g(\mathbf{W}_{ya}\mathbf{a}^{<t>} + \mathbf{b}_y), \tag{2.16}$$

where $\mathbf{x}$ and $\hat{\mathbf{y}}$ are a set of input and output vectors of set size T respectively, $\mathbf{W}_a = [\mathbf{W}_{aa}, \mathbf{W}_{ax}]$ and $\mathbf{W}_{aa}, \mathbf{W}_{ax}$ and $\mathbf{W}_{ya}$ are weight matrices for the associative vector multiplications between either the activations, input or output, $\mathbf{a}^{<t>}$ is the output activation at time t, g is the activation function and $\mathbf{b_a}$ and $\mathbf{b_y}$ are the bias vectors for the activation and output. In Figure 2.4, $z^{-1<T>}$ indicates a delay for one time cycle up to a total time T.

2.3.2.1 Gated Recurrent Unit and Long Short-Term Memory Cells

In practice, this simplified version of a recurrent cell requires modification to enhance connection between time distant dependencies. Two widely used RNNs exist that perform this through the addition of a memory cell c, these are the Gated Recurrent Unit (GRU) and the Long Short-Term Memory (LSTM) [29, 30]. Each memory cell implementation varies, providing trade-offs between speed, scalability and flexibility.

GRU Cell Equations:

The equations describing GRU functionality are represented by Equations (2.17) through (2.21).

$$c^{<t>} = a^{<t>} \tag{2.17}$$

$$\tilde{c}^{<t>} = g(W_c[\Gamma_r \odot c^{<t-1>}, x^{<t>}] + b_c) \tag{2.18}$$

$$\Gamma_r = \sigma(W_r[c^{<t-1>}; x^{<t>}] + b_r) \tag{2.19}$$

$$\Gamma_u = \sigma(W_u[c^{<t-1>}; x^{<t>}] + b_u) \tag{2.20}$$

$$c^{<t>} = \Gamma_u \odot \tilde{c}^{<t>} + (1 - \Gamma_u) \odot c^{<t-1>} \tag{2.21}$$

Equation (2.17) implies that the memory cell output $c^{<t>}$ assumes the role of the activation $a^{<t>}$, and propagates between each RNN time instant. An updated memory cell candidate is represented by $\tilde{c}^{<t>}$ and is dependant on relevance gate Γ_r. If the old memory cell value $c^{<t-1>}$ is determined relevant, Γ_r will be close to 1 and weigh heavily on the updated memory cell candidate value. To determine if the memory cell output is updated and by how much, a second gate Γ_u is employed. A sigmoid activation function is referred to as σ. An output prediction $y^{<t>}$ can be made Equations (2.16) and (2.17).

LSTM Cell Equations:

The LSTM cell equations are given in (2.22) through (2.27).

$$\tilde{c}^{<t>} = g(W_c[a^{<t-1>}, x^{<t>}] + b_c) \tag{2.22}$$

$$\Gamma_u = \sigma(W_u[a^{<t-1>}; x^{<t>}] + b_u) \tag{2.23}$$

$$\Gamma_f = \sigma(W_f[a^{<t-1>}; x^{<t>}] + b_f) \tag{2.24}$$

$$\Gamma_o = \sigma(W_o[a^{<t-1>}; x^{<t>}] + b_o) \tag{2.25}$$

$$c^{<t>} = \Gamma_u \odot \tilde{c}^{<t>} + \Gamma_f \odot c^{<t-1>} \tag{2.26}$$

$$a^{<t>} = \Gamma_o \odot \tanh c^{<t>} \tag{2.27}$$

The differences between the GRU and LSTM cells is that (i) the LSTM propagates both $a^{<t>}$ and $c^{<t>}$ through time, with $a^{<t>}$ resuming its place as per the original RNN Equation (2.15), and (ii) the relevance gate Γ_r has been removed and two additional gates specified. The forget gate Γ_f replaces the $(1 - \Gamma_u)$ term thus providing more control regarding when the memory cell output is updated. The output gate Γ_o again provides additional control over when the output activations are updated.

Therefore in general, GRU cells may be selected for applications where speed or a large neural network is required due to their simpler structure with only two gates present. Where application flexibility is preferred, the LSTM with three gates is more suitable. Typically for both GRU and LSTM, $g(z) = \tanh$. Backpropagation in an RNN is performed the same way as per the FFNN. In this thesis, the LSTM cell was implemented within an RNN architecture to achieve a multi-CIR algorithm for the accurate prediction of room dimensions.

Chapter **3**

Literature Review

With the definition of Channel State Information (CSI) now established, this chapter provides a comprehensive literature review on how multipath information contained within CSI, can be used to perform localisation, characterisation and environment mapping. Moreover, the application of Artificial Intelligence (AI) and machine learning to localisation and characterisation using wireless waveforms is reviewed. From this, it is concluded very little information exists relating directly to the use of machine learning when utilising wireless information for environment mapping.

3.1   Localisation Techniques

Localisation exploits wireless information and geolocation techniques to derive the position of either a transmitter or receiver. Depending on the type of wireless information used, several techniques exist to perform localisation, each with varying degrees of accuracy. We begin by exploring these techniques in detail.

3.1.1   Received Signal Strength Indicator Techniques

Wi-Fi is one of the most ubiquitous wireless protocols, found in almost every household or workplace. As a result, there has been much research conducted on Wi-Fi, exploring ways in which it can be leveraged for a variety of applications other than its original function of Wireless Local Area Network (WLAN) high speed data delivery [2, 4, 5, 6, 31, 32, 33, 34, 35, 36, 37, 38, 39].

For example, training data present in the Wi-Fi packet preamble has been applied to both transmitter and receiver localisation. Such training data includes the Legacy non-high throughput Long Training Field (L-LTF) used for packet detection, timing, carrier frequency offset estimation and channel estimation. A Received Signal Strength Indicator (RSSI) value can be calculated based on the channel estimates **h** derived through equalisation of this field.

17
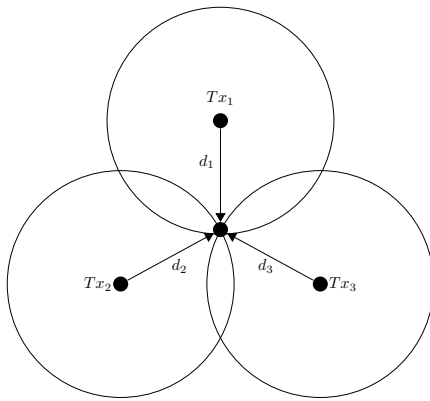
FIGURE 3.1: Time-of-Arrival (ToA)
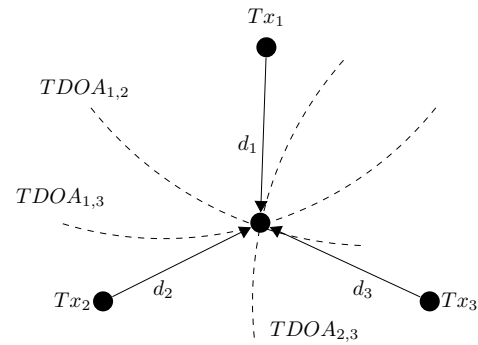
FIGURE 3.2: Time-Difference-of-Arrival (TDoA)

$$r_{RSSI}(dB) = 10log_{10} \parallel \mathbf{h} \parallel_2 + \mathbf{w}, \tag{3.1}$$

where $\mathbf{w}$ represents zero-mean additive white Gaussian noise with variance σ^2. Feeding this RSSI value into a path-loss equation such as the COST Hata model for urban environments, an approximate distance between the receiver and transmitter can be obtained. This process is repeated for different receiver locations until a transmitter location error ellipse is minimised.

RSSI based localisation techniques work well in high Signal to Noise Ratio (SNR) environments and where direct Line-of-Sight (LoS) is available. However in an indoor environment rich in multipath, or where a transmitter shares no LoS with a receiver, RSSI can produce erroneous results leading to poor localisation accuracy [2, 3, 6, 31].

3.1.2 Time-of-Arrival and Time-Difference-of-Arrival Techniques

An alternative method for determining distance between a transmitter and receiver is by directly measuring the time taken for a signal to propagate between each node. This is referred to as Time-of-Flight (ToF), alternatively known as Time-of-Arrival (ToA) (Figure 3.1). ToA methods rely on the accurate determination of arrival time for the leading LoS path. In an indoor environment, the presence of dense multipath can interfere with this time estimate by 'blurring' the LoS peak within the Channel Impulse Response (CIR).

Humphrey and Hedley [40] explore super-resolution techniques to separate multipath and reveal the true LoS CIR peak. Super-resolution techniques allow the extrapolation of high frequency, high-time resolution spectrum from low frequency, low-time resolution data [41]. Two common signal-subspace, super-resolution methods examined here are MUltiple SIgnal Classification (MUSIC) and Matrix Pencil Method (MPM).

MUSIC provides super-resolution as it can be evaluated at any frequency rather than those at

well-defined intervals under a Discrete Fourier Transform (DFT) [8]. Firstly, a set of steering vectors (manifold) is created, responsible for the search space for super-resolution signal detection. Each steering vector comprises of complex exponentials, that depending on the application e.g. ToA or Angle of Arrival (AoA), are a function of predetermined time delays or incident angles respectively. MUSIC then performs a Singular Value Decomposition (SVD) on the autocorrelation matrix of a received waveform to determine its signal and noise subspaces. As each steering vector exists in the signal subspace, a peak will be produced when a selected steering vector becomes mutually orthogonal to the noise subspace i.e. their dot product equals to zero.

The MPM uses a set of damped complex sinusoids to approximate a measured discrete waveform [7]. In the presence of noise, the MPM also requires use of an SVD. Firstly, to pre-filter the measured waveform and estimate the signal subspace size, and secondly to determine the poles of each damped complex sinusoid via the eigen-values of a matrix pencil. A matrix pencil consists of two matrices $\mathbf{Y}_1$ and $\mathbf{Y}_2$, which are functions of the singular vectors from the pre-filtered SVD stage. These combine linearly form $\mathbf{Y}_2 - \lambda \mathbf{Y}_1$, where λ is a scalar parameter. With the poles for each complex sinusoid known, respective amplitudes and phases may be determined via a least squares solution.

Humphrey and Hedley compare MUSIC, MPM and the non super-resolution Fast Fourier Transform (FFT) algorithm and demonstrate that an FFT can be used to obtain accurate ToA. This however relies heavily on further processing that sets a magnitude threshold of the LoS path, to a percentage of the leading rising edge in the CIR.

Super-resolution techniques described by Humphrey and Hedley provide a method for extracting multipath components from wireless system channel estimates. In a complex environment, the leading edge of the CIR may not provide enough information to extrapolate a probable room geometry.

Increased time resolution comes at the cost of increased signal bandwidth, as bandwidth is inversely proportional to the pulse width of a signal. Wi-Fi 802.11ac permits a contiguous instantaneous bandwidth of 160 MHz occupying a total of 35 channels in the Industrial, Scientific and Medical (ISM) frequency band of 5 GHz. Only 100 MHz bandwidth is permitted at 2.4 GHz [12, 13, 42, 43]. To obtain bandwidths larger than 160 MHz and therefore finer time resolution for ToA estimation at a decimeter level, Vasisht et al. [44] propose a technique utilising ToFs for all 35 Wi-Fi channels. ToFs are estimated for each channel by performing a non-uniform DFT at each frequency, followed by an Inverse Discrete Fourier Transform (IDFT) to obtain time delay CIR information. ToA is calculated from the LoS path only in CIRs that are sparse i.e. where only a few multipaths are present and each multipath component presents a single distinct time delay peak. The overall true ToA is estimated as the point all frequency band ToFs closely match. Random phase offsets incurred as a result of non-contiguous frequency channel band jumping are compensated for by assuming reciprocity in the forward and reverse channels.

Due to this reciprocity, existing techniques for increased bandwidth and fine-time resolution CIR estimation using Wi-Fi are only applicable to slow varying signal subspace applications, as a channel must remain constant during the entire information exchange round trip time. Use is also limited to high SNR, low-complex environments characterised by both few and sparse multipath.

If multiple time-synchronised receivers are available, Time-Difference-of-Arrival (TDoA) may be used to localise a transmitted waveform (Figure 3.2). TDoA differs from ToA in that it uses the difference between arrival times of the transmitted waveform at each synchronised receiver rather than the absolute received time.TDoA methods do not require knowledge of the time a waveform was transmitted and therefore are more versatile than ToA techniques. Shi et al. [45] compare two super-resolution techniques, MUSIC and Volume Cross-Correlation (VCC) for TDoA estimation. VCC is similar to MUSIC in that they both exploit the subspace structure of the waveform using an SVD. The singular values however are combined through multiplication to define the matrix volume. This is then used to generate the VCC function. It is demonstrated that the VCC function presents lower Mean Squared Error (MSE) for slow-varying subspace signals whilst detecting more multipath delays for fast-varying subspace signals when used in combination with signal interpolation prior to SVD.

VCC can therefore be used to reveal additional and weaker multipaths from a complex environment over widely used techniques such as MUSIC. Such information will assist environment reconstruction algorithms reliant on multipath detection, however its use is limited to applications where multiple receivers are available and TDoA techniques employed.

3.1.3 Fingerprinting Techniques

Channel Fingerprinting is an alternative technique for receiver localisation. These techniques comprise of two key stages; a prior training phase, and a post estimation phase. The training phase or offline phase, involves recording wireless information at various locations in the environment as a prior step before localisation estimation. Such information may include the CIR and Channel Frequency Response (CFR), or statistics derived from these such as mean, variance and peak value etc. This information therefore acts as a wireless "fingerprint" and is unique to each recorded location. Fingerprints are stored in a database for future reference.

The post estimation phase, or online phase, for wireless fingerprinting compares previously stored fingerprints acquired during the offline phase, to new fingerprinting information gathered by a receiver. Location estimation may then be achieved by selecting a location whereby the respective fingerprint maximises the cross-correlation between itself and the new online information. It is important to note that variation will exist between multiple fingerprints at the same location and between online information derived periodically over time due to physical or atmospheric changes of the environment.

There are a multitude of different ways to generate the wireless fingerprint and also compute the

location estimate. Chapre et al. [37] present a simplistic approach for fingerprint generation by creating a single complex value representative of a pseudo averaged CFR. Here the CFR matrix is summed across all Multiple-In Multiple-Out (MIMO) antennas to create a single averaged subcarrier channel vector. The difference between each subcarrier value is taken as a means of data compression and the average of this vector used to obtain the final complex fingerprint. To determine a location estimate, both the simple machine learning algorithm k-Nearest Neighbours (kNN) and the probabilistic Bayes rule are explored.

kNN is a supervised machine learning algorithm which can be used to classify new data into known data classes. The k parameter in the instance of localisation using fingerprints, determines how many nearby (most similar) fingerprints gathered during the offline training phase, we compare our new online fingerprint with. A class decision and therefore the determined location, is then formed based on which location is most common [25]. To determine a localised receiver position, Chapre et al. set the value of k to 1. This reduces the location decision to the minimum Euclidean distance between a single offline stored fingerprint and the new online fingerprint. In the event of significant noise, the k parameter can be increased to include more location comparisons. The drawback of this would be poorer location estimation resolution.

Bayes rule is used to calculate the conditional probabilistic likelihood of a receiver existing at a past location $\{l_i : i \in 0, \ldots, L-1\}$ given the new online fingerprint H_{test} i.e $P(l_i|H_{test})$, where L is the total number of measured locations. Bayes rule states

$$P(B|E) = \frac{P(E|B)P(B)}{P(E)},$$

(3.2)

where $P(B)$ is the probability of the belief i.e. hypothesis, and $P(E)$ is the probability of the evidence. Bayes rule is useful as the probability of a hypothesis is not only based on how well data fits the hypothesis $P(E|B)$, but also how sensible it is through a-priori knowledge i.e. $P(E)$ and $P(B)$ [26]. In this case, prior probabilities $P(l_i)$ and $P(H_{test})$ as well as likelihood $P(H_{test}|l_i)$ are determined by correlation between the offline and online fingerprints and direct evaluation.

Representing the complex CSI fingerprint as a single complex value through averaging, all frequency selective fading effects over each Wi-Fi subcarrier are lost. For localisation applications where fine spatial resolution is required, such techniques reduce localisation estimation accuracy. However where prior training requires a large amount of template storage, such representation would provide significant data compression capability.

Fang et al. [46] also use a probabilistic approach for location likelihood estimation using Wi-Fi CSI. Offline fingerprints however, are generated through a three step process using a wavelet transform. There are two types of wavelet transforms, the Continuous Wavelet Transform (CWT) and

Discrete Wavelet Transform (DWT). Both types of transform conceptually decompose a waveform into a combination of wavelets, finite time oscillations of known frequencies. Each wavelet may vary in both amplitude and time duration. The wavelet time duration is determined by factors of $\{2^i : i = 1, 2, 3, \dots\}$ and is referred to as an octave. Each increment of i results in a time window octave reduction of 2^i and therefore an inverse increase in wavelet frequency. The main difference between CWT and DWT lies in that for the CWT, each octave may be further sub-divided into intermediate wavelets known as voices. Because DWT is constrained to wavelet durations multiples of 2^i, it may be efficiently implemented using a cascade set of high and low pass Finite Impulse Response (FIR) filter banks known as the Multilevel Discrete Wavelet Transform (MDWT).

Fang et al. first performed the MDWT on the received CSI information and normalised the wavelet coefficients using a process called Histogram Equalisation (HEQ). HEQ takes a cumulative distribution of a set of data points, in this case the wavelet coefficients, then transforms it into a uniform distribution. This process is widely adopted in image processing to improve contrast of an image that is saturated by light areas or deep shadows. The localisation fingerprint is then generated by performing an inverse MDWT on the equalised wavelet coefficients. Using this method, mean localisation error is shown to be significantly reduced over RSSI and raw CSI fingerprinting methods.

For applications using physically obtained CSI, performing noise removal on channel estimates during preprocessing would likely provide significant benefit. Such techniques could involve wavelet transformations for situations where the number of multipaths are unknown, or signal subspace methods for example MUSIC if an estimate for the number of multipaths can be achieved. The use of HEQ in addition to MDWT by Fang et al. evokes interest for such general CSI preprocessing noise removal applications.

In situations where access to an area is unavailable, fingerprints must be generated using methods that aim to predict the wireless Wi-Fi channel in an indoor environment. Indoor channels may be estimated using ray optic techniques such as Ray-Tracing (RT), or through statistical methods. RT fingerprinting is demonstrated by Tseng et al. [47] and generates the CIR using the Shooting and Bouncing Rays (SBR) technique. RT is designed to characterise the deterministic aspects of CSI. In this case, for an Orthogonal Frequency Division Multiplexing (OFDM)-like 100 MHz bandwidth waveform. To capture the random aspects of CSI, additional fingerprints are generated by directly obtaining the CIR using a channel sounder. An assisted ray-tracing algorithm is then achieved by merging both fingerprint types. This fingerprint hybrid is demonstrated to reduce conventional CIR fingerprinting localisation error by 25% [47].

This technique however requires physical access to an environment to directly obtain the CIR. Random CSI perturbation however, can be created to more closely model physical environment characteristics. This could be performed for example, by placing probability densities on RT material parameters such as permittivity, permeability and conductivity, in addition to including modelled

indoor noise.

Humphrey and Hedley [48] generate CIR templates based on Monte Carlo simulation of the Saleh-Valenzuela propagation model. The Saleh-Valenzuela model is a statistical time-domain model developed at Bell Laboratories for the simulation of indoor radio channels [49, 50]. It describes the concept of different multipath components arriving as clusters modelled with a Poisson distribution with an averaged inter-arrival time. Additionally, clusters are also described by a Poisson distribution with a larger inter-arrival time. The leading term of the generated CIR was then exploited for ToA estimation.

In summary, templating techniques for environment mapping may not be ideal if access to an area is restricted. Deterministic modelling of the wireless channel for template acquisition can be achieved through methods such as ray-tracing. Its use however, will more likely be limited to only a small set of known environment. Statistical time domain modelling on the other hand provides benefit for areas where fine details about an environment are unknown. However, they may not correctly represent an individual environment to produce accurate channel state information estimates. Therefore, careful consideration must be used when selecting either a statistical model, deterministic model, or a combination of these, for application in environment mapping.

3.2   Simultaneous Localisation and Mapping

Simultaneous Localisation and Mapping (SLAM) is a process by which an autonomous device can jointly build a map of an environment whilst estimating its own location from this map [51, 52, 53]. There are several classes of SLAM algorithms, one of which is feature based SLAM that uses sparse landmark information to generate 2D or 3D high resolution environment maps [54].

Leitinger et al. [55] use feature based probabilistic SLAM in combination with multipath CSI. The estimated map is extracted from features described as virtual anchors, also known as virtual Access Points (APs). The locations of the virtual APs represent the mirror images of true Wi-Fi APs with respect to each wall present in the environment and thus, also represents the main one-bounce reflected signal path directed from the respective Wi-Fi AP to the autonomous device.

The authors use a standard Markovian motion model to determine new location and orientation vectors based only on a past location estimate $\mathbf{x}_{k-1}$ and a recently applied control input $\mathbf{u}_k$. However, they also modify this model to only rely on the set of all landmark observations without conditioning on the control inputs to compute the probability distribution functions. Distance estimates using the determined multipath components for all virtual APs are fed into an Extended Kalman Filter (EKF) to compute the measurement model.

The authors extend their work to develop a feature based SLAM algorithm using CSI without using prior knowledge of the environment [56]. This algorithm uses Joint Probabilistic Data Association (JPDA) based on Belief Propagation (BP) and Ultra-Wideband (UWB) transmit signals. Distance information regarding each multipath component is fused with additional sensor outputs to create a multi-sensor approach to support feature association when multipath component associations are ambiguous. BP factorises the joint posterior Probability Density Function (PDF) into a factor graph which is a function of the agent state, feature states, associations between the agent and features, feature detection probabilities and multipath component distance measurements. In graphical form, BP can be used to determine estimates for the marginal probabilities and hence joint posterior PDF since the resulting graph contains loops. This is referred to as Loopy Belief Propagation (LBP).

An interesting assumption to note is that each environment feature e.g. a wall, can generate at most one measurement and each measurement can be generated by at most one feature. This works well if all multipath components are generated only through reflection, which based on the environment, may or may not be the case.

In SLAM, the ability to identify that a past location has been visited is known as loop closing. Loop closing is important in large scale mappings, as failure to achieve closure results in the inability to correct for trajectory drift accumulated over time. Kudo and Miura [39] utilise Wi-Fi RSSI fingerprints in conjunction with the Wi-Fi APs unique Basic Service Set Identifier (BSSID) to aid in loop closure. Wi-Fi signal data is smoothed and fed into an Iterative Closest Points (ICP) algorithm for pose sequence matching between current and past positions. Graph based SLAM is finally used to generate the final map output.

The pose sequence matching algorithm used by Kudo and Miura is reliant on Wi-Fi RSSI. Therefore, significant distance measurement error may occur if an AP is obstructed. Using Wi-Fi training data that encodes multipath information as an alternative, or increasing the number or APs and thereby unique RSSI and BSSID pairs, may help to reduce localisation error.

Graph based SLAM models the probability distribution function as a Dynamic Bayesian Network (DBN) [53]. The DBN is represented via a graph and describes the temporal structure of SLAM as the device updates to a new pose via control and observation inputs. The spatial constraints between poses created through observations and odometry measurements may also be encoded in an alternative graphical representation. An example includes Wi-Fi Graph-SLAM which uses Received Signal Strength (RSS) measurements with pedometry and gyroscopic information to estimate both the trajectory and map to achieve an accuracy of 2.23m ±1.25m [3].

3.3 Characterisation Techniques

Wireless characterisation techniques aim to extract information about the physical environment from how the signal interacts with objects present within this environment. This interaction could describe static objects via reflection, transmission or diffraction properties, similar to the mapping process of SLAM. Or it might also describe mobile objects through temporal changes to their CSI.

Holl and Reinhard [38] use Wi-Fi holography to generate a three dimensional image of static objects within a room. This 3D image represents a composite of 2D images or holograms, described by the amplitude and phase variations of a propagating Wi-Fi 802.11ac waveform using reference and scanning antennas. To obtain the hologram complex field, the scanning antenna is normalised by the reference waveform (homodyne scheme). Objects are identified through shadows in the light cone of the generated hologram.

Amplitude and phase variations of a propagating Wi-Fi 802.11ac waveform may also be obtained directly from transmitted Wi-Fi training data. Karanam and Mostofi [57] use Wi-Fi RSSI to directly generate a three-dimensional through-wall imaging framework based on Markov Random Field (MRF) modelling, LBP and sparse signal processing. The framework begins by modelling the power of the received signal as a function of objects in the area. To do this, a linear based LoS approximation of the volume-integral wave equations were performed. The 3D environment model area is then partitioned into cubic volumes known as volumetric pixels or voxels. These voxels provide an object map representation of the area.

Sparse signal processing is then used to minimise the total variation of this object map, obtaining grey scale intensities for each voxel. LBP with MRFs, captures spatial dependencies amongst local voxel neighbours. These are both applied to obtain binary conversion of the grey scale intensities for the final 3D object map. Using Wi-Fi RSSI measurements only for holography, models objects in the area as a function of the received signal power. An alternative approach could exploit multipath information present in additional Wi-Fi training fields.

Analysis of the temporal variation of Wi-Fi CSI for the use of human detection has also been widely explored [33, 58, 59, 60]. Zhang et al. [33] define a centimetre scale accuracy algorithm for human activity sensing using 802.11n. The high accuracy localisation method uses the concept of Fresnel zones, which are a series of concentric ellipsoids of alternating strength arising from constructive and destructive interference of multipath Electromagnetic (EM) waves. An object moving through these zones will create a series of valleys and peaks proportional to the signal wavelength. Fine-grained accuracy is achieved through difference variations between Fresnel zones of all Wi-Fi subcarriers, where each subcarrier is transmitted on a different frequency thus incurring larger or smaller Fresnel zones.

Zhang et al. note that it is difficult to correlate changes of the CSI waveform to human movement in a dynamic geometric space. Therefore, detecting any phase change incurred through a dynamic reflected wave front must occur in a static environment.

Zeng et al. [60] develop a system for person identification based on walk and gait analysis of Wi-Fi CIRs. Their method consists of three key stages: pre-filtering of the CIR data to remove distant multipath i.e. multipath components with a delay in excess of 0.5 microseconds; walk detection using a threshold FFT technique; and gait analysis using statistical data of the CFR and CIR in combination with a machine learning classifier. Temporal fluctuations in the CSI translate themselves into low frequency components less than 10 Hz. Performing an FFT on this data reveal slow, medium and fast human movement.

An alternative to CSI temporal variation for human detection is to observe correlations between each Wi-Fi subcarrier. Zhu et al. [32] also pre-filter CSI data prior to analysis. Pre-filtering is performed through a multi-stage process that first removes CSI outliers for each subcarrier using a Hampel identifier. High frequency components are then removed via a wavelet transform with the CSI data subsequently interpolated to smooth any missing subcarriers. Correlations between each subcarrier are found using Principal Component Analysis (PCA) and Eigenvalue Decomposition (EVD). Human detection thresholding is then performed via a machine learning Support Vector Machine (SVM) classifier. Robustness of human detection is increased through a majority voting scheme between each spatial stream of a MIMO Wi-Fi channel.

3.4 Localisation and Characterisation using Machine Learning

With the rise of AI through machine and deep learning algorithms, many authors have recently begun to apply these techniques to problems in wireless localisation and characterisation. These algorithms are designed to select the best feature set that represents variation in the wireless data (auto-encoders), learn from this feature set (neural networks), and output information that informs us in a way that otherwise would have been untenable just through raw input observation or by other more traditional analytical means.

Wang et al. [61] use machine learning for CSI based fingerprinting in an indoor environment. A four hidden layer deep Feed-Forward Neural Network (FFNN) is constructed to learn position locations based on CSI amplitude only data. Neuron count in each of the four layers are 500, 300, 150 and 50 respectively. The authors select weights present at each hidden layer neuron as the final position fingerprint.

Training of the FFNN is performed in two stages; an initial unsupervised stage, followed by

supervised learning for weight fine tuning. In this instance, Restricted Bolzmann Machines (RBMs) are used to first train the network in a recursive manner. Once neuron weights are approximated, these weights are used to construct an approximate representation of CSI amplitudes which are used for standard supervised learning backpropagation with the input training data. 1000 Wi-Fi packets are collected at each location for offline training.

The localisation accuracy of the FFNN corresponds to an error of less than 1m for 60% of the test locations. This is significantly better than a traditional CSI amplitude localisation algorithm Fine-grained Indoor Fingerprinting System (FIFS), with an accuracy of 25% of locations with an error under 1m. FIFS requires CSI amplitudes across multiple antennas to be averaged prior to location determination resulting in information loss. Both FIFS and the FFNN implemented by Wang et al. are limited to 30 subcarriers across 3 antennas as their input data.

Several of these authors have used a similar neural network for machine learning fingerprinting based on CSI phase information [62]. Prior to passing phase information into the machine learning algorithm, sanitisation of this data must be undertaken. Phase sanitisation includes removing errors through random phase offsets due to sampling frequency offsets and non-synchronous clocks, in addition to unwrapping the phase which is performed by subtracting multiples of 2π.

The deep neural network consists of three hidden layers with 60, 30 and 15 nodes in each layer respectively. The input layer takes a vector of 90 elements representing CSI sanitised phase data points. 800 Wi-Fi packets are collected at each location for offline training.

It is our understanding that this paper by Wang et al. is the first paper to explore the use of CSI phase information as a feature set for localisation. In addition, Wang et al. also verify that positional wireless information is both encoded in the phase component of CSI and can be extracted through sanitisation and machine learning processes.

A different approach to localisation using deep learning is performed by Zhou et al. [63], who use both Support Vector Regression (SVR) and Support Vector Classification (SVC) for location coordinate determination using de-noised CSI amplitude data. SVMs are generally used for classification problems whose goal is to separate two clusters of data by a hyperplane that minimises the distance between them. If the constraints for this hyperplane are changed slightly, the hyperplane will attempt to minimise an error between all points within an error margin and thus, provide a regression estimate. Kernels can be used to fit the hyperplane on higher dimensional data. PCA was used to compress the raw CSI feature set comprising of amplitudes for 30 Wi-Fi subcarriers across 4 spatial streams and two APs, to a feature vector of length 93 and setting a cumulative contribution rate of 99%.

The authors first train two SVR models to establish a functional relationship between the feature set and each respective location axis. The CSI fingerprints and locations are considered non-linear and therefore a Radial Basis Function (RBF) is used to map fingerprints to a higher dimensional feature

space. To determine a location using SVC, a total of 190 SVM classifiers are used to support multi-class classification. The same localisation processes were applied for RSSI data and the localisation accuracy across all methods and feature sets compared. It was shown that SVR performs better than SVC for both CSI and RSSI data.

It is clear from both [62] and [63] that filtering of raw, physically obtained input data prior to machine learning is an important step. PCA is shown as a useful algorithm to both filter CSI amplitude data in addition to reducing the dimension of the feature space, which can lead to a simplification of the entire neural network. It is also interesting to note that only 100 packets were selected from the center of a larger pool containing 1200 packets as the input data for each position. Increasing this number may assist in neural network training and help to prevent future overfitting.

Machine Learning algorithms have further been used for the purpose of human movement classification [34, 35, 64, 65]. Zhao et al. [66] use machine learning for through-wall human pose estimation. Two machine learning algorithms are used in tandem for cross-model supervision, a process by which output from one algorithm (teacher network) is used to train the other (student network). In this instance, a teacher network fed by camera RGB images performs initial 2D pose estimation whose output comprises of key-point confidence maps. Processing is performed by a two branch multi-stage Convolutional Neural Network (CNN). The first branch predicts confidence maps for human body location and the second stage predicts Part Affinity Fields (PAFs), a set of 2D vector fields that encode the location and orientation of limbs over the image domain. A CNN has two main parts; a feature extraction stage, and a classification stage. The feature extraction stage uses sliding matrix convolutional filters followed by pooling to reduce the amount of input information passed to a fully connected classifier network.

The student network is designed to convert the Radio Frequency (RF) data also into keypoint confidence maps. This is performed using vertical and horizontal RF encoders with a successive pose decoder, and utilises both strided and fractionally-strided CNNs. The stride of the convolutional filter defines the step size when shifting the kernel along the input data matrix. A stride size greater than 1 leads to downsampling. Fractionally-strided convolution attempts to reconstruct spatial resolution similar to deconvolution, however using a different process called transposed convolution. The network is trained using 70% of data from the visible scene with the remaining 30% used for cross-validation.

To achieve fine pose estimation, the authors use a radar based chirp waveform instead of Wi-Fi. This waveform occupies a much larger bandwidth, between 5.46 GHz and 7.24 GHz, than Wi-Fi and therefore provides higher time-frequency bandwidth resolution. The use of a teacher machine learning algorithm for the training of a student network is a unique solution and could be applied to autonomous systems with multiple information sources.

Chapter **4**

Room Dimension Estimation using Feed-Forward Neural Networks

This chapter presents our design methodologies and results for the challenging problem of room dimension estimation using wireless Channel State Information (CSI). More specifically, we simultaneously estimate a room's width, breadth and height using two developed Feed-Forward Neural Network (FFNN) algorithms given either a single or multiple Channel Impulse Responses (CIRs) respectively. Here, a single CIR describes the channel between one Transmitter (Tx) location and one Receiver (Rx) location and was generated using Ray-Tracing (RT) software. Each CIR was sampled from 118 different room sizes and was unaffected by bandwidth and noise limitation.

Accuracies as a percentage of successful room dimension predictions are presented rounded to one decimal place and demonstrate that multipath present in a CIR can be exploited using deep learning techniques. The level of accuracy in this case was dependant on a successful prediction threshold value, either 0.5m or 1m, and indicates the maximum prediction error permitted for each of the three room dimensions. A percentage of successful predictions is then obtained by registering multiple successes for all neural network input testing CIRs.

We begin by detailing our developed deep learning system model to enable large-scale CIR room dimension estimation. Such a model was required to facilitate the acquisition of over $850,000$ wireless CIRs, obtained from many transmitter and receiver location combinations, for multiple room sizes.

4.1   System Model

The multi-staged system model implemented to generate CSI and process it using deep learning is shown in Figure 4.1. Firstly, CIRs for a large number of different room sizes were generated using RT software. Metadata obtained from the RT output, namely room dimensions with transmitter and receiver locations, was then used within a CIR preprocessing stage, to reformat each CIR and generate

the necessary Machine Learning (ML) labels required for supervised deep learning. The CIR data along with its respective labels were then read, preprocessed using critical algorithms for ML training, then trained upon using deep neural network algorithms developed in this thesis. TensorFlow, an open source ML Application Programming Interface (API), was used to construct our deep neural network algorithms [67]. Results were evaluated using Matlab.
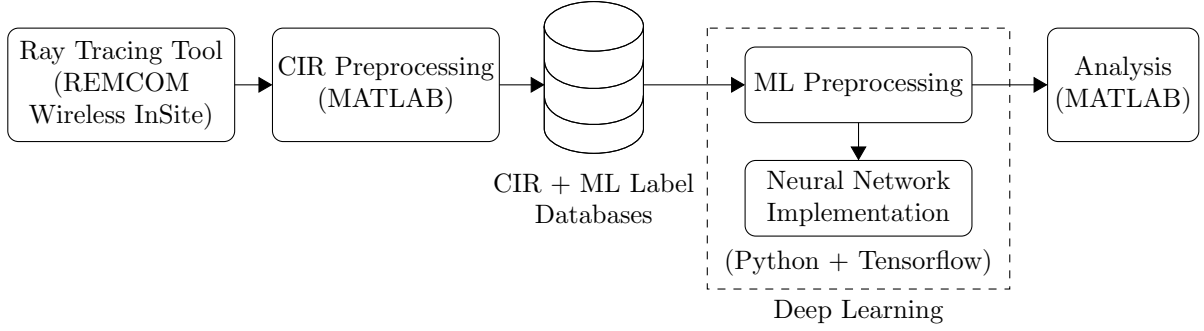
FIGURE 4.1: Software Model for Channel Impulse Response (CIR) analysis using Machine Learning

4.1.1 Ray-Tracing

RT was used to generate $857,859$ CIRs across 118 different room sizes, with possible transmitter and receiver locations arranged according to a grid spacing of 1m×1m. The software selected to perform indoor ray-tracing was Wireless InSite by REMCOM. This industry leading software allows both indoor and outdoor wireless 3D propagation modelling. It further includes a materials database to model the exact propagation effects due to interaction between materials and waveforms at various frequencies, plus provides a high degree of 3D model tailoring due to a large amount of available tuning parameters. Refer Table A.1 for a complete RT parameter listing used in this thesis.

Each output CIR multipath is characterised by three parameters; phase (degrees), exact delay (seconds) and magnitude (dBm). Consequently, with reference to the time-bandwidth product of a signal, each CIR may therefore denote a channel through which an infinite bandwidth signal has been transmitted. To limit computational complexity, a limit of 25 paths was selected. This captured the 25 strongest multipaths and resulted in an effective received path dynamic range of between 26.1633 dBm (strongest path) and -76.1872 dBm (weakest path).

Each indoor room model was constructed to the basic geometric shape of a rectangle or square. The height of all rooms was fixed to 2.4m, with width and breadth ranging from 1m×2m through 15m×15m to simulate small to large rooms respectively. Common building materials such as ceiling and floor boards were selected, along with common drywall for the walls. Material characteristics were chosen to model the environment in the unlicensed 5 GHz band used for high-speed consumer communications e.g. Wi-Fi 802.11ac. The waveform type, frequency and bandwidth, as well as the transmit power were also selected to approximate consumer device specifications operating within this frequency band. Root-Raised Cosine (RRC) shaping was applied to the transmitting wave pulse

so as to provide favourable characteristics, such as limiting out-of-band frequency emissions and simplifying end-to-end communication system filter complexity. A half-wave dipole antenna was included within the model as it is ubiquitous in consumer wireless devices, presenting a practical implementation of an idealised isotropic antenna with a doughnut radiation pattern.

The selected RT algorithm was X3D, which combines both the Shooting and Bouncing Rays (SBR) and image method propagation models presented in Chapter 2, in a hybrid approach for faster computation time (SBR) with high accuracy (image method). The X3D model supports Graphics Processing Unit (GPU) implementation and Central Processing Unit (CPU) multi-threading enabling speed optimised runtimes. Transmissions i.e. signal penetration through a material, were not enabled as all measurements were confined to a single room. The number of reflections and diffractions were set to 4 and 1 respectively. These were selected as a trade off between accuracy and processing time.

Ray spacing refers to the minimum angle between adjacent rays during the transmitter shoot procedure. Its value can be determined by $\Delta\phi \approx R_{collect}/D_{max}$, where $R_{collect}$ is the receiver collection sphere radius and D_{max} is the maximum room distance, in this case 15m [11]. The larger the receiver collecting sphere radius, the greater the probability of receiving a ray path. However, this can lead to the collection of too many inaccurate rays, resulting in a misleading CIR.

4.1.2 CIR Preprocessing

An essential part of deep learning involves both the generation of format consistent data and the accurate creation of associative supervised learning labels. The CIR Preprocessing stage was developed to implement both these tasks, refer Figure 4.2. Consistent ML training data was created by first reading all RT output CIR text files into MATLAB and confirming that 25 multipaths exist for each CIR. If this was not the case, the respective RT model was re-evaluated, otherwise each CIR was converted into a cell format to promote straightforward importation to the Python environment present in a subsequent system model stage.

To generate each ML label, information regarding the respective room size, CIR transmitter location and receiver location were required. Room dimensions were extracted from the folder location string, transmitter locations from the CIR text file name string, and receiver locations from the information contained within the RT output file. The room dimensions, transmitter location and receiver location were then wrapped in eXtensible Markup Language (XML) formatting and stored along with the reformatted CIR as a file database. Refer Appendix B for an example of our implemented machine learning label including additional unused fields.

Before we examine both the ML preprocessing algorithms and neural network implementations presented in this thesis, we first ask the question; how well can a neural network estimate room dimension given a single clean and undistorted receiver CIR measurement?

31

Read RT .txt Files

↓

Check All Paths Present

↓

Extract Room Dimensions

Room Loop

Extract Tx/Rx Locations

CIR Loop

Compress path information into Struct

Generate ML Label (XML)

↓

Write CIR.mat + label.xml
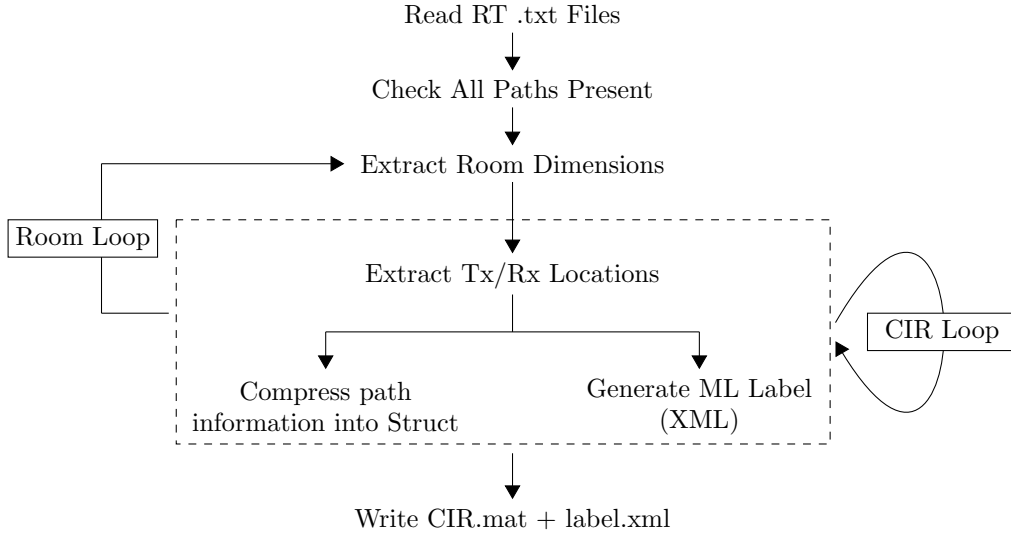
FIGURE 4.2: Procedure for converting Ray-Tracing output for Machine Learning usage

4.2 Room Size Estimation using a single undistorted CIR

In this scenario, we develop a deep FFNN to predict a room's width, breadth and height given a single randomly selected undistorted CIR from the set of CIRs for all transmitter and receiver locations across all room sizes. Due to the randomised nature of the learning problem, we perform the ML preprocessing workflow summarised in Figure 4.3.

Data Import → CIR Removal (Same Tx/Rx Pos.) → Standardisation → Feed-Forward Neural Network (1 CIR)
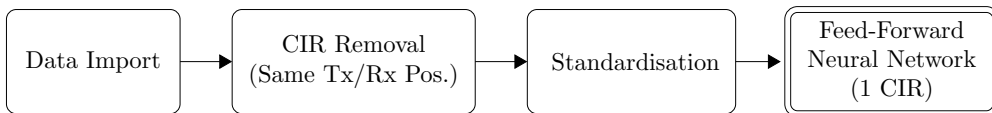
FIGURE 4.3: Preprocessing Workflow for a Feed-Forward Neural Network with a Single CIR

After data and label importation, it is critical that each learning feature be scaled to a similar range to aid neural network learning. A common scaling method is standardisation and is the process of rescaling data to have a mean of 0 and a standard deviation of 1. This is described by

$$\hat{\mathbf{x}} = \frac{\mathbf{x} - \boldsymbol{\mu}}{\boldsymbol{\sigma}}, \tag{4.1}$$

where x is a matrix of feature vectors, μ and σ are vectors representing the mean and standard deviation for each feature across the set of features i.e. feature vector. The feature vector representative of each CIR comprised of a 75 element array, where each feature represents either path delay, magnitude or phase for one of the 25 CIR multipaths respectively.

Standardisation may produce erroneous results if strong outliers exist for any feature. To prevent

this, all CIRs with both the transmitter and receiver co-located were first removed from the dataset in a CIR Removal stage. This eliminates strong Line-of-Sight (LoS) path outliers defined by a minimal transmitter and receiver separation distance, thereby increasing standardised feature dynamic range.

A FFNN architecture was selected as a basis for the developed deep learning algorithm due to the following properties:

- The architecture supports supervised learning which is ideal for regression applications,
- Such a neural network can be trained iteratively on isolated labelled data samples,
- FFNNs support deep learning implementations.

The input layer consists of 75 neurons enabling full support of the input feature vector. A deep neural network was then constructed by specifying 3 hidden layers of 500 neurons each, and an output layer of 3 neurons for room width, breadth and height estimation. Furthermore, the Rectified Linear Unit (ReLU) activation function and Mean Squared Error (MSE) cost function were used to support regression estimation. Refer Table C.1 for the complete list of FFNN parameters.

The widely used Adam optimiser was selected over stochastic gradient descent as it is shown to significantly improve neural network convergence rates [68, 69]. The split between training, cross-validation and testing data was allocated to 80%, 10% and 10% respectively. For repeatability, data randomisation at each epoch (entire passes through the training dataset) was pre-defined and all neuron weight matrices initialised using He initialisation with a seed equal to 1 [70].

The ability to accurately predict room dimensions using our FFNN algorithm given a single randomised undistorted CIR is represented by Table 4.1. Each result indicates a separately trained deep FFNN. For any result indicating the number of paths present, each respective neural network has been trained and tested using this number with all remaining paths zero padded to maintain a fixed neural network input dimension. Similarly, any result indicated by: delay; delay and magnitude; or delay, magnitude and phase; represents the respective path components utilised during neural network training and testing and are referred to as the path data types.

In general for all data types, as the number of usable paths increases so too does the predicted room dimension accuracy. Approaching the limit of 25 available paths however, a small reduction of $1 - 2\%$ accuracy is present. This is explained via a small increase in training cost error during the final epoch as the neural network slightly overshoots the global minima. Terminating the neural network training earlier, or reducing the learning rate will prevent this artefact.

With respect to data type utilisation, it is observed that including both delay and magnitude as CIR learning features improves accuracy over using path delay information only. However, including path phase information as a third additional learning feature contrary to expectation, does not improve room dimension estimation. Examination of the normalised phase features with respect to room size

Feed-Forward Neural Network: Single-CIR					
No. Paths	Data Types				
	0.5m Error Threshold (W, B, H)			1m Error Threshold (W, B, H)		
	Del.	Del.+ Mag.	Del. + Mag. + Pha.	Del.	Del. + Mag.	Del. + Mag. + Pha.
1	0	2.7	2.7	0	10.5	10.8
2	3.0	3.2	3.3	11.8	12.4	12.9
3	4.3	5.2	5.3	16.5	17.6	18.3
4	7.0	8.9	8.8	23.7	26.1	26.4
5	10.9	16.0	15.2	34.1	38.8	39.1
6	18.1	27.1	27.1	46.3	54.8	55.1
7	28.2	43.4	41.7	59.8	69.4	68.9
8	37.7	56.5	54.8	70.4	79.8	79.1
9	49.0	67.3 [2]	66.0	80.2	87.7 [2]	87.2
10	52.7 [3]	75.3	72.6	85.0 [3]	92.6	92.1
11	59.6	79.8	78.7	88.4	95.1	94.9
12	64.4	83.5	79.8	91.2	96.6	96.2
13	67.2	85.8 [1]	84.2	93.2	97.7 [1]	97.6
14	69.7	86.7	83.6	94.3	98.2	97.9
15	70.6	85.5	87.0	94.8	98.4	98.5
16	71.2	85.9	84.6	95.2	98.5	98.1
17	73.0	88.9	86.6	96.0	99.0	98.7
18	75.9	86.9	86.9	96.6	98.9	98.7
19	76.6	88.7	88.0	97.0	99.0	98.8
20	76.6	89.2 [1]	86.0	97.0	99.0 [1]	98.6
21	75.1	86.8	87.5	96.5	98.9	99.0
22	76.9	87.8	84.9	97.1	98.9	98.5
23	72.2	86.2	85.8	96.1	98.6	98.6
24	73.2	85.5	86.7	96.2	98.6	98.6
25	74.8	86.2	88.5	96.3	98.6	99.0

[1] Referred to in Section 4.2.
[2] Referred to in Section 4.3.
[3] Referred to in Section 5.1.

TABLE 4.1: Feed-Forward Single CIR Percentage Accuracies

reveals a high degree of randomness. This is likely attributed to the cyclic rotation of the phase as a signal propagates through space according to $\theta = \mathrm{mod}(2\pi f d/c, 2\pi)$, where f represents the wireless

signal frequency, d is the distance travelled and c is the speed of light.

A maximum accuracy of 89.2% was achieved when predicting all room size dimensions within 0.5m for the case utilising 20 paths as well as the delay and magnitude components of each CIR. For the case where the accuracy constraint for room size prediction is relaxed to 1m, a prediction accuracy of 99.0% was achieved. Reducing the number of paths constraint from 25 to 13, a moderate accuracy result can still be reached of 85.8% and 97.7% for an error threshold of 0.5m and 1m respectively when utilising both path delay and magnitude.

With our developed FFNN algorithm establishing a baseline for room dimension estimation using a single CIR, the question must be asked; can a deep neural network improve on its predictive performance given many CIR samples within a room?

4.3 Room Size Estimation using multiple undistorted CIRs

In this scenario, we develop a new deep FFNN algorithm capable of learning from multiple CIRs. It is demonstrated that this alternative algorithm, based only on a small modification to the previous deep FFNN's input layer, increases the accuracy of predicting a room's width, breadth and height, with optimum results achieved when exploiting 2 CIRs.

Taking the previous single CIR neural network implementation, we can extend its input layer to a length of any multiple of a single feature vector. An arbitrary multiple of 5 was selected, therefore supporting up to 5 CIRs for a total feature vector length of 375. Three CIR pairing algorithms were also developed in this thesis based on the transmitter and receiver locations namely: random transmitter and random receiver; fixed transmitter and random receiver; and fixed transmitter with a random walk receiver; as presented in Figure 4.4. A random walk receiver scenario was implemented to simulate a plausible information gathering scenario e.g. person walking in a room with a mobile wireless device.
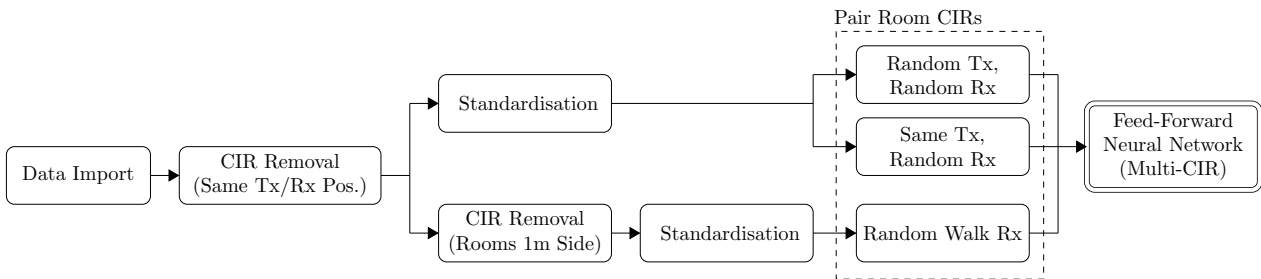
FIGURE 4.4: Preprocessing Workflow for a Feed-Forward Neural Network with Multi-CIR Training

The algorithms for random transmitter and random receiver, plus fixed transmitter and random
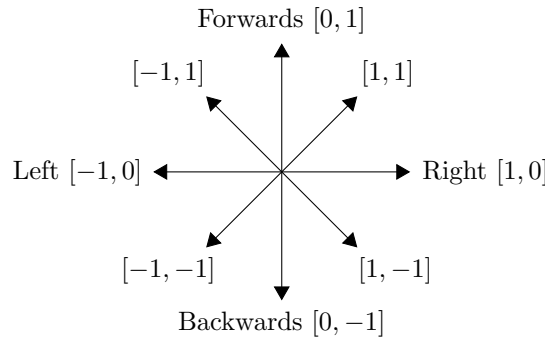
FIGURE 4.5: Random Walk Next Position Selection

receiver pairing are summarised by Algorithm 1 in Appendix E. Here, selected pairing is performed referencing a unique array list consisting of a concatenation between the transmitter location, receiver location and room dimension, each extracted from its respective machine learning label. Then based on either unique room dimension indices only, or transmitter location and room dimension indices, a new data matrix is formed of appropriate size with multiple CIRs paired per single data sample.

To summarise the algorithm for the random receiver walk, we can refer to Figure 4.5. Each successive position in the random walk is selected from a random choice of the available directions. Three caveats exist however to warrant non-selection:

1. The next receiver position cannot be the transmitter location,
2. The random walk cannot double back on itself i.e. the new position cannot be the previous position,
3. The random walk must remain within the boundaries of the current room size.

Due to limitation number 2, any room that presents a width or breadth of 1m must be removed in a second staged CIR removal function. This is to prevent the random walk algorithm from entering into a case whereby there are no valid next positions.

Results for three deep FFNN algorithms supporting multiple CIRs are as follows when combining: CIRs from a common room given random transmitter and random receiver positions, refer Table 4.2; fixed transmitter and random receiver positions, refer Table 4.3; and a fixed transmitter with the receiver undergoing a random walk, refer Table 4.4.

For the case of random transmitter and receiver locations, pairing of 2 CIRs improves prediction accuracy performance. In particular, for the CIR configurations comprising of 9 and 13 multipaths respectively, room dimension prediction accuracy increased significantly from 67.3% to 83.7%, and 85.8% to 93.4% respectively for an error threshold of 0.5m. Relaxing the error threshold to 1m results in an increase from 87.7% to 95.5%, and 97.7% to 99.0% for 9 and 13 multipaths respectively. Similar increases in accuracy of between approximately 5% to 10% occur for the cases where 17 to

25 multipaths are present and an error threshold of 0.5m.

Feed-Forward Neural Network Multi-CIR Pairing: Random Tx, Random Rx								
No. Paths	No. CIRs							
	0.5m Error Threshold (W, B, H)				1m Error Threshold (W, B, H)			
	2	3	4	5	2	3	4	5
1	4.0	4.1	4.0	4.0	15.2	15.4	15.5	15.6
5	22.5	26.9	31.3	33.4	53.6	61.4	67.0	70.1
9	83.7 [1]	81.0	82.2	81.1	95.5 [1]	95.5	96.5	96.3
13	93.4 [1]	91.1	88.3	86.7	99.0 [1]	98.7	98.6	98.4
17	94.5	91.2	88.6	85.8	99.3	99.0	98.8	98.4
21	95.1	91.6	90.1	87.9	99.5	99.1	99.0	98.8
25	95.1 [1]	91.8	89.9	86.4	99.5	99.2	99.0	98.7

[1] Referred to in Section 4.3.

TABLE 4.2: Feed-Forward Neural Network Multi-CIR Percentage Accuracies: Random Tx and Random Rx Pairing using CIR Delays and Magnitudes

These increases in prediction accuracy indicate that including multiple CIRs from many receiver positions adds diversity to the neural network feature vector, therefore increasing the ability of such networks to estimate complex regression patterns. However a limitation appears to exist for FFNNs regarding how much benefit multiple CIRs present. Expanding the number of paired CIRs to 3, a reduction in overall successful room dimension estimation is evident. This accuracy reduction trend continues when pairing 4 or 5 CIRs respectively. Intuitively, this result is unsurprising knowing that FFNNs are unable to share neuron features learnt across different layer positions and therefore are unable to extract relationships between paired CIRs. Contrary to this statement however, where the number of available paths are limited i.e. 1 and 5, increased accuracy is present for all successive CIR number pairing though total predictive accuracy remains low.

For the fixed transmitter and random receiver case, the same general trends are observed; namely, as the number of available CIR paths increases, so too does the accuracy. In addition, pairing multiple CIRs significantly improves accuracy initially, then gradually reduces thereafter. Furthermore, maintaining a fixed transmitter position results in similar accuracies to that of random transmitter and random receiver CIR pairing.

When changing the selected receiver positions to a random walk pattern, the overall room dimension prediction accuracy is seen to reduce compared to both (i) the random transmitter and random receiver case, and (ii) the fixed transmitter and random receiver case. More specifically where the number of available CIR paths is 25 and an error threshold of 0.5m is selected, accuracy falls from (i) 95.1% and (ii) 95.2%, to 72.0%.

Feed-Forward Neural Network Multi-CIR Pairing: Fixed Tx, Random Rx							
No. Paths	No. CIRs						
	0.5m Error Threshold (W, B, H)				1m Error Threshold (W, B, H)		
	2	3	4	5	2	3	4	5
1	3.5	3.5	3.6	3.6	13.6	13.9	14.0	14.0
5	24.3	28.6	31.7	35.0	55.9	63.1	67.8	71.2
9	83.8	82.4	83.9	83.0	95.8	96.4	96.9	97.1
13	92.4	89.9	88.5	87.2	99.0	98.7	98.7	98.7
17	94.4	91.9	90.0	88.0	99.3	99.2	99.1	98.9
21	95.0	91.4	89.3	89.1	99.5	99.2	98.9	99.1
25	95.2 [1]	92.4	90.1	85.8	99.5	99.5	99.1	98.6

[1] Referred to in Section 4.3.

TABLE 4.3: Feed-Forward Neural Network Multi-CIR Percentage Accuracies: Fixed Tx and Random Rx using CIR Delays and Magnitudes

Feed-Forward Neural Network Multi-CIR Pairing: Fixed Tx, Random Rx Walk							
No. Paths	No. CIRs						
	0.5m Error Threshold (W, B, H)				1m Error Threshold (W, B, H)		
	2	3	4	5	2	3	4	5
1	2.6	2.7	2.6	2.7	10.2	10.4	10.4	10.6
5	16.2	17.1	17.0	18.3	42.4	44.9	45.9	48.4
9	58.4	57.8	56.8	55.2	85.9	86.3	86.7	86.1
13	69.2	68.5	65.0	59.6	91.9	92.3	90.9	89.5
17	71.8	70.3	64.5	66.0	93.3	92.8	91.4	91.9
21	71.8	70.1	66.8	65.7	93.1	93.0	92.0	92.1
25	72.0 [1]	70.7	67.7	65.1	93.6	93.1	92.9	91.9

[1] Referred to in Section 4.3.

TABLE 4.4: Feed-Forward Neural Network Multi-CIR Percentage Accuracies: Fixed Tx and Random Rx Walk using CIR Delays and Magnitudes

A random walk is designed to maximise paired CIR relationships by selecting consecutive spatially adjacent receiver positions. Recalling previously that FFNNs are unable to extract such relationships, this forms one possible explanation for the marked decrease in accuracy. In addition, with the constraint that each new receiver position must be adjacent to the previous, a reduction in receiver spatial diversity will be present as CIRs can no longer be sampled randomly from all possible room locations. Such spatial reduction will be further compounded as the room size increases.

Chapter **5**

Room Dimension Estimation using Recurrent Neural Networks

This chapter extends that of Chapter 4, presenting the design methodology and results of developed deep Recurrent Neural Network (RNN) algorithms for the simultaneous estimation of a room's width, breadth and height using multiple Channel Impulse Responses (CIRs). Furthermore, the algorithms presented in this chapter are shown to be capable of estimating room dimensions to a high degree of accuracy, approximately 90% or above, when exploiting multiple CIRs limited by bandwidth or multipath strength to that of practical systems such as Wi-Fi 802.11ac.

To better exploit the temporal and spatial CIR relationships that will exist during a Receiver (Rx) random walk, an alternative more appropriate neural network implementation to the Feed-Forward Neural Network (FFNN) is the RNN. As previously stated in Chapter 2, the RNN is able to maintain knowledge of long-term sequenced data through the addition of the gated memory cell thereby sharing features learnt across different positions in the data.

5.1   Room Size Estimation using multiple undistorted CIRs

Our baseline RNN algorithm using undistorted CIRs is shown via the upper most path in Figure 5.1. Machine Learning (ML) preprocessing remains the same as our previous multi-CIR FFNN algorithm utilising a random walk receiver described in Chapter 4.

Implementation of the developed RNN is summarised by Table D.1. To permit sequenced CIR feature vectors, the input layer size was reduced back to length 75. Therefore, the training data structure was changed from two dimensions to three, where dimension one represents the total number of sequenced training examples, dimension two represents the maximum supported sequence length and dimension three is the feature vector length. This is known as a batch-major arrangement, with each batch consisting of a random selection of 32 training examples from the total training dataset.
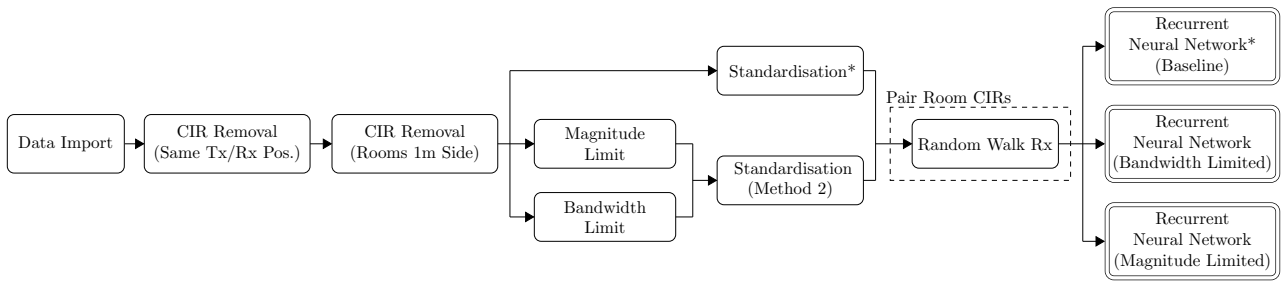
FIGURE 5.1: Preprocessing Workflow for Recurrent Neural Networks with Multi-CIR Training

The number of hidden layers selected for our RNN was 4, each containing a Long Short-Term Memory (LSTM) cell with 512 neurons which being a power of 2, optimises calculation performance. This increase in hidden layer depth in addition to cell neurons, increases the risk of data overfitting i.e. neural network prediction of training examples significantly outperforms the testing dataset. To minimise this, a regularisation technique known as dropout was included with a neuron randomised dropout rate of 20% [69]. All neurons included the tanh activation function and were initialised using Xavier initialisation with a seed equal to 1 [71]. Furthermore, due to the increased processing cost incurred by the RNN time-sequenced structure, a Graphics Processing Unit (GPU) implementation was employed. This utilises the NVIDIA CUDA Deep Neural Network Library (cuDNN) framework optimised for RNN LSTM performance.

Tables 5.1 and F.1 display RNN accuracies for error thresholds 0.5m and 1m respectively where only CIR delays have been utilised. The number of paired CIRs i.e. random walk length, are selected from the set $N_{CIRs} \in \{4, 8, 12, 16, 20\}$, with the number of paths available for neural network training selected from $N_{paths} \in \{5, 10, 15, 20, 25\}$.

Recurrent Neural Network Multi-CIR Pairing: Delays Only, 0.5m Error Threshold (W, B, H)					
No. Paths			No. CIRs		
	4	8	12	16	20
5	19.8	31.6	43.0	49.6	56.1
10	83.3 [1]	93.3	96.0 [1]	97.1	97.4 [1]
15	93.4	97.0	97.0	97.8	98.4 [1]
20	92.7	97.0	95.0	98.3	97.7
25	94.3 [1]	96.4	97.1	97.6	96.4

[1] Referred to in Section 5.1.

TABLE 5.1: RNN Multi-CIR Percentage Accuracies: CIR Delays Only, 0.5m Threshold

Firstly it is observed that room size prediction accuracy in general, increases for all random walk lengths in contrast to the multi-CIR FFNN implementation in Chapter 4. Furthermore, comparing

RNN delay only results with our FFNN algorithm utilising a single CIR in Chapter 4 (Table 4.1), selecting 10 CIR paths produced an FFNN accuracy of 52.7% when the prediction error is limited to 0.5m. With an RNN architecture, this accuracy can be significantly increased to 83.3% or 97.4% when exploiting a minimum of 4 or a maximum of 20 CIRs respectively. Relaxing the prediction error to 1m, only 4 CIRs are required to achieve 96.9% accuracy for the same use case, up from 85.0385% for a single CIR.

The maximum accuracy achieved for room dimension estimation using CIR path delay information only is 98.4% and occurs when exploiting 15 paths with a random walk length of 20. A small reduction in accuracy is evident when combining the maximum walk length, with the maximum available paths and is once again the result of a small increase in neural network training error as the network fluctuates around the global minima. Reducing the RNN learning rate will improve this.

Including CIR delays and magnitudes as RNN features results in Tables 5.2 and F.2 for error thresholds 0.5m and 1m respectively. Following on from the previous delay only case, RNN accuracy also increases with respect to the number CIRs i.e. receiver positions, CIRs used in the random walk. As we saw in the single CIR FFNN case, adding the magnitude components significantly improved prediction accuracy. However, with the RNN implementation this is only the case where both the number of paths present is small e.g. 10 or less, and where the number of paired CIRs is also small e.g. 4. Taking the number of CIR paths as 10 and the number of CIRs as 4, accuracy with an error threshold at 0.5m increases from 83.3% to 93.4% when shifting from delay information only, to delay and magnitude. Increasing the number of usable paths to 25 sees a prediction accuracy increase from 94.3% to 95.7% for the same respective data type utilisation. Likewise increasing the usable CIRs to 12 with 10 paths sees prediction accuracy increase from 96.0% to 97.4%.

Recurrent Neural Network Multi-CIR Pairing: Delays and Magnitudes, 0.5m Error Threshold (W, B, H)					
No. Paths	No. CIRs				
	4	8	12	16	20
5	29.9	42.0	57.7	54.9	64.2
10	93.4 [1]	96.8	97.4[1]	97.5	97.5
15	96.4	98.1	98.3 [1]	98.0	96.6
20	96.2	97.2	97.1	96.9	96.6
25	95.7 [1]	97.0	97.0	96.2	95.5

[1] Referred to in Section 5.1.

TABLE 5.2: RNN Multi-CIR Percentage Accuracies: CIR Delays and Magnitudes, 0.5m Threshold

Maximum accuracies of 98.3% and 99.8% are achieved when exploiting both delay and magnitude CIR information for the configuration utilising 12 CIRs with; 15 paths and an error threshold of 0.5m,

and 20 paths for an error threshold of 1m, respectively. This is similar to the delay only case of 98.3827% and 99.7856% with 15 paths and 20 CIRs. Once again, accuracies for increased path numbers and CIRs can be improved through a reduction in neural network learning rate.

Adding the phase CIR component to include path delays, magnitudes and phases produces Tables 5.3 and F.3 for error thresholds 0.5m and 1m respectively. Similar to the FFNN arrangement, addition of the phase component does not provide any advantage to room dimension predictive accuracy with both delay and magnitudes only, and delay, magnitudes and phase cases producing approximately equivalent accuracy results. For example selecting an error threshold of 0.5m, the number of CIR paths as 10 and the number of CIRs as 4, accuracies of 93.4% and 91.9% are achieved for: delay and magnitudes only; and delay, magnitudes and phase cases respectively. Increasing the number of CIR paths to 25, obtained accuracies are 95.7% and 94.0% respectively.

Recurrent Neural Network Multi-CIR Pairing: Delays, Magnitudes and Phases, 0.5m Error Threshold (W, B, H)					
No. Paths	No. CIRs				
	4	8	12	16	20
5	28.4	41.9	51.6	57.4	64.3
10	91.9 [1]	95.7	96.7	97.3	97.2
15	95.3	97.2	97.8	97.4	97.1
20	94.8	96.4	96.9	95.4	96.4
25	94.0 [1]	96.3	96.1	96.8	94.0

[1] Referred to in Section 5.1.

TABLE 5.3: RNN Multi-CIR Percentage Accuracies: CIR Delays, Magnitudes and Phases, 0.5m Threshold

Post examination of all multi-CIR RNN implementations has led to the following generalisation. If either a high number of CIR paths are available or if many receiver locations can be sampled, using CIR delay information only will be sufficient to accurately predict room dimensions. This is useful for either case where it is difficult to extract accurate and meaningful magnitude information from super-resolution algorithms such as MUltiple SIgnal Classification (MUSIC) and Matrix Pencil Method (MPM), or where additional processing requirements needed for magnitude analysis are high.

In practice, the ability for a receiver to detect and extract multiple paths from any given CIR highly depends on several factors; the weakness of the multipaths in a noise limited environment, and the available bandwidth of the transmitted signal (as reducing bandwidth will reduce a receiver's ability to distinguish between adjacent multipaths). Based on this, we now investigate how well neural networks are able to handle such limitations brought upon by physical systems?

5.2   Room Size Estimation using Bandwidth Limited CIRs

An RNN algorithm developed in this thesis demonstrates that room dimension estimation can be achieved given multiple CIRs limited in bandwidth to that of practical systems. The method used to band-limit each CIR was previously described in Section 2.1.1. A bandwidth lower bound was selected at 160 MHz, corresponding to the maximum available by Wi-Fi 802.11ac devices. An upper limit of 25 GHz was selected, as significantly increasing this figure resulted in no loss of multipath.

A bandwidth limited CIR can be calculated as a convolution between the time-domain impulse response and a bandwidth limited sinc filter function, refer Equation (2.10). The frequency-domain function of a sinc filter is rectangular, and therefore a practical drawback lies in that such a filter is unrealisable. Furthermore, sharp cutoff frequency transitions result in extensive time-domain filter ripple. In the case where weak multipaths exist near in time to a single strong path, such large filter ripple will mask these weaker multipaths.

Based on this, the sinc filter was replaced by a windowed Raised Cosine (RC) filter. This filter is practically realisable and possesses reduced time-domain ripple favourable for minimising masking loss of weak multipath. A Blackman window was added to smooth filter time-domain band edges to zero further reducing ripple. The difference between filter time-domain ripple amplitude and decay rate for various filters is illustrated in Appendix J. Furthermore the complete algorithm detailing the bandwidth limitation process is presented by Algorithm 2 in Appendix E. If the total number of new paths detected was less than originally present for the infinite bandwidth case, the remaining CIR paths were padded to zero.

Due to the zero padding of CIR path information, a slight modification to the original standardisation algorithm is required. In this case, the mean and standard deviation of each feature within the feature vector must be determined excluding all zero elements and is referred to as Standardisation (Method 2). The CIR pairing random walk algorithm remains unchanged as described in Chapter 4 and the complete workflow for the new physically limited RNN neural network algorithm is depicted in Figure 5.1.

To begin, each CIR was bandwidth limited according to the bandwidth selections indicated by Table 5.4. At a bandwidth of 160 MHz, when utilising both delay and magnitude information in addition to 20 CIRs, an accuracy of 90.0% is achieved with an error threshold of 0.5m. Increasing the bandwidth to 500 MHz sees accuracies above 90% when exploiting either delays only and 16 CIRs, or delays and magnitudes and 12 CIRs. Above 1 GHz, 90% accuracy can be achieved using a random walk length of only 4.

To place these results in context with respect to previously obtained undistorted infinite bandwidth

Bandwidth Limited CIRs: 0.5m Error Threshold (W, B, H)						
Bandwidth	Data Types	No. CIRs				
		4	8	12	16	20
160 MHz	Del.	27.2	43.7	57.2	68.7	77.1
	Del. + Mag.	43.2	67.7	79.6	84.3	90.0 [1]
	Del. + Mag. + Pha.	44.3	65.6	79.3	85.2	89.8
500 MHz	Del.	56.9	76.7	87.9	90.6	93.4
	Del. + Mag.	68.0	84.4	90.5	93.5	92.4
	Del. + Mag. + Pha.	65.5	84.3	89.6	92.0	91.2
1 GHz	Del.	71.0	88.0	91.7	95.6	95.7
	Del. + Mag.	80.2	89.6	94.1	94.8	94.6
	Del. + Mag. + Pha.	75.7	87.5	90.1	92.8	93.7
5 GHz	Del.	89.4	94.6	96.0	96.9	97.0
	Del. + Mag.	90.7	94.8	96.3	95.7	95.8
	Del. + Mag. + Pha.	88.1	92.5	94.6	93.9	93.7
10 GHz	Del.	90.5	95.9	97.1	96.2	96.3
	Del. + Mag.	92.3	95.3	93.3	96.2	95.8
	Del. + Mag. + Pha.	89.3	92.4	94.3	94.8	94.7
25 GHz	Del.	91.3	95.4	96.6	94.2	97.2 [1]
	Del. + Mag.	93.3	95.4	93.1	96.0	95.7
	Del. + Mag. + Pha.	90.8	93.8	94.9	94.4	93.6

[1] Referred to in Section 5.2.

TABLE 5.4: RNN Bandwidth Limited Multi-CIR Percentage Accuracies: 0.5m Threshold

based CIR accuracies, we can refer to Figure J.1. Presented are histograms detailing the number of CIR paths available given the respective bandwidth limitations for all CIRs. At 160 MHz, the mean available paths is approximately 5. For the remaining bandwidths of 500 MHz, 1 GHz, 5 GHz, 10 GHz and 25 GHz the mean paths available are approximately 8, 12, 21, 22 and 24 respectively. From this it is clear that overall accuracies for the bandwidth limited case, 160 MHz excluded, are marginally less than via simple selection of $N \in \{5, 10, 15, 20, 25\}$ undistorted CIR paths. Possible explanation for this is twofold. Firstly, as a CIR becomes increasingly band-limited, individual paths become merged resulting in a new path subset, each with new delays, amplitudes and phases. Therefore no longer are we comparing equivalent CIRs. Secondly, the number of CIRs available under the bandwidth limited case is represented by a normal distribution with known mean and standard deviation. Therefore, each successive CIR may present a different number of paths within the distribution, resulting in complexity that will affect neural network training.

A maximum overall accuracy of 97.2% was achieved at 25 GHz bandwidth using delay path information only and 12 CIRs. Increasing either the bandwidth or number of CIRs sees a small reduction in accuracy and is once more attributed to a small increase in training cost error for the final training epochs. Fine-tuning of the neural network learning rate or total epochs trained will help improve these results.

Relaxing the room dimension error threshold to 1m, upwards of 98% accuracy can be achieved for all bandwidth cases with varying random walk lengths, refer Table G.1. An accuracy of 98.2% is achieved at 160 MHz with 20 CIRs using delay and magnitude path information. This is reduced to 12 CIRs at 500 MHz for 98.2% accuracy, 8 CIRs at 1 GHz for 98.7% accuracy or 4 CIRs at 5 GHz for 99.1% accuracy.

In conclusion when utilising bandwidth limited CIRs, increasing bandwidth in addition to the number of CIRs improves RNN performance. Moreover, the same accuracy generalisation with respect to utilisation of only the CIR delay component and multipath availability can be reached. In that for most bandwidth and random walk lengths, sufficient accuracy of greater than 90% can be achieved when exploiting only the CIR delay component.

5.3   Room Size Estimation using Magnitude Limited CIRs

Another RNN algorithm developed in this thesis demonstrates that room dimension estimation can also be achieved when exploiting CIRs limited by multipath signal strength i.e. magnitude, whilst maintaining infinite bandwidth. This represents a physical noise limited environment, reducing the ability of a receiver to extract weak multipaths from noise.

Tables 5.5 and G.2 detail RNN accuracies where each CIR has been filtered to only include received multipaths equal to, or greater than a magnitude threshold $T_{mag} \in \{-40, -45, -50, -55, -60, -65, -70\}$ dBm. Exceeding this range would result in either all or extremely few multipaths present for neural network training. Paths weaker than the specified threshold value were eliminated by zero padding. It is important to note that a fixed transmitter power of 30 dBm was maintained during the Ray-Tracing (RT) process. This ensures consistency of multipath threshold selection for all CIRs across all room models.

Similarly for the bandwidth restricted case, histograms detailing the number of available paths post magnitude thresholding are presented in Figure J.2. Observation of each histogram reveals that restricting CIR multipaths to -40 dBm results in very few available i.e. 3 or less, for neural network training. This is typified by room dimension prediction accuracies, less than 6% and 20% for error thresholds of 0.5m and 1m respectively. Reducing the multipath limit to -45 dBm results in a small

and insignificant increase to accuracy, once again limited by the number of available multipaths for training with a histogram mean of approximately 3.

Magnitude Limited CIRs: 0.5m Error Threshold (W, B, H)						
Magnitude Threshold	Data Types	No. CIRs				
		4	8	12	16	20
−40 dBm	Del.	4.2	4.6	5.0	5.3	5.5
	Del. + Mag.	4.4	5.0	5.4	5.7	5.8
	Del. + Mag. + Pha.	4.7	5.0	5.5	5.8	6.1
−45 dBm	Del.	12.5	17.3	21.3	24.3	28.4
	Del. + Mag.	14.8	19.3	23.1	27.1	30.2
	Del. + Mag. + Pha.	14.2	18.8	24.1	27.9	30.8 [1]
−50 dBm	Del.	76.6	88.5	93.1	94.7	96.0
	Del. + Mag.	87.6	92.8	95.0	95.9	96.2 [1]
	Del. + Mag. + Pha.	86.5	91.8	94.1	95.6	95.9
−55 dBm	Del.	94.1	97.0	97.0	97.8	97.7
	Del. + Mag.	96.2 [1]	97.4 [1]	97.8	97.3	96.6
	Del. + Mag. + Pha.	94.2	96.4	97.1	96.5	95.9
−60 dBm	Del.	94.2	95.2	97.6	97.6	96.8
	Del. + Mag.	95.6	95.4	97.2	96.7	95.9
	Del. + Mag. + Pha.	94.2	96.1	96.6	95.5	95.5
−65 dBm	Del.	93.4 [1]	96.1	97.3	97.1	96.6
	Del. + Mag.	95.8	97.2	95.9	96.9	97.0
	Del. + Mag. + Pha.	94.5	96.0	95.9	93.9	95.7
−70 dBm	Del.	94.2	96.2	97.0	97.2	95.9
	Del. + Mag.	94.8	97.0	97.2	96.7	95.3
	Del. + Mag. + Pha.	94.3	95.7	96.5	94.9	94.5

[1] Referred to in Section 5.3.

TABLE 5.5: RNN Magnitude Limited Multi-CIR Percentage Accuracies: 0.5m Threshold

At −50 dBm, a marked increase in accuracy is evident, from a maximum of 30.8% in the previous case of −45 dBm, to 96.2% utilising 20 CIRs and both delay and magnitude data types at an error threshold of 0.5m. Examination of the respective histogram identifies approximately 8 to 9 paths on average are available for training with significant variation due to a large standard deviation. Relaxing the dimension error to 1m produces an accuracy of 99.4%, refer Table G.2.

Limiting the multipath strength to −55 dBm or above once again witnesses a stark increase in

room dimension successful prediction rates. With an average of approximately 18 CIR paths available distributed between an approximate range of 14 and 24, 96.2% accuracy can be achieved at 0.5m error with as little as 4 sampled CIRs. Sampling 8 CIRs increases the obtainable accuracy to 97.4%. With a 1m dimension error threshold, any CIR configuration with the -55 dBm multipath strength limitation produces accuracies greater than 99%.

Chapter **6**

Conclusion

Traditionally, the use of wireless Channel State Information (CSI) has been to correct distortions introduced by signal propagation in a complex environment. These corrections are designed to improve communication Bit Error Rate (BER) and/or Signal to Noise Ratio (SNR), through equalisation and beamforming techniques. Resent research has investigated the use of such CSI representations e.g. Channel Frequency Response (CFR) and Channel Impulse Response (CIR), for applications pertaining to situational awareness in particular, localisation, Simultaneous Localisation and Mapping (SLAM) and object characterisation. Furthermore, the range of Artificial Intelligence (AI) applications has exploded for its ability to draw inference from big data where traditional statistic techniques fail. Despite both recent research in situational awareness and AI, very little has been published regarding the combined use of wireless CSI and deep learning for environment mapping.

This thesis makes several contributions to the field of wireless environment mapping by investigating how deep learning can be used to draw inference from wireless channel measurements, in particular using a CIR to estimate indoor room dimensions, namely a room's width, breadth and height. A Feed-Forward Neural Network (FFNN) algorithm was developed in Chapter 4 that predicts empty room dimensions to 89% and 99% given a single CIR with 25 multipaths unaffected by bandwidth or noise limitation, and estimation dimension error margins of 0.5m and 1m respectively. A key finding concluded that on average, utilisation of the CIR path phase component in addition to both path delay and magnitude, provides no improvement to overall neural network algorithm performance.

A second FFNN algorithm was developed in Chapter 4, resulting in improved room dimension estimation when pairing multiple single room CIR measurements. Accuracies of 95% and 99% were obtained for error margins of 0.5m and 1m respectively when utilising 2 clean CIRs. However, further investigation showed pairing of greater than 2 CIRs resulted in reduced accuracy proportional to the number of additional CIRs paired.

Three different methods for CIR pairing were also developed in Chapter 4. These showed that for FFNNs, pairing according to both random transmitter and random receiver locations, in addition to fixed transmitter and random receiver locations performed better than pairing according to a fixed

transmitter and random walk receiver. This behaviour was attributed to a FFNN's inability to learn from temporal and spatial relationships which a random walk inherits.

Furthermore, in physical systems limited by both bandwidth and SNR, the number of paths available for neural network training will become reduced. Our FFNN algorithms showed that significantly reducing the number of available CIR paths results in a marked decrease in neural network performance. To successfully combine multiple CIR measurements and therefore counteract this effect, several Recurrent Neural Network (RNN) algorithms were created in Chapter 5.

The first developed RNN algorithm improved room dimension estimation over our previous FFNN algorithms given clean, undistorted CIRs. Accuracies greater than 96% were produced when exploiting 8 CIRs with 10 CIR paths, or 4 CIRs and 15 paths with delay and magnitude information and an error margin of 0.5m. Maximum performance corresponded to accuracies of 98% and 99% when utilising both CIR delay and magnitude components for error margins 0.5m and 1m. A fixed transmitter and random walk receiver CIR pairing process was utilised.

An important conclusion using the RNN algorithm was reached given various combinations of receiver walk lengths and CIR paths exploited. This stated that if either a high number of CIR paths are available or if many receiver locations can be sampled, using CIR delay information only will be sufficient to accurately predict room dimensions. This is useful where magnitude information cannot be extracted from super-resolution algorithms such as MUltiple SIgnal Classification (MUSIC) and Matrix Pencil Method (MPM), or where processing requirements for magnitude analysis is expensive.

A second RNN algorithm demonstrated that CIRs band-limited to that of practical systems can be exploited to achieve room dimension estimation. A range of bandwidths were considered, from 160 MHz, which is the maximum supported by Wi-Fi 802.11ac, to 25 GHz used forUltra-Wideband (UWB) systems. Accuracies of approximately 90% and greater were obtained utilising various receiver walk lengths and CIR data type combinations with a 0.5m dimension error margin. Accuracies of 98% may be achieved when relaxing the error margin to 1m. Relating bandwidth to CIR path availability using a developed bandwidth limiting algorithm, showed path availability becomes represented by a normal distribution, with its mean and standard deviation dependent on the bandwidth selected.

A third developed RNN algorithm estimated room dimensions using CIRs restricted according to path magnitude. This demonstrated that once again, path availability becomes represented by a normal distribution with mean and standard deviation dependent on the magnitude limit selected. Results further indicate that a high degree of room dimension estimation accuracy can be achieved, 96% and above, when all CIRs paths are limited to a magnitude of -50 dBm. Setting a magnitude limit greater than this, significantly reduced the number of multipaths available for RNN training and indicates that for practically obtained CIRs, a similar scenario likely exists whereby a significant proportion of multipaths arrive with similar strength. Therefore, any wireless receiver used to perform room

dimension estimation should be capable of extracting multipath at suitably low magnitude levels.

6.1 Future Work

This thesis creates potential for further investigation into several aspects of environment mapping using wireless CIRs. Perhaps the first logical step would be to utilise the existing tools and techniques as developed in this thesis, applied to CIRs captured by real, physical systems.

Such an investigation would involve an extensive measurement campaign requiring several hundred thousand measurements of a vast number of room sizes and shapes. Limitation further exists by requiring systems capable of performing wideband or UWB bandwidth coverage. Several current consumer technologies exist which are capable of supporting relatively wide bandwidths e.g. 802.11ac and 802.11ax at 160 MHz, and should make such a measurement campaign realisable.

Noting advancements in existing consumer wireless standards, the use of Multiple-In Multiple-Out (MIMO) systems is further becoming ubiquitous. MIMO exploits spatial diversity to increase total system throughput through spatial coding or beamforming techniques. It does this by adding multiple antennas to both the transmitter and receiver. Therefore, for every combination of transmitter and receiver antenna there exists a new and separate channel with multipath. Taking Wi-Fi as an example, CSI may be obtained in a similar way described in this thesis i.e. directly through a received signals CIR or CFR. This is referred to as Implicit CSI. Alternatively, some modern communications systems transmit a quantised version of CSI back to the transmitter when operating in MIMO and is known as Compressed Beamforming Feedback. Use of both CSI forms would provide additional information for machine learning algorithms to increase wireless environment mapping performance.

Lastly, it is clear that not all practically utilised rooms are empty. Further statistical analysis of how multipath changes for a large number of possibly furnished rooms would be required. Such analysis would need to be extensive as under-sampling possible furniture room combinations could result in datasets with different distributions for machine learning training and testing. For a machine learning algorithm to generalise for any possible future combination, differences between these distributions need to be minimised. Furthermore, experimentally obtaining a large enough training dataset may be difficult, if not impossible. One possible technique borrowed from the machine learning space which may help solve this is data augmentation. Knowing approximate statistics on how multipath delays, magnitudes and or phases change in general with respect to room size and furniture occupation, one can generate many new CIRs from a known set, where each path is altered according to these statistics. In reality, the distribution of data generated through data augmentation will not exactly match that gathered experimentally. However, it may dramatically assist training of a practical and generalisable environment mapping neural network system.

Bibliography

[1] J. A. del Peral-Rosado et al. "Survey of Cellular Mobile Radio Localization Methods: From 1G to 5G". In: *IEEE Communications Surveys Tutorials* 20.2 (Secondquarter 2018), pp. 1124–1148. ISSN: 1553-877X. DOI: 10.1109/COMST.2017.2785181.

[2] A. Yassin et al. "Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications". In: *IEEE Communications Surveys Tutorials* 19.2 (Secondquarter 2017), pp. 1327–1346. ISSN: 1553-877X. DOI: 10.1109/COMST.2016.2632427.

[3] C. Laoudias et al. "A Survey of Enabling Technologies for Network Localization, Tracking, and Navigation". In: *IEEE Communications Surveys Tutorials* 20.4 (Fourthquarter 2018), pp. 3607–3644. ISSN: 1553-877X. DOI: 10.1109/COMST.2018.2855063.

[4] A. A. Panchpor, S. Shue, and J. M. Conrad. "A survey of methods for mobile robot localization and mapping in dynamic indoor environments". In: *2018 Conference on Signal Processing And Communication Engineering Systems (SPACES)*. Jan. 2018, pp. 138–144. DOI: 10.1109/SPACES.2018.8316333.

[5] C. BASRI and A. El Khadimi. "Survey on indoor localization system and recent advances of WIFI fingerprinting technique". In: *2016 5th International Conference on Multimedia Computing and Systems (ICMCS)*. Sept. 2016, pp. 253–259. DOI: 10.1109/ICMCS.2016.7905633.

[6] F. Zafari, A. Gkelias, and K. K. Leung. "A Survey of Indoor Localization Systems and Technologies". In: *IEEE Communications Surveys Tutorials* (2019), pp. 1–1. ISSN: 1553-877X. DOI: 10.1109/COMST.2019.2911558.

[7] Y. Hua and T. K. Sarkar. "Matrix pencil method and its performance". In: *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*. Apr. 1988, 2476–2479 vol.4. DOI: 10.1109/ICASSP.1988.197145.

[8] R. Schmidt. "Multiple emitter location and signal parameter estimation". In: *IEEE Transactions on Antennas and Propagation* 34.3 (Mar. 1986), pp. 276–280. ISSN: 1558-2221. DOI: 10.1109/TAP.1986.1143830.

[9]   G. de la Roche, A. Alayón-Glazunov, and B. Allen, eds. *LTE-Advanced and Next Generation Wireless Networks: Channel Modelling and Propagation*. John Wiley and Sons Ltd, 2012. ISBN: 978-1-119-97670-7.

[10] Z. Yun and M. F. Iskander. "Ray Tracing for Radio Propagation Modeling: Principles and Applications". In: *IEEE Access* 3 (2015), pp. 1089–1100. DOI: 10.1109/ACCESS.2015.2453991.

[11] Remcom Inc. *Wireless InSite: Reference Manual*. Version 3.2.0. Feb. 2018. ISBN: 9787777777779.

[12] E. Perahia and R. Stacey. *Next Generation Wireless LANs: 802.11n and 802.11ac*. 2nd ed. Cambridge University Press, 2013. DOI: 10.1017/CBO9781139061407.

[13] M. S. Gast. *802.11ac: A Survival Guide*. O'Reilly Media, Inc., Aug. 2013. ISBN: 9781449343149. URL: https://www.safaribooksonline.com.

[14] H. T. Friis. "A Note on a Simple Transmission Formula". In: *Proceedings of the IRE* 34.5 (May 1946), pp. 254–256. ISSN: 2162-6634. DOI: 10.1109/JRPROC.1946.234568.

[15] D. Tse and P. Viswanath. *Fundamentals of Wireless Communications*. Cambridge University Press, Aug. 2005. ISBN: 9780521845274.

[16] J. B. Schneider. "Understanding the Finite-Difference Time-Domain Method". 2010. URL: www.eecs.wsu.edu/~schneidj/ufdtd.

[17] R. Mittra, ed. *Computational Electromagnetics: Recent Advances and Engineering Applications*. Springer-Verlag New York, 2014. ISBN: 978-1-4614-4381-0.

[18] D. Shi et al. "A GPU Implementation of a Shooting and Bouncing Ray Tracing Method for Radio Wave Propagation". In: *ACES Journal* 32.7 (July 2017).

[19] K. Yee. "Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media". In: *IEEE Transactions on Antennas and Propagation* 14.3 (May 1966), pp. 302–307. ISSN: 1558-2221. DOI: 10.1109/TAP.1966.1138693.

[20] E. Nichols. "An Introduction to the Theory of Optics by Arthur Schuster". In: *apj* 21 (May 1905), p. 382. DOI: 10.1086/141227.

[21] J. B. Keller. "Geometrical Theory of Diffraction". In: *J. Opt. Soc. Am.* 52.2 (Feb. 1962), pp. 116–130. DOI: 10.1364/JOSA.52.000116. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-52-2-116.

[22] J. O. Perrine. "Fermat's Principle of Least Time and Snell's Law of Refraction of Light". In: *School Science and Mathematics* 61.1 (1961), pp. 5–14. DOI: 10.1111/j.1949-8594.1961.tb08497.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1949-8594.1961.tb08497.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1949-8594.1961.tb08497.x.

[23] H. Ling, R. .-. Chou, and S. .-. Lee. "Shooting and bouncing rays: calculating the RCS of an arbitrarily shaped cavity". In: *IEEE Transactions on Antennas and Propagation* 37.2 (Feb. 1989), pp. 194–205. ISSN: 1558-2221. DOI: 10.1109/8.18706.

[24] W. G. Hatcher and W. Yu. "A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends". In: *IEEE Access* 6 (2018), pp. 24411–24432. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2830661.

[25] S. Marsland. *Machine Learning: An Algorithmic Perspective*. Second. CRC Press, 2015. ISBN: 9781466583283.

[26] J. Mueller and L. Massaron. *Machine Learning: for dummies*. John Wiley & Sons, Inc., 2016. ISBN: 9781119245513.

[27] C. J. C. H. Watkins and P. Dayan. "Q-learning". In: *Machine Learning* 8 (May 1992), p. 292. ISSN: 1573-0565. DOI: 10.1007/BF00992698. URL: https://doi.org/10.1007/BF00992698.

[28] S. Russell. "Learning agents for uncertain environments (Extended Abstract)". In: *Proceedings of the Eleventh Annual Conference on Computational Learning Theory*. ACM Press, 1998, pp. 101–103.

[29] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

[30] J. Chung et al. *Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling*. 2014. arXiv: 1412.3555 `[cs.NE]`.

[31] L. N. Kandel and S. Yu. "Indoor Localization Using Commodity Wi-Fi APs: Techniques and Challenges". In: *2019 International Conference on Computing, Networking and Communications (ICNC)*. Feb. 2019, pp. 526–530. DOI: 10.1109/ICCNC.2019.8685501.

[32] H. Zhu et al. "R-TTWD: Robust Device-Free Through-The-Wall Detection of Moving Human With WiFi". In: *IEEE Journal on Selected Areas in Communications* 35.5 (May 2017), pp. 1090–1103. ISSN: 0733-8716. DOI: 10.1109/JSAC.2017.2679578.

[33] D. Zhang, H. Wang, and D. Wu. "Toward Centimeter-Scale Human Activity Sensing with Wi-Fi Signals". In: *Computer* 50.1 (Jan. 2017), pp. 48–57. ISSN: 0018-9162. DOI: 10.1109/MC.2017.7.

[34] Z. Wang et al. "Wi-Fi CSI-Based Behavior Recognition: From Signals and Actions to Activities". In: *IEEE Communications Magazine* 56.5 (May 2018), pp. 109–115. ISSN: 0163-6804. DOI: 10.1109/MCOM.2018.1700144.

[35] L. Chen et al. "Human Behavior Recognition Using Wi-Fi CSI: Challenges and Opportunities". In: *IEEE Communications Magazine* 55.10 (Oct. 2017), pp. 112–117. ISSN: 0163-6804. DOI: 10.1109/MCOM.2017.1700081.

[36] Y. Zou et al. "Wi-Fi Radar: Recognizing Human Behavior with Commodity Wi-Fi". In: *IEEE Communications Magazine* 55.10 (Oct. 2017), pp. 105–111. ISSN: 0163-6804. DOI: 10.1109/MCOM.2017.1700170.

[37] Y. Chapre et al. "CSI-MIMO: Indoor Wi-Fi fingerprinting system". In: *39th Annual IEEE Conference on Local Computer Networks*. Sept. 2014, pp. 202–209. DOI: 10.1109/LCN.2014.6925773.

[38] P. M. Holl and F. Reinhard. "Holography of Wi-fi Radiation". In: *Phys. Rev. Lett.* 118 (18 May 2017), p. 183901. DOI: 10.1103/PhysRevLett.118.183901. URL: https://link.aps.org/doi/10.1103/PhysRevLett.118.183901.

[39] T. Kudo and J. Miura. "Utilizing WiFi signals for improving SLAM and person localization". In: *2017 IEEE/SICE International Symposium on System Integration (SII)*. Dec. 2017, pp. 487–493. DOI: 10.1109/SII.2017.8279260.

[40] D. Humphrey and M. Hedley. "Super-Resolution Time of Arrival for Indoor Localization". In: *2008 IEEE International Conference on Communications*. May 2008, pp. 3286–3290. DOI: 10.1109/ICC.2008.618.

[41] Y. Chen. *Super Resolution*. Lecture Slides. ELE 538B: Sparsity, Structure and Inference. 2017. URL: www.princeton.edu/~yc5/ele538b_sparsity/lectures/super_resolution.pdf.

[42] I. Standard. "IEEE Standard for Information technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements–Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications–Amendment 4: Enhancements for Very High Throughput for Operation in Bands below 6 GHz." In: *IEEE Std 802.11ac(TM)-2013 (Amendment to IEEE Std 802.11-2012, as amended by IEEE Std 802.11ae-2012, IEEE Std 802.11aa-2012, and IEEE Std 802.11ad-2012)* (Dec. 2013), pp. 1–425. DOI: 10.1109/IEEESTD.2013.7797535.

[43] I. Standard. "IEEE Standard for Information technology–Telecommunications and information exchange between systems Local and metropolitan area networks–Specific requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications". In: *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)* (Dec. 2016), pp. 1–3534. DOI: 10.1109/IEEESTD.2016.7786995.

[44] D. Vasisht, S. Kumar, and D. Katabi. "Decimeter-Level Localization with a Single WiFi Access Point". In: *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*. Santa Clara, CA: USENIX Association, 2016, pp. 165–178. ISBN: 978-1-931971-29-4. URL: https://www.usenix.org/conference/nsdi16/technical-sessions/presentation/vasisht.

[45] H. Shi, H. Zhang, and X. Wang. "A TDOA Technique with Super-Resolution Based on the Volume Cross-Correlation Function". In: *IEEE Transactions on Signal Processing* 64.21 (Nov. 2016), pp. 5682–5695. ISSN: 1053-587X. DOI: 10.1109/TSP.2016.2548988.

[46] S. H. Fang et al. "Channel State Reconstruction Using Multilevel Discrete Wavelet Transform for Improved Fingerprinting-Based Indoor Localization". In: *IEEE Sensors Journal* 16.21 (Nov. 2016), pp. 7784–7791. ISSN: 1530-437X. DOI: 10.1109/JSEN.2016.2602840.

[47] P. H. Tseng et al. "Ray-Tracing-Assisted Fingerprinting Based on Channel Impulse Response Measurement for Indoor Positioning". In: *IEEE Transactions on Instrumentation and Measurement* 66.5 (May 2017), pp. 1032–1045. ISSN: 0018-9456. DOI: 10.1109/TIM.2016.2622799.

[48]   D. Humphrey and M. Hedley. "Prior Models for Indoor Super-Resolution Time of Arrival Estimation". In: *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*. Apr. 2009, pp. 1–5. DOI: 10.1109/VETECS.2009.5073817.

[49]   T. Paul and T. Ogunfrunmiri. "Wireless LAN Comes of Age: Understanding the IEEE 802.11n Amendment". In: *IEEE Circuits and Systems Magazine* 8.1 (First 2008), pp. 28–54. ISSN: 1531-636X. DOI: 10.1109/MCAS.2008.915504.

[50]   K. Pahlavan and A. H. Levesque. *Wireless Information Networks (Wiley Series in Telecommunications and Signal Processing)*. John Wiley & Sons, 2005. ISBN: 0471738646.

[51]   H. Durrant-Whyte and T. Bailey. "Simultaneous localization and mapping: part I". In: *IEEE Robotics Automation Magazine* 13.2 (June 2006), pp. 99–110. ISSN: 1070-9932. DOI: 10.1109/MRA.2006.1638022.

[52]   T. Bailey and H. Durrant-Whyte. "Simultaneous localization and mapping (SLAM): part II". In: *IEEE Robotics Automation Magazine* 13.3 (Sept. 2006), pp. 108–117. ISSN: 1070-9932. DOI: 10.1109/MRA.2006.1678144.

[53]   G. Grisetti et al. "A Tutorial on Graph-Based SLAM". In: *IEEE Intelligent Transportation Systems Magazine* 2.4 (winter 2010), pp. 31–43. ISSN: 1939-1390. DOI: 10.1109/MITS.2010.939925.

[54]   M. E. Jefferies and W.-K. Yeap. *Robotics and Cognitive Approaches to Spatial Mapping*. 1st. Springer Publishing Company, Incorporated, 2010. ISBN: 3642094627, 9783642094620.

[55]   E. Leitinger et al. "Simultaneous localization and mapping using multipath channel information". In: *2015 IEEE International Conference on Communication Workshop (ICCW)*. June 2015, pp. 754–760. DOI: 10.1109/ICCW.2015.7247272.

[56]   E. Leitinger et al. "Factor graph based simultaneous localization and mapping using multipath channel information". In: *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*. May 2017, pp. 652–658. DOI: 10.1109/ICCW.2017.7962732.

[57]   C. R. Karanam and Y. Mostofi. "3D Through-Wall Imaging with Unmanned Aerial Vehicles Using WiFi". In: *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. Apr. 2017, pp. 131–142.

[58]   J. Lv et al. "Robust WLAN-Based Indoor Fine-Grained Intrusion Detection". In: *2016 IEEE Global Communications Conference (GLOBECOM)*. Dec. 2016, pp. 1–6. DOI: 10.1109/GLOCOM.2016.7842238.

[59]   M. A. A. Al-qaness et al. "Device-Free Home Intruder Detection and Alarm System Using Wi-Fi Channel State Information". In: *International Journal of Future Computer and Communication* (Aug. 2016), pp. 180–186. ISSN: 2010-3751. DOI: 10.18178/ijfcc.2016.5.4.468.

[60] Y. Zeng, P. H. Pathak, and P. Mohapatra. "WiWho: WiFi-Based Person Identification in Smart Spaces". In: *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. Apr. 2016, pp. 1–12. DOI: 10.1109/IPSN.2016.7460727.

[61] X. Wang et al. "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach". In: *IEEE Transactions on Vehicular Technology* 66.1 (Jan. 2017), pp. 763–776. ISSN: 0018-9545. DOI: 10.1109/TVT.2016.2545523.

[62] X. Wang, L. Gao, and S. Mao. "CSI Phase Fingerprinting for Indoor Localization With a Deep Learning Approach". In: *IEEE Internet of Things Journal* 3.6 (Dec. 2016), pp. 1113–1123. ISSN: 2327-4662. DOI: 10.1109/JIOT.2016.2558659.

[63] R. Zhou et al. "CSI fingerprinting with SVM regression to achieve device-free passive localization". In: *2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. June 2017, pp. 1–9. DOI: 10.1109/WoWMoM.2017.7974313.

[64] S. Yousefi et al. "A Survey on Behavior Recognition Using WiFi Channel State Information". In: *IEEE Communications Magazine* 55.10 (Oct. 2017), pp. 98–104. ISSN: 0163-6804. DOI: 10.1109/MCOM.2017.1700082.

[65] D. Wu et al. "Device-Free WiFi Human Sensing: From Pattern-Based to Model-Based Approaches". In: *IEEE Communications Magazine* 55.10 (Oct. 2017), pp. 91–97. ISSN: 0163-6804. DOI: 10.1109/MCOM.2017.1700143.

[66] M. Zhao et al. "Through-Wall Human Pose Estimation Using Radio Signals". In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2018.

[67] M. Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[68] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. 2014. arXiv: 1412.6980 `[cs.LG]`.

[69] A. Géron. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*. O'Rielly Media, Inc., 2017. ISBN: 9781491962299.

[70] K. He et al. *Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*. 2015. arXiv: 1502.01852 `[cs.CV]`.

[71] X. Glorot and Y. Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: *In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10). Society for Artificial Intelligence and Statistics*. 2010.

Appendix A: Ray-Tracing Parameters

Parameters	Values		
Room Sizes			
Rectangular	1m×2m → 14m×15m		
Square	3m×3m → 15m×15m		
Height	2.4m		
Material	**Permittivity (F/m)**	**Conductivity (S/m)**	**Thickness (m)**
ITU Floorboard 5 GHz	3.66	0.003874	0.03
ITU Ceiling Board 5 GHz	1.5	0.003252	0.0095
ITU Single Layer Drywall	2.8	0.001	0.013
Waveform			
Type [1]	Root-Raised Cosine		
Frequency	5.775 GHz		
Bandwidth [1]	80 MHz		
Antenna			
Type	Half-wave Dipole		
Polarisation	Vertical		
Voltage Standing Wave Ratio (VSWR)	1		
Receive Power Threshold	−250 dBm		
Receiver Properties			
Noise Figure	3 dB		
Collection Sphere Radius	0.25m		
Height from Floor	1m		
Spacing	1m		
Transmitter Properties			
Input Power	30 dBm		
Height from Floor	1m		
Propagation Model			
Algorithm	X3D (Shoot and Bounce + Exact Path Correction)		
Max Paths Computed	25		
Max Reflections	4		
Max Transmissions	0		
Max Diffractions	1		
Ray Spacing	0.9°		
Accelerated Graphics (AGP)	Enabled		
Atmosphere			
Temperature	22.2°C		
Pressure	1013 mbar		
Humidity	50%		
Output			
Channel Impulse Response (CIR)	Phase (Degrees), Delay (s), Magnitude (dBm) per path		
Total CIRs Generated	857, 859		

TABLE A.1: REMCOM Wireless InSite Ray-Tracing Parameters

---

[1]Value selected arbitrarily and had no effect on generated CIR path output.

Appendix B: Machine Learning Label

Listing B.1: Machine Learning Label for a single Receiver Location

```xml
<?xml version="1.0" encoding="utf-8"?>
<labels>
    <rx_label>
        <shape>Rectangle</shape>
        <square_similarity>0.5</square_similarity>
        <room_dimension>
            <x>2</x>
            <y>1</y>
            <z>2.4</z>
        </room_dimension>
        <tx_location>
            <x>0.5</x>
            <y>0.5</y>
            <z>1</z>
        </tx_location>
        <rx_location>
            <x>0.5</x>
            <y>0.5</y>
            <z>1</z>
        </rx_location>
        <N_antennas>1</N_antennas>
        <N_walls>4</N_walls>
        <perp_dist>
            <wall_1>0.5</wall_1>
            <wall_2>0.5</wall_2>
            <wall_3>1.5</wall_3>
            <wall_4>0.5</wall_4>
        </perp_dist>
        <floor_area>2</floor_area>
        <room_volume>4.8</room_volume>
        <furniture_model>0</furniture_model>
    </rx_label>
</labels>
```

Appendix C: Feed-Forward Neural Network Parameters

Parameter	Value
Feed-Forward Neural Network: Single Channel Impulse Response (CIR)	
No. Data Samples	$850,047$
No. Inputs	75 (25 Paths, 3 Data Types - Delay, Magnitude, Phase)
No. Outputs	3 (Room - Width, Breadth, Height)
No. Hidden Layers	3
No. Nodes per Hidden Layer	500
Neuron Activation Function	Rectified Linear Unit (ReLU)
Neuron Initialisation	He (Variance Scaling Initialiser) Seed = 1
Cost Function	Mean Squared Error (MSE)
Learning Schedule	
Decay Function	Exponential
Initial Learning Rate	0.001
Decay Steps	1000
Decay Rate	1/10
Optimiser	Adam
Training, Cross-Validation, Testing Data Split	80/10/10 %
Training Batch Size	32
Training Epochs	20
Feed-Forward Neural Network: Multi-CIR	
No. Data Samples (Random Tx/Rx Pairing)	$850,047$
No. Data Samples (Same Tx/Random Rx Pairing)	$850,047$
No. Data Samples (Random Rx Walk)	$848,945$
No. Inputs	375 (75×5, Maximum 5 CIRs paired)

TABLE C.1: Feed-Forward Neural Network Parameters

Appendix D: Recurrent Neural Network Parameters

Parameter	Value
Recurrent Neural Network: Multi-Channel Impulse Response (CIR)	
No. Data Samples	$848,945$
No. Inputs	75 (25 Paths, 3 Data Types - Delay, Magnitude, Phase)
No. Outputs	3 (Room - Width, Breadth, Height)
Output Neuron Activation Function	Rectified Linear Unit (ReLU)
Output Neuron Initialisation	He (Variance Scaling Initialiser). Seed = 1
No. Hidden Layers	4
Learning Cell Type	Long Short-Term Memory (LSTM)
Cell Activation Function	tanh
Propagation Direction	Forward only
Cell Initialisation	Xavier. Seed = 1
Regularisation	Dropout = 0.2 (20%)
No. Nodes per Cell	512
GPU Acceleration	Nvidia CUDA RNN-LSTM
Cost Function	Mean Squared Error (MSE)
Learning Schedule	
Decay Function	Linear
Initial Learning Rate	0.001
Decay Steps	1000
Decay Rate	1/10
Optimiser	Adam
Training, Cross-Validation, Testing Data Split	80/10/10 %
Training Batch Size	32
Training Epochs	25
Magnitude Restrictions (dBm)	$[-40, -45, -50, -55, -60, -65, -70]$
Bandwidth Restrictions (GHz)	$[0.16, 0.5, 1, 5, 10, 25]$
Bandwidth Filter Function	Blackman Windowed Raised Cosine (RC)

TABLE D.1: Recurrent Neural Network Parameters

Appendix E: Algorithms

Algorithm 1 Pair Room Channel Impulse Responses (CIRs):
Pair Random Transmitter (Tx) and Random Receiver (Rx) + Same Tx and Random Rx Locations

---

Require:

    x, Data matrix containing CIRs [depth, CIR width].

    txrxroom, Label information matrix containing concatenation of Tx location, Rx location, room dimension (3 dimensions each) [depth, 9].

    N_cirs, Maximum number of CIRs to Pair

  1: Initialise memory for new matrix $\mathbf{x_{new}}$ with size [depth($\mathbf{x}$), CIR width $\times N\_cirs$]

  2: Obtain column indices of either room dimensions, OR (Tx locations AND room dimensions) respectively from *txrxroom* (Random Tx and Random Rx or Same Tx and Random Rx Locations)

  3: Find the unique set of labels given the previous pairing indices

  4: **for all** $u \in$ unique set **do**

  5:     Obtain data indices where Labels equal u

  6:     **for all** $idx \in$ indices **do**

  7:         Randomly sample *N_cirs* worth of data indices

  8:         Assign original data in $\mathbf{x}$ indexed by randomly sampled data indices to $\mathbf{x_{new}}$

  9:     **end for**

10: **end for**

Algorithm 2 Bandwidth Limitation

Require:

    x, Data matrix containing CIRs

  1: Extract Delays from $\mathbf{x}$

  2: Extract Magnitudes from $\mathbf{x}$ and convert to absolute value using $\sqrt{10^{((CIR_{mag}-30)/10)}}$

  3: Extract Phases from $\mathbf{x}$ and convert to radians

  4: Generate the bandwidth filter Impulse Response

  5: **for all** $cirs \in \mathbf{x}$ **do**

  6:     Construct upsampled complex CIR

  7:     Convolve CIR with Filter

  8:     Find peaks

  9:     Cull peaks such that only peaks close to original CIR delays are present

10:     Find unique peak subset due to bandwidth restricted CIR paths merging

11:     Sort new CIR Delays, Magnitudes and Phases according to Descending Magnitudes

12:     Convert Magnitudes, Delays and Phases back into dBm, seconds and degrees respectively

13: **end for**

Appendix F: Clean-CIR RNN Results: 1m Threshold

Recurrent Neural Network Multi-CIR Pairing: Delays Only, 1m Error Threshold (W, B, H)					
No. Paths			No. CIRs		
	4	8	12	16	20
5	48.9	61.8	70.7	75.3	79.6
10	96.9 [1]	98.9	99.4	99.5	99.6
15	99.1	99.6	99.7	99.8	99.8
20	99.0	99.7	99.4	99.8	99.7
25	99.4	99.7	99.7	99.8	99.4

[1] Referred to in Section 5.1.

TABLE F.1: RNN Multi-CIR Percentage Accuracies: CIR Delays Only, 1m Threshold

Recurrent Neural Network Multi-CIR Pairing: Delays and Magnitudes, 1m Error Threshold (W, B, H)					
No. Paths			No. CIRs		
	4	8	12	16	20
5	57.3	68.2	75.5	78.3	83.4
10	98.8	99.7	99.6	99.6	99.7
15	99.5	99.8	99.7	99.8	99.4
20	99.7	99.7	99.8 [1]	99.6	99.6
25	99.6	99.7	99.7	99.5	99.5

[1] Referred to in Section 5.1.

TABLE F.2: RNN Multi-CIR Percentage Accuracies: CIR Delays and Magnitudes, 1m Threshold

Recurrent Neural Network Multi-CIR Pairing: Delays, Magnitudes and Phases, 1m Error Threshold (W, B, H)					
No. Paths			No. CIRs		
	4	8	12	16	20
5	57.3	69.1	76.0	79.9	83.9
10	98.9	99.4	99.7	99.5	99.4
15	99.6	99.6	99.7	99.7	99.7
20	99.4	99.5	99.7	99.2	99.4
25	99.4	99.7	99.6	99.7	99.1

TABLE F.3: RNN Multi-CIR Percentage Accuracies: CIR Delays, Magnitudes and Phases, 1m Threshold

Appendix G: Limited-CIR RNN Results: 1m Threshold

Bandwidth Limited CIRs: 1m Error Threshold (W, B, H)						
Bandwidth	Data Types	No. CIRs				
		4	8	12	16	20
160 MHz	Del.	61.2	76.9	86.1	91.3	94.5
	Del. + Mag.	78.6	91.9	95.9	97.1	98.2 [1]
	Del. + Mag. + Pha.	79.8	91.6	95.9	97.1	98.2
500 MHz	Del.	87.7	95.3	98.0	98.6	99.0
	Del. + Mag.	93.0	96.8	98.2 [1]	99.2	99.0
	Del. + Mag. + Pha.	91.9	97.6	98.5	98.5	98.6
1 GHz	Del.	94.5	98.2	99.0	99.5	99.3
	Del. + Mag.	97.0	98.7 [1]	99.3	99.3	99.2
	Del. + Mag. + Pha.	96.0	98.6	98.0	98.3	99.2
5 GHz	Del.	98.5	99.3	99.5	99.4	99.4
	Del. + Mag.	99.1 [1]	99.3	99.7	99.6	99.6
	Del. + Mag. + Pha.	98.5	98.9	99.4	99.3	99.5
10 GHz	Del.	99.1	99.5	99.7	99.6	99.6
	Del. + Mag.	99.1	99.5	99.4	99.7	99.6
	Del. + Mag. + Pha.	99.0	99.2	99.4	99.4	99.2

[1] Referred to in Section 5.2.

TABLE G.1: RNN Bandwidth Limited Multi-CIR Percentage Accuracies, 1m Threshold

Magnitude Limited CIRs: 1m Error Threshold (W, B, H)						
Magnitude Threshold	Data Types	No. CIRs				
		4	8	12	16	20
−40 dBm	Del.	14.7	16.0	17.1	17.7	18.3
	Del. + Mag.	14.9	16.3	17.5	18.4	19.1
	Del. + Mag. + Pha.	15.7	17.0	18.4	19.5	20.3
−45 dBm	Del.	33.9	40.3	45.6	48.8	53.4
	Del. + Mag.	35.9	42.1	46.8	51.3	55.3
	Del. + Mag. + Pha.	36.4	42.9	48.4	52.8	56.5
−50 dBm	Del.	95.1	97.8	98.9	99.2	99.3
	Del. + Mag.	97.5	99.0	98.9	99.4	99.0
	Del. + Mag. + Pha.	97.2	98.8	99.1	99.3	99.4 [1]
−55 dBm	Del.	99.4	99.6	99.6	99.7	99.7
	Del. + Mag.	99.6	99.8	99.8	99.6	99.6
	Del. + Mag. + Pha.	99.3	99.6	99.7	99.6	99.5

[1] Referred to in Section 5.3.

TABLE G.2: RNN Magnitude Limited Multi-CIR Percentage Accuracies: 1m Threshold

Appendix H: Software and Hardware Parameters

Software/Hardware	Version
Software	
REMCOM Wireless InSite	3.2.0.6
MATLAB	R2018a
CUDA	10.0.130
NVIDIA CUDA Deep Neural Network Library (cuDNN)	7.6.0
NVIDIA Display Driver	431.02
Operating System	Windows 10 Pro for Workstations
Hardware	
Computer Type	Dell Precision Tower 7910
Graphics Card	NVIDIA Quadro P6000
Installed RAM	128 GB
Processor	Intel(R) Xeon(R) E5 − 2660 v4 @ 2.00 GHz, 2.00 GHz

TABLE H.1: Computer Software and Hardware Versions
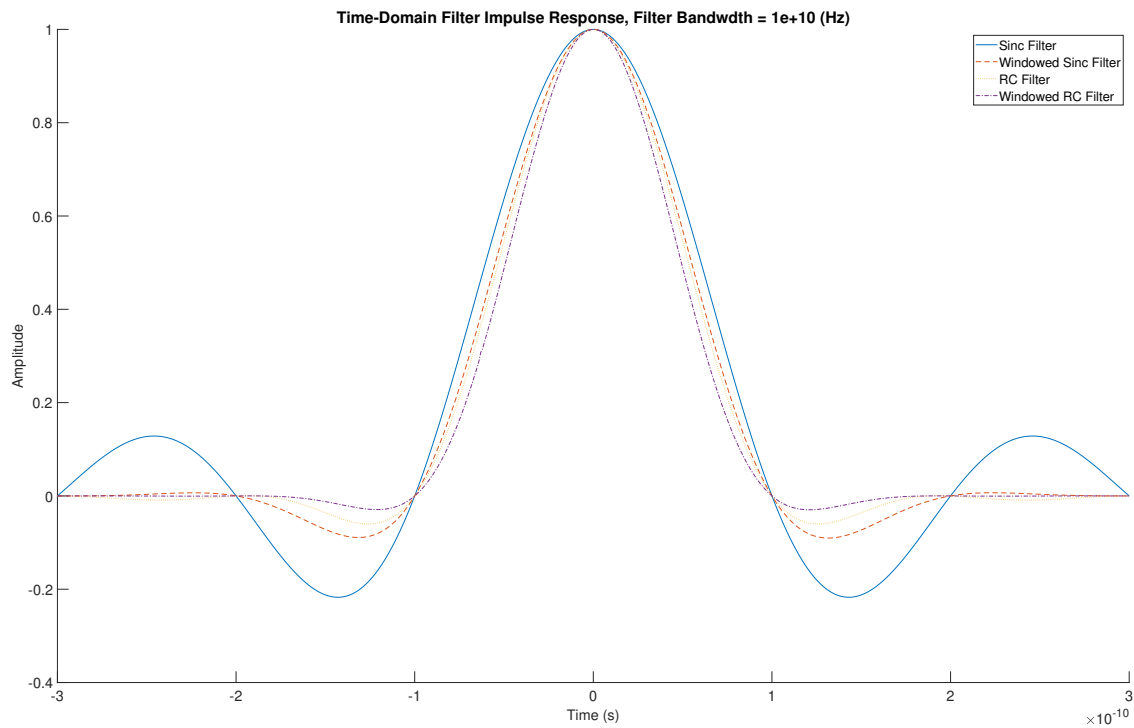
Appendix I: Bandwidth Filter Comparison

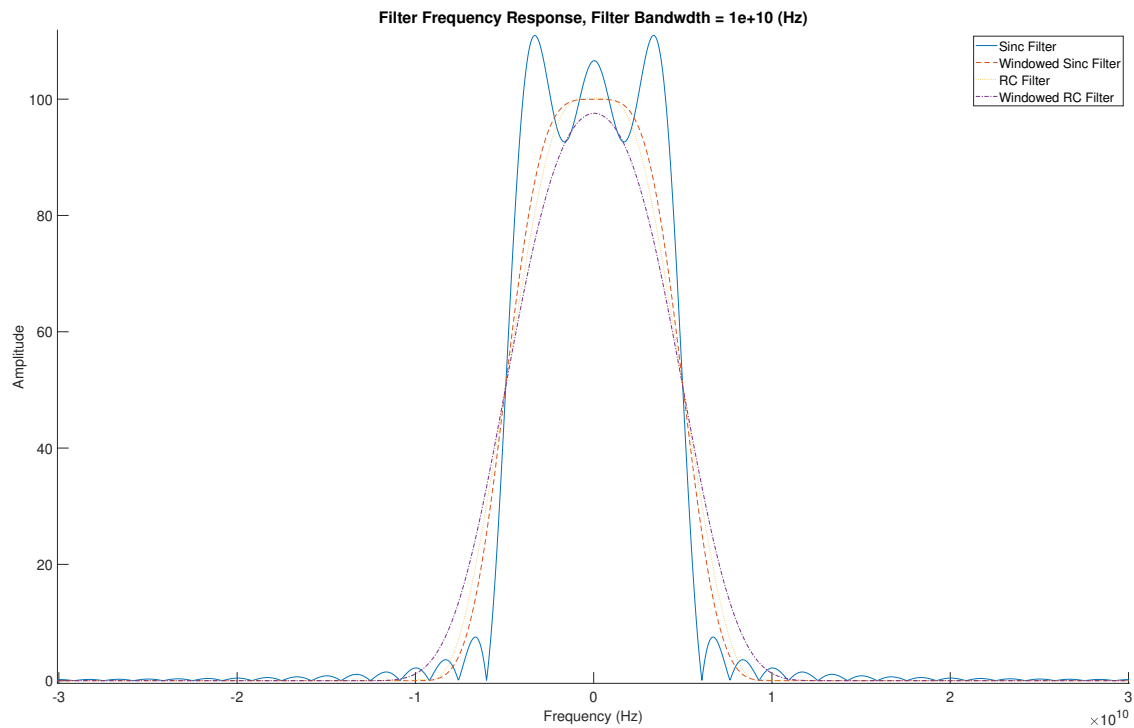FIGURE I.1: Impulse Responses for different Filter Types

FIGURE I.2: Frequency Responses for different Filter Types
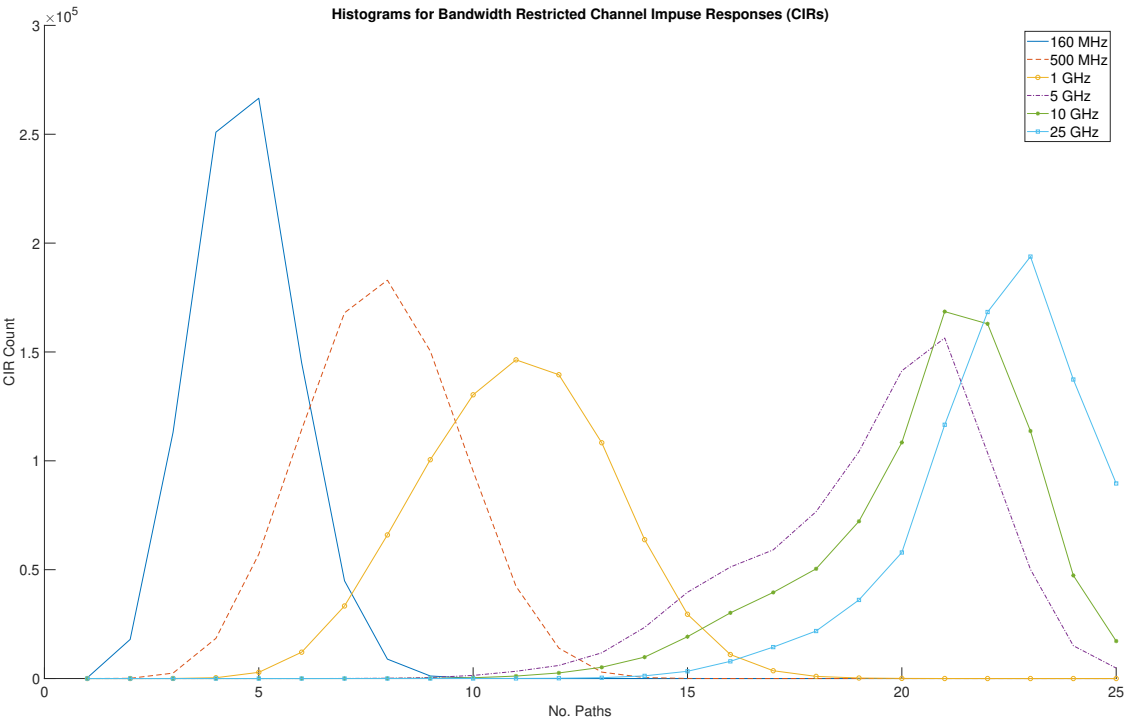
Appendix J: Histograms

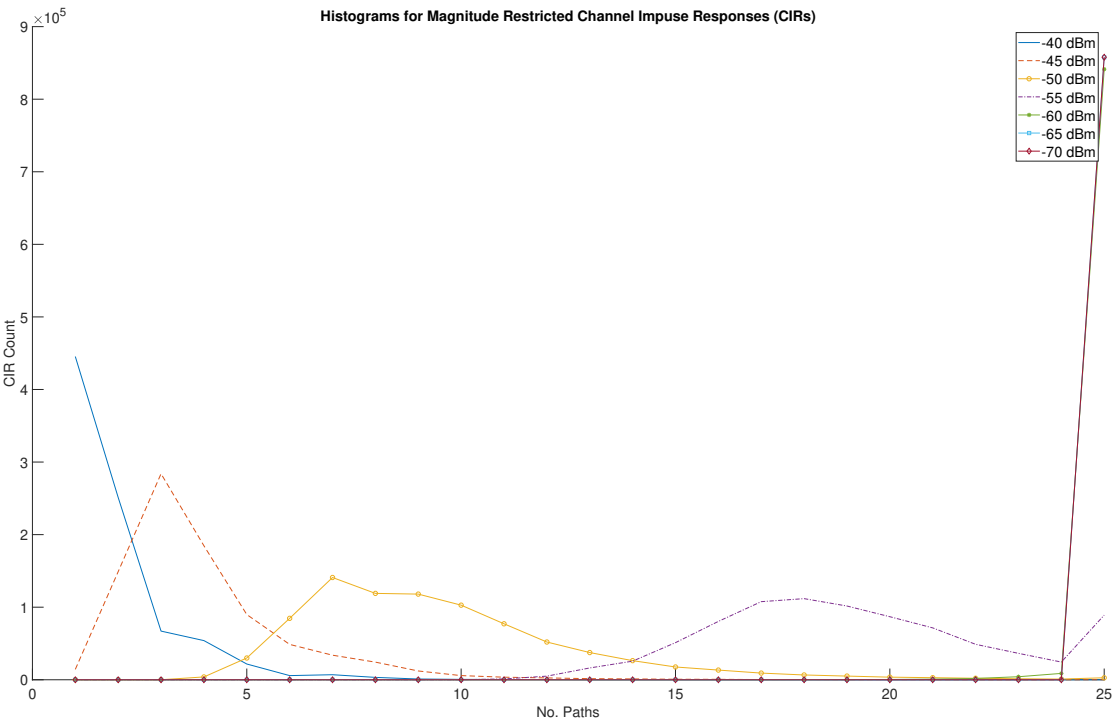FIGURE J.1: No. Path Histograms for Bandwidth Limited CIRs

FIGURE J.2: No. Path Histograms for Magnitude Limited CIRs