Computer Science Master - Engineering Track
UBINET

Internship Final Report

Data Augmentation for Wi-Fi Fingerprint-Based Localization in Indoor Environments

Student: Angelo Rodio
Academic Supervisor: Giovanni Neglia
Internship Supervisors: Arnaud Legout, Chadi Barakat

August 2020

Contents

Abstract

Precise localization of smartphones in indoor environments is a challenge. Accurate outdoor localization has been exhaustively implemented through Global Navigation Satellite Systems. On the contrary, indoor localization still continues to be a challenge even though it has been exploited during the last several years. Many indoor localization techniques have been explored in the scientific literature. Most of them require the installation of dedicated infrastructures that is expensive to perform and hard to scale. Indoor localization systems that reuse the building existing infrastructures, on the other hand, can provide high localization performance but require a large effort for the fingerprint database construction. This task is time-consuming and labor-intensive [1]. The excessive time to perform fingerprint data collection affects the popularity of fingerprint-based localization.

In my internship, I developed three methods for fingerprints augmentation, namely WF-LERP, WF-DCGAN, and WF-RND, to truly expand the fingerprinting database while reducing the human effort, by virtually increasing the number of fingerprints collected at each location.

The proposed solutions reduce the time expense involved in the data collection phase of a factor 20, or, under the same time, increase the accuracy performance of the indoor positioning algorithm up to the 52%. Tests are conducted with fresh fingerprints and in presence of scans that suffer a 4-months time-aging. Extensive experiments and results prove the effectiveness and the advantages of the proposed methods over the existing state-of-the-art solutions.

Keywords: Indoor Positioning System, Wi-Fi Fingerprinting, Generative Adversarial Networks.

1 Introduction

1.1 Context: a world in development

Accurate, reliable and real-time positioning systems play an important role in our life nowadays [2]. The progress in chip miniaturization and the advent of modern electronic devices allowed the development of new positioning technologies. However, in the past, the use of such technologies was restricted to military and professional applications, such as the navigation of ships and aircraft or land surveying. Only after 2000, the Global Positioning System (GPS) has been made accessible for civilian use [3]. Since then, positioning technologies have become an important requirement to ordinary people's daily activities.

1.2 Motivations: need for location-based services

Location Based Services (LBSs) are applications integrating geographic location (i.e., spatial coordinates) with the general notion of services and find wide use in military, navigation, monitoring and tracking, healthcare, retail, inventory, and marketing [4]. With the development of mobile communication and devices, these applications have become a novel challenge both conceptually and technically. A positioning system enables a mobile device to determine its position and makes the position of the device available for LBSs. In recent years, the development of new technologies (5G and Augmented Reality, among all) is revolutionizing the market and enables integration of traditional applications with LBSs. The increasing demand for localization services asks for innovations in positioning technologies.

1.3 Research objectives, requirements and challenges

To address the increasing demand for location-based services, in my internship I decided to face the challenge of precise localization in indoor environments. The objective of my research activity deals with the question: is it possible to estimate the position of a mobile device in an indoor environment, when the following requirements are satisfied?

1. Do not require dedicated hardware. Prefer reusing existing infrastructures;
2. Do not require any knowledge about the indoor environment structure (e.g. floor-plan of the buildings, position of the APs);
3. Do not require the user to be paired in advance with the positioning base station. Support unassociated clients.

The challenge of my work is to develop a positioning system that satisfies all these requirements together. Relaxing one or more requirements would enormously simplify my research activity, but also add limiting constraints on the usability of the positioning system.

1.4 Personal contributions

The personal contributions of my internship are:

1. A comprehensive analysis of the state-of-the-art for Indoor Positioning Systems, with particular focus on advantages and disadvantages of each solution. The re-implementation of a state-of-the-art milestone article about Indoor Positioning Systems based on Wi-Fi Fingerprinting, that enables a comprehensive comparison of positioning algorithms on a well-known challenging dataset;
2. The identification of bottlenecks in the current Wi-Fi fingerprinting positioning algorithms: the high cost of the data collection process. Experiments to show the impracticability of current fingerprinting techniques.
3. I proposed three methods for fingerprint generation that are inspired from image augmentation techniques and allow to cut-off the data collection cost in terms of time and human labor. Extensive experiments and results prove the effectiveness and the advantages of the proposed methods.

1.5 Structure of the thesis

The remainder of this thesis is organized as follows. Chapter 2 reviews the state-of-the-art, presenting particular attention on the motivations that lead to my choice of studying Indoor Positioning Systems based on Wi-Fi fingerprinting. Chapter 3 presents a re-implementation of a deep neural network positioning algorithm, and its comparison with other current Wi-Fi fingerprinting positioning methods. Chapter 4 analyzes the data collection cost of current fingerprinting implementations, showing the impracticability of the existing solutions. In Chapter 5, I present the three methods developed to reduce the data collection cost through virtual fingerprint generation, together with extensive experiments that prove their effectiveness. Finally, Chapter 6 concludes the thesis.

2 State of the art

2.1 Basics of Wireless Positioning

Wireless Positioning Systems

A Wireless Positioning System (WPS) can be defined as a system that makes use of signals transmitted and/or received over a wireless communication channel to determine the position of a target device [5].

Base stations and mobile devices

In wireless positioning systems, terminologies like mobile station and base station are often used [6]. A mobile station is the object to be located, such as a person with a mobile phone or a positioning tag, an animal with an attached tracking device, or a vehicle. A base station is a fixed point of communication that transmits and receives signals, usually of known position. In Global Navigation Satellite System (GNSS), a satellite may be called a base station, while any mobile object that carries a GNSS receiver (a chip in a smartphone) may be called a mobile station. In indoor environments, a base station can be associated with a fixed ultra-wideband device, a Wi-Fi Access Point, or a mobile phone cellular network.

2.1.1 Wireless Positioning Systems: Performance Metrics

Wireless Positioning Systems use numerous positioning mechanisms that vary in terms of cost, accuracy, precision, technology, scalability, robustness, and security [7]. Some applications may require low-cost WPS, whereas others may require high accuracy WPS, such as military and medical ones.

Accuracy. The higher the accuracy, the better the positioning results of the method, thus the greater the likelihood of the estimated position being close to the mobile device's real position. In this thesis, the positioning accuracy is defined as the number of times the position of the target device is correctly estimated over the number of total predictions:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{1}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives.

Availability. Time percentage during which the positioning service is available, taking into consideration the required accuracy and integrity. The integrity is defined as the confidence of the WPS output. Generally, availability can be seen as three levels: low availability ($< 95\%$), regular availability (between 95% and 99%), and high availability ($> 99\%$) [8].

Coverage Area. Area that is covered by the WPS. The levels of coverage can be: local, scalable, and global [9]. Local coverage refers to a limited area that is well-defined and is not extendable such as a building, while scalable coverage refers to the ability of a system to increase the area by adding hardware. On the other hand, global coverage refers to a system that has a worldwide area such as Global Navigation Satellite System. At the state of the art, for indoor positioning, the coverage area ranges from 5 to 50 m. The design of building systems with > 60 m coverage area is challenging [8].

Complexity. The complexity of a positioning system can be evaluated considering the complexity of its positioning algorithm and the level of its hardware. The spatial and time complexity of the algorithm used in the positioning method can directly affect the performance of the positioning system. Due to the limited processing capacity of wireless terminal equipment, especially mobile devices, the low spatial complexity of a positioning method represents low energy and memory consumption, while low time complexity represents low positioning delay and high dynamic adaptability.

Scalability. Scalability describes whether the positioning system can maintain good performance when the system expands. A wireless position system can be considered scalable in two ways: considering the coverage area; considering the number of mobile devices. The scale of the coverage area considers the case when the geographic area increases (e.g., from buildings to a campus, within a metropolitan area). The scale of the number of users indicates that the number of units located per time period per geographic area increases [10].

Cost. The cost of a WPS can be evaluated in terms of money, time, space, and energy. These can be affected at various levels of the system: system installation and maintenance, infrastructure components, and positioning devices [10], [11]. The cost for system installation and maintenance includes cost required for installation, and any expenses that are required to maintain the system functionality; whereas the cost for infrastructure components and positioning devices may include the costs of buying components and preparing them, as well as the space and energy needed to run those components. Some WPSs, particularly those that reuse existing infrastructures such as the network, are more cost-effective.

Privacy. Privacy should really be of concern for individuals who use WPSs. Users' personal and sensible information is collected and used by wireless positioning systems [11]. In order to improve users' privacy, security mechanisms should be implemented and maintained to protect data from intrusion, theft, and misuse. Unfortunately, the privacy aspect of WPSs was not addressed sufficiently in the indoor positioning literature [9].

2.2 Outdoor Positioning: GPS

Without a doubt, first thing people usually think of when talking about Wireless Positioning Systems is GPS. GPS stands for Global Positioning System, and was developed for military purposes. In May 2000, it became available to civilian uses [3]. Thanks to the progresses in chip miniaturization and the advent of modern electronic devices, nowadays all smartphones embed a GPS receiver and include a navigation software that allows users to find the phone's location on a map or a path to a nominated destination.
GPS can be classified as a Global Navigation Satellite System (GNSS) [12]. Global positioning is enabled to provide services with global coverage by use of satellites. The characteristics of GNSS are:

1. The satellite infrastructure is built and maintained typically by a nation's military, and then made available for civilian uses [12].

2. The civilian design function is limited to the mobile device, and most of the other technological and operational characteristics are beyond the scope of civilian designers [5]. The civilian aspects that are challenging for researchers have been restricted to developing the cheap GPS receiver hardware, and developing the mathematical algorithms and enhanced signal processing required for improving GPS performance [12].

GPS: the best choice in outdoor environments

In outdoor environments, GNSS is the first and probably the best choice for positioning and navigation. Since the satellite networks are already deployed and the rapid development of microelectronics is no longer a major limitation, GNSS outperforms the other wireless positioning systems in terms of positional accuracy, mobile size, battery life, positioning time and cost [5]. As a matter of fact, all modern smartphones include an embedded GNSS chip. In the air, similar to the open space on the ground, GNSS is the best option for positioning and navigation of aircraft, UAVs, and missiles.

GPS: working principle

GPS system has three basic components: satellites, ground stations, and receivers [12]. Grounds stations use radars to find out if the satellites really are where they're supposed to be. A receiver in the mobile device measures signals from satellites to determine its distance from them. When the distance information from four or more GPS satellites is available, it can easily estimate the position of the mobile device, with an accuracy within meters.
The GPS system has 32 active satellites orbiting the Earth. 24 of them are core satellites, and the rest serve as emergency replacements. They need constant maintenance and sometimes repairs, but even with all that, they only last about 10 years. GPS works in any weather, rain or shine; but there is one important condition: a receiver on Earth has to see at least 4 satellites to calculate an accurate location because the GPS uses a trilateration mechanism.

Distance estimation The satellite sends intermittent radio signals down to Earth. These radio signals contain the exact time the signal was sent, and the position of the satellite. The GPS receiver analyzes radio signals from the GPS satellites to figure out two important things: the location of at least 4 satellites in space, and the distance between the receiver and those satellites. Since radio waves travel at the speed of light, the receiver receives the signal after a certain time duration. By finding out the difference between the sent and received times, and multiplying it by the speed of light, the receiver will be able to find out the distance between itself and the satellites [13].

Trilateration Trilateration is about calculating latitude, longitude, and altitude position on a map. At least 4 satellites are required. The method basically consists in considering distance information and draw spheres on a map. Let us assume that the receiver knows it is at distance R1 from satellite A. So, its position is somewhere within the sphere of center satellite A and radius R1. Same reasoning for the distance between receiver and satellite B. This means the actual location should satisfy both these spheres. The intersection between two spheres is a circle. Using the distance from satellite C to build another sphere, there are two points of intersection. A simple trick is to consider the Earth's surface as the forth circle, and eliminate the improbable solution. The result is the estimated position of the mobile device on the map [13].

GPS: need for very accurate time measurements

In GPS, the time measurement has to be very accurate. Even an error of microseconds will give an error in the range of kilometers, since the speed of light is huge. Here comes the main issues:

1. Receivers can not have highly accurate clocks. Satellites have atomic clocks that keep the most precise time, but it would be impossible to install these clocks in every receiver. They cost somewhere between $50000 and $100000, so it would make mobile phones really, really expensive. Today, receivers use crystal clocks that are not accurate when compared to atomic clocks.

The difference between the actual time measured by the satellites and the time measured by the receiver is called time offset. This time offset cause a huge error in GPS calculations. How do we overcome this issue? Since the time offset of the receiver with all three of the satellites is the same, the time offset of the mobile device becomes the forth unknown. This means, we need an extra satellite measurement to solve this fourth unknown, and that is why we need four satellites to measure the user's location [13].

2. Second complication: time is not absolute. According to Einstein's theory of special relativity, a fast moving clock will slow down. The atomic clocks, which are moving at speed of 14000 kilometers per hour, will slow down by 7 microseconds every day. At an altitude of 20000 kilometers above the Earth, the satellites experience one quarter of the Earth's gravity: their clocks will tick around 45 microseconds faster every day. This means a net 38 microseconds offset is created every day in the atomic clock. To compensate for this, a theory of relativity equation is integrated into the computer chips, and adjusts the rates of the atomic clocks. Without this application of the theory of relativity, the GPS would have produced an error of 10 kilometers every day [13].

GPS: not applicable in 3 scenarios

In a perfect world, GNSS would be the best option for positioning and navigation everywhere. Unfortunately - that is, fortunately for researchers - there are three scenarios where GNSS is not applicable [5].

1. **Underground/Underwater.** Radio signals cannot be used due to the very short transmission range, particularly in salty water. Acoustic signals are almost universally used, together with Inertial Navigation Systems.

2. **Deep space.** In the deep space, where GNSS is not applicable, positioning and navigation rely on locating reference objects such as bright stars, pulsars or quasars for accurate time reference.

3. **Indoor environments.** When walking from an outdoor open space into a building, even a large one, or into a dense forest, GNSS signals are blocked. This proclaims the unsuitability of global navigation satellite systems for indoor applications.

2.2.1 Why GPS is not suitable for Indoor Positioning

Since GPS uses a trilateration mechanism, a fundamental requirement for GPS to estimate a meaningful location is that the receiver positioned on the Earth can receive electromagnetic signals from at least four satellites. Whenever this condition is not satisfied, global navigation satellite systems are not able to calculate an accurate location of the mobile device. Each of the four communications required between the transmitter, i.e., the satellite, and the receiver on the Earth, require the existence of a Line-of-Sight (LoS) propagation path.

Indoor environments cannot always guarantee the existence of a direct path between satellites and mobile device. More often, no direct path signal is received by the mobile equipment. Differently from outdoor environments, in indoor environments the receiver captures a number of multiple paths from the transmitter after they go through multiple reflections and attenuation. The presence of multiple objects (such as people, walls, and pieces of equipment) reflects signals and leads to multi-path and delay problems. As a result, indoor environments are typically characterized by Non-Line-of-Sight (NLoS) propagation, which causes inconsistent time delays at the receiver.

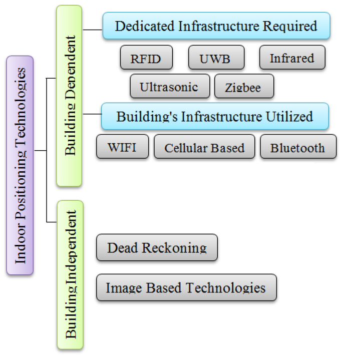


Figure 1: Classification of Indoor Positioning Technologies [14].

In the section about the GPS working principle, I motivated how important is for the time measurements to be accurate. Computing the time interval needed by an electromagnetic wave to travel from the transmitter to the receiver is possible by exploiting the phase information. The phase information extracted by the receiver from multiple waves that went through multiple reflections is still a mathematical quantity, but deprived of any physical meaning. A time measurement that is certainly incorrect propagates its error to the distance estimation between satellite and receiver. Indeed, considering that radio waves travel at the speed of light, from the relation $distance = speed \times time$, it is evident that the time estimation gets multiplied for the speed of light, $c \approx 3 \times 10^8$ m/s, thus resulting in a considerable error in the distance estimation. The position returned by satellite-based systems when used in indoor environments would be completely wrong.

2.3 Indoor Wireless Positioning

GPS and in general GNSS satellite systems do not work in indoor environment [5]. This problem was clear for researchers since the early developments of the GNSS satellite systems, but it remained unfaced for a very long time.
In my academic career, it happened many times for me to observe that scientific interest on a research topic is usually influenced by the outside world and its technological innovations. Ideas that seemed futuristic and infeasible in a historical period can gain interest in a relatively short amount of time. This is exactly the case of indoor localization systems.
In the 2010s, the rapid development of smartphones and their progressive entrance in people's everyday lives created a new generation of mobile devices that mostly moves in indoor environments and enabled a new need for services based on the location of the mobile device owners.
Over the years, many different solutions and implementation have been proposed both by the research and industry environment [5]. The development of Indoor Positioning Systems (IPSs) investigates the science of indoor radio propagation, and the development of algorithms that mitigate the effects of the complex indoor propagation environment. What is also challenging in this research effort is the engineering required for cheap effective hardware for terrestrial applications which GPS cannot satisfy.

2.3.1 Indoor Positioning Technologies

A number of indoor positioning systems have been developed [5]. In the literature, different kinds of signals have been employed for positioning. Among these: radio, acoustic, light, and magnetic

signals. The signal selection for a particular requirement depends on several factors, such as the application scenario, the performance requirements (such as accuracy and coverage), and the cost.
Indoor positioning technologies are classified by researchers in many different ways [14]. A very interesting classification work, that merges and complements heterogeneous and independent contributions from many authors, is provided by Alafiri et al. in the article *"Ultra wideband indoor positioning technologies: Analysis and recent advances"*, 2016 [14].
As shown in Figure 1, the authors classify indoor positioning technologies into two main classes: indoor positioning technologies that are dependent on the building and indoor positioning technologies that are building independent.

Building Dependent Indoor Positioning Technologies

Building dependent indoor positioning technologies refer to technologies that depend on the building that they operate in. They depend either on an existing technology in the building or on the map of the building. Building dependent indoor positioning technologies can be further divided into two major classes: indoor positioning technologies that require dedicated infrastructure and indoor positioning technologies that utilize the building's infrastructure.

Dedicated Infrastructure Required

Indoor positioning technologies that require dedicated infrastructure are radio frequency that is either RFID (1) or UWB (2); infrared (3); ultrasonic (4); and Zigbee (5) [14].

(1) Radio Frequency Identification (RFID). Radio Frequency Identification uses radio waves to transmit the identity of an object (or person) wirelessly. It exchanges radio signals between readers and tags. Tags are attached to all the objects that need to be tracked. Although different positioning methods can be used with RFID, proximity is the most used one and it senses the presence of RFID tags rather than the exact position [9], [10], [15].

(2) Ultra Wideband (UWB). UWB is whatever RF signal occupying a portion of the frequency spectrum greater than 20% of the center carrier frequency, or with a bandwidth greater than 500 MHz. This allows UWB transmitters to transmit large amounts of data while consuming little transmit energy [16]. UWB can be used for positioning by utilizing the time difference of arrival (TDOA) of the RF signals to obtain the distance between the base station and the target [17].

(3) Infrared (IR). Infrared wireless communication makes use of the invisible spectrum of light just below the red edge of the visible spectrum [9], [16]. IR uses wide angle LEDs which emit signals in many directions. Thus, it does not require direct line of sight, which makes this technology less intrusive than indoor positioning based on visible light. Proximity, differential phase-shift, and angle of arrival (AoA) positioning methods are frequently used with Infrared technology [18], [19], [20].

(4) Ultrasonic. An ultrasound wave is a mechanical wave, that is an oscillation of pressure transmitted through a medium [9]. It does not interfere with electromagnetic waves and has relatively short range. The relative distance between the different devices can be estimated using time of arrival (ToA) measurements of ultrasound pulses traveling from emitters to the receivers [16].

(5) Zigbee. ZigBee is a standard for short distance and low rate wireless personal area networks [9]. A ZigBee node is small and has low complexity and cost. It is designed for applications that require low power consumption and low data throughput. Usually, RSS values are used to estimate a distance between Zigbee nodes [16].

Reuse Existing Building Infrastructure

Indoor positioning technologies that utilize the building's infrastructure are Wi-Fi (6); cellular based (7); and Bluetooth (8).

(6) Wireless Local Area Network (WLAN). IEEE 802.11 is considered the dominant local wireless networking standard. The IEEE 802.11 WLAN standard defines the protocol and compatible interconnection of data communication equipment via the air in a local area network (LAN) using the carrier sense multiple access protocol with collision avoidance (CSMA/CA) medium sharing mechanism [9]. Using WiFi in indoor positioning and navigation systems depends on knowing a list of wireless routers that are available in an area in which the system operates. The most popular WLAN positioning method is received signal strength (RSS) which is easy to extract in 802.11 networks and could run on off-the-shelf WLAN hardware [16]. Using RSS, the accuracy of WLAN positioning systems is around 3 to 30 m. Time of arrival (ToA), time difference of arrival (TDoA), and angle of arrival (AoA) mechanisms are less common in WLAN because of the angular measurements and time delay complexity [10].

(7) Cellular Based. Global System for Mobile Communications (GSM) networks are available in most countries and can outreach the coverage of WLAN with lower positioning accuracy. The most common method of GSM indoor positioning is fingerprinting, which is based on the power level (RSS) [16].

(8) Bluetooth. Bluetooth is a proprietary format managed by the Bluetooth Special Interest Group (SIG) and a standard for wireless personal area networks (WPANs) [16]. Bluetooth is designed to be a very low power technology for peer-to-peer communications, and it operates in the 2.4-GHz ISM band. In comparison with WLAN, the gross bit rate is lower and the range is shorter (approximately 10 cm to 10 m). The Bluetooth SIG groups include a local group that investigates the use of Bluetooth wireless technology for positioning. Bluetooth technology commonly uses proximity and RSS methods to estimate positions [9], [10].

Building Independent Indoor Positioning Technologies

Indoor positioning technologies that are building independent do not require any information about the building. In dead reckoning, an object can determine its current position by knowing its past position, its speed and the direction in which it is moving [21]. In many applications, the knowledge of the mobile device's past position is not available [22]. Image based technologies mainly rely on a camera (e.g., sensor and image processing), and make use of a map of the building [23].

2.3.2 Advantages of Indoor Positioning Technologies based on Wi-Fi

Disadvantage of requiring a dedicated infrastructure

Indoor positioning technologies that are building dependent require some sort of information about the building. Indoor positioning systems based on technologies as RFID or UWB obtain this information through the use of dedicated hardware [16]. Thanks to the use of dedicated hardware, this set of techniques can guarantee the best results in terms of positioning accuracy. However, the deployment of dedicated infrastructure is considered to be expensive, either in terms of costs and time [9]. Not only the purchase, but mainly the installation costs, justify the choice for this solution only when the accuracy requirements are in the order of one meter, or less.

Advantage of reusing the existing building infrastructure

Indoor positioning systems based on technologies that reuse the existing building's infrastructure - i.e., Wi-Fi, cellular based, and Bluetooth - still require piece of information from the indoor environment. This kind of information can be either obtained through measurement piece of hardware, this time installed in the mobile equipment - this first solution still requires the use of dedicated hardware that is not yet integrated in today's mobile devices; or through the collection of fingerprints of the existing EM signals measured in the environment - advantage of totally reusing the infrastructure already present in the environment, but requires a data collection phase [16].

My choice of studying methods that allow to reuse the existing building infrastructure

At the state of the art, research developed on indoor positioning systems based on technologies that can reuse the existing building's infrastructure allows these methods to achieve accuracy performance almost comparable to the ones obtained when dedicated infrastructure is deployed. This motivation, together with the convenience of reusing the building existing infrastructure and the possibility to use commercially available off-the-shelf (COTS) mobile devices, conditioned my choice to focus my study on indoor positioning technologies that reuse the building existing infrastructure.

The choice of focusing my work on Wi-Fi based methods

Cellular based indoor positioning systems usually estimate the mobile device position in the coverage area of the cell of the mobile network, and achieve bad performance in terms of accuracy. Research on Wi-Fi and Bluetooth-based methods is more advanced. Either Wi-Fi and Bluetooth are promising technologies for indoor positioning systems. Both of them reuse the existing building infrastructure and achieve satisfactory results in terms of accuracy. However, the Wi-Fi infrastructure is deployed more than the Bluetooth one. According the data collected by ElectroSmart [66], in nearly every house, today, is possible to measure Wi-Fi EM signals emitted by wireless access points deployed in the same building and neighborhoods. This is the reason why I decided to focus my work on indoor positioning technologies based on Wi-Fi. However, since for both technologies it is possible to measure the Received Signal Strength (RSS) in dBm, the work presented for Wi-Fi based indoor positioning systems can be implicitly extended for Bluetooth based ones.

2.4 Indoor Positioning Methods based on Wi-Fi

The potential of commercial off-the-shelf indoor positioning systems has greatly improved with the wide distribution of Wi-Fi. The high precision, low-power consumption, and low-cost has made it an area of great interest in indoor positioning technology research [9]. The IEEE 802.11 standards enable a Wi-Fi wireless network card with the function of measuring the intensity of radio frequency signals transmitted by wireless access points. Therefore, users can utilize mobile devices such as smartphones, laptops, tablet PCs, and others to achieve indoor positioning with Wi-Fi and certain algorithms [10].

So far, there are two types of Wi-Fi-based positioning technologies: methods based on computation that allow to achieve a fine-grained localization accuracy and methods based on estimation that allow to achieve a coarse-grained localization accuracy.

2.4.1 Wi-Fi fine-grained localization methods based on computation

Wi-Fi based positioning system methods based on computation that allow to achieve a fine-grained localization accuracy rely on the measure of Time and Space Attributes of the Received Signal (TSARS). TSARS-based positioning methods mainly use time and space attributes of received signal such as spatial distance, the time signal sending-receiving consumes and the spatial angle of signal to estimate the distance between the mobile device and a wireless access point. Once the position of at least three wireless access points is known and their distance values with respect to the mobile devices are estimated, the trilateration method - the one used in GPS - allows to analytically compute the position of the user. The methods used to estimate the distance between the mobile device and a wireless access point includes Time of Arrival (1), Angle of Arrival (2), and Time Difference of Arrival (3) [24].

(1) Time of Arrival (ToA). Time of Arrival consist in measuring the absolute time that the radio signal transmitted from the wireless access point takes to travel to the mobile device that acts as remote receiver. The distance can be directly calculated from the time of arrival as signals travel with a known velocity [25].

(2) Angle of Arrival (AoA). The angle of arrival of a signal is the direction from which the signal is received. Measurement of AoA can be done by determining the direction of propagation of a radio-frequency wave incident on an antenna array or determined from maximum signal strength during antenna rotation. The mobile device would calculate the AoA of signals transmitted by multiple access points in the environment [27]. This information would be then combined to the knowledge of the position of the wireless access points determine the user's location.

(3) Time Difference of Arrival (TDoA). Time Difference of Arrivals measures the delay time of the arrival signal. The idea is that the mobile device measures the time of arrival of RF signals at several points in space and then compares the time difference between each wireless access point. By knowing the location of each access point, an estimate of the location of the mobile device can then be deduced provided all access points are time synchronized [26].

2.4.2 Wi-Fi coarse-grained localization methods based on estimation

Wi-Fi based positioning system methods based on estimation that allow to achieve a coarse-grained localization accuracy rely on the measure of the Received Signal Strength (RSS) of Wi-Fi signals. The RSS-based positioning methods utilizes the strength of the received signal to obtain the position of the user. Known the measure of received signal strength from multiple access points at the mobile device's location, the position of the user can be estimated through the use of propagation models (1) or through a data driven approach based on fingerprinting (2) [10].

(1) Propagation models and trilateration

Suppose that exists a very accurate path loss propagation model for indoor environments that, given as input the RSS value in dBm measured at the mobile device, is able to return a proper estimation

of the distance between the mobile device and the wireless access point. Suppose that the received signal strength measured at the mobile device from at least three different wireless access point is known. Suppose that the position of the three wireless access point is known. Combining the propagation model with the trilateration technique, this method would be able to exactly compute the user's position.
In reality, the relation between received signal strength and distance in indoor environments is very difficult to predict. The complexity of indoor space can have a great impact on the RF signal, and the conversion of signal strength to spatial distance inevitably produces errors. The result is that, even using trilateration, the final position is obtained as estimation instead of computation. As a result, the Wi-Fi RSS-based positioning systems that use this method produce significantly large positioning errors, with a positioning accuracy in the range 8–10 m [28].

(2) Fingerprinting

Fingerprint matching is a very promising technique for indoor localization. It consists in the collection of received signal strength values from different access points at same location with the aim of creating fingerprints that are unique identifiers of that specific location. The physical location in which the signal strength value is measured is called Reference Point (RP). Fingerprint-based localization consists of two phases [10].

1. In an offline data acquisition phase, fingerprints are collected at each reference point to create what is called fingerprint database or radio map.

2. In the online matching phase, the mobile device measures real-time RSS values and runs a positioning algorithm to compare the measure with the fingerprints stored in the database. The reference point that most closely matches the received data is assumed to be device's location.

2.4.3 Pro and Cons of Indoor Positioning Methods based on Wi-Fi

The market of indoor positioning system has gradually started to move towards solutions based on Wi-Fi [10]. The great advantage of reusing the infrastructure already existing in the building has made it an area of great interest in indoor positioning technology research. The possibility to localize commercial off-the-shelf smartphones with relatively low effort has encouraged studies. As a result of these works, the precision of indoor positioning methods based on Wi-Fi is progressively increasing.

Disadvantages of Wi-Fi fine-grained localization methods based on computation

Wi-Fi fine-grained localization methods based on computation rely on the measure of time and space attributes of the received signal (TSARS). The measure of time of arrival, angle of arrival, or time difference of arrival requires dedicated hardware devices that should be added to the mobile equipment. Commercial off-the-shelf smartphones are not suitable as mobile equipment. The deployment of the new hardware has a not negligible cost. The trilateration method requires the exact knowledge of the wireless access points position: this information is not always available. Computing time and space attributes require the user to be associated to the wireless access points. The user should be aware about the service set identifiers (SSID) and passwords of all the access points used for positioning: a significant problem concerning security and privacy of administrators.

My choice of studying Wi-Fi coarse-grained localization methods based on estimation

Due to the limitations of the TSARS-based positioning methods, I have decided to orient my research effort on Wi-Fi coarse-grained localization methods based on estimation.
The great advantage of receive signal strength (RSS) methods over TSARS ones is the easiness of the measurement process. While the collection of time and space attributes requires dedicated hardware, a commercial off-the-shelf smartphone with existing software is more than enough to collect RSS values from multiple access points. The reuse of already available hardware, together with the advantage of Wi-Fi based positioning systems that reuse the existing building infrastructure, plays an important role in the potentiality of this method. What is more, differently from TSARS-based methods, Wi-Fi localization methods based on RSS support unassociated clients. The measure of RSS values can be done through a commercial wireless network interface card without the user being aware of the network SSIDs and passwords.

Disadvantages of Wi-Fi RSS-based localization methods that make use of propagation models and trilateration

Path loss propagation models aim at finding a relation between the intensity of RF signals transmitted by the access points and measured at the mobile equipment and their distances. Outdoor propagation is fairly predictable: many outdoor propagation models have been proposed in the literature: Hata – Okumura, Walfisch - Ikegami, Lee, and others. On the contrary, indoor propagation models are empirical in nature. RF signal propagation in indoor environments is complex and can be affected by the multipath effect, human body interference, and other factors. This creates many difficulties when trying to find a mathematical formulation for the characterization of radio wave propagation as a function of distance. Indoor environments are nearly unpredictable: a location that is close to the transmitter may measure a RSS-value that is low; a location that is far from the transmitter may measure a RSS-value that is high. There is no path loss propagation model for indoor environments that, given as input the RSS value in dBm measured at the mobile device, is able to return a proper estimation of the distance between the mobile device and the wireless access point. Positioning methods based on propagation models usually results in low position accuracy, in the range 8–10 m.
In addition, positioning methods based on propagation models require trilateration. The main drawback of trilateration is demanding the knowledge of the position of all the access points in the environment. This is unpractical, since many signals are emitted by access points from other buildings and surroundings.

2.5 Indoor Positioning Systems based on Wi-Fi Fingerprinting

In fingerprinting-based methods, the location of the mobile device is determined by comparing observed RSS with a radio map of previous collected RSS [6]. Fingerprinting-based methods are generally divided in two phases (Figure 2): an offline phase of data collection and an online phase in which the position of the mobile device is estimated.

2.5.1 Offline Phase: Data Collection

In the offline phase, a number of scans are collected. In each scan, the mobile device acts as a receiver and measures the RSS-values from M access points (APs). Each scan can be represented as an array of length M:

$$S_i = [rss_{i,1}, rss_{i,2}, \cdots, rss_{i,j}, \cdots, rss_{i,M}] \,, \tag{2}$$

where $rss_{i,j}$ is the RSS-value measured by the mobile device during the i-th scan from the j-th access point, $i = 1, \ldots, N$ and $j = 1, \ldots, M$.
Let us assume that the total number of scans collected during the data collection phase is N. Each scan is associated to a physical location, that is the location in the indoor environment where the scan has been measured. The location identifier of a scan is called Reference Point (RP). For room-based indoor positioning, the RP is stored in the form of a label entered by the user during the data collection phase.
In my notation, I define as fingerprint any pair (Reference Point, Scan). The generic fingerprint F_i is indicated by:

$$F_i = (RP_i, S_i) = (RP_i, [rss_{i,1}, rss_{i,2}, \cdots, rss_{i,j}, \cdots, rss_{i,M}]) \,, \tag{3}$$

with $i = 1, \ldots, N$.
The fingerprint database, or radio map, is constructed as a matrix that includes all the fingerprints from all the scans in all the reference points:

$$FD = \begin{bmatrix} (F_1) \\ (F_2) \\ \vdots \\ (F_N) \end{bmatrix} == \begin{bmatrix} (RP_1, S_1) \\ (RP_2, S_2) \\ \vdots \\ (RP_N, S_M) \end{bmatrix} = \begin{bmatrix} rss_{1,1} & rss_{1,2} & \ldots & rss_{1,M} \\ rss_{2,1} & rss_{2,2} & \ldots & rss_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ rss_{N,1} & rss_{N,2} & \ldots & rss_{N,M} \end{bmatrix} . \tag{4}$$

In general, more than one scan is measured for one reference points. All scans are stored in the fingerprinting database.

2.5.2 Online Phase: Positioning Algorithms

Whenever the user demands to find his location, a new Wi-Fi scan is measured by the mobile device in the position of the user. This target location is called Query Point (QP), and the scan associated to the QP is called Query Scan (QS):

$$QS = [rss_{QS,1}, rss_{QS,2}, \cdots, rss_{QS,j}, \cdots, rss_{QS,M}] \,. \tag{5}$$

In the online phase, a positioning algorithm is executed with the aim of estimating the position of the QS. The estimation returns the identifier of the environment where the query scan has more likely been taken. The target location can be considered as the result of a classification problem, where the scans in the fingerprinting database are the feature vectors and the classes are the labels that uniquely identify the environments, i.e., the Reference Points.

2.5.3 Challenges of Indoor Positioning Systems based on Wi-Fi Fingerprinting

My choice of studying data-driven RSS localization methods based on fingerprinting

Data-driven techniques are totally empirical [29]. In my opinion, data-driven approaches inherently reflect the physics of indoor environments, that are empirical by nature. The use of propagation models try to fit the unpredictability of indoor propagation into a mathematical formulation. On the contrary, the fingerprinting approach just accept the randomness of indoor environments and tries to take advantage of it to create a unique identifier of the specific location. Furthermore, the fingerprinting approach inherits the advantages of the Wi-Fi RSS-based localization techniques, that

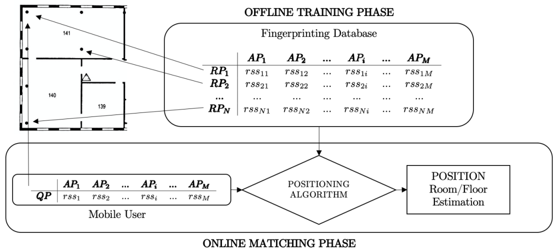


Figure 2: A diagram of the fingerprinting method.

are: 1) reusing the existing WLAN infrastructure; 2) not requiring additional positioning devices; 3) supporting unassociated clients.
All these advantages are the reason why this approach has attracted progressively more and more attention by researchers since it was first developed [10]. Without requiring new hardware or external infrastructure, the location of the AP, or the specific distribution of interior space environment, but just exploiting the surrounding APs, Wi-Fi fingerprinting localization techniques can obtain satisfactory positioning results.

Challenges of Wi-Fi positioning systems based on RSS-fingerprinting techniques

After reading the list of the advantages, one may think that Wi-Fi localization techniques based on the fingerprinting approach are the solution to the indoor localization problem. Unfortunately, at the state of the art, no fingerprinting positioning method has been implemented in the applications available on the well-known software stores for smartphones. Research on indoor fingerprinting technique has started only in the last decade, and many open problem still need to be solved [29]:

1. **Development of efficient positioning algorithm: from coarse-grained to fine-grained positioning accuracy.** Indoor fingerprinting methods are in the family of the RSS-based positioning systems, that are based on estimation and usually provide a coarse-grained positioning accuracy. Research is focusing on the challenge of the development of efficient positioning algorithms to endow these methods with fine-grained positioning accuracy [30]. Considering its data-driven nature, the fingerprinting approach benefit from developments made by data scientists in machine learning algorithms. Computation consumption and algorithm complexity can be relatively high. Furthermore, the final positioning accuracy is greatly affected by the indoor environment changes. Understanding the influence factors of the environment on the Wi-Fi signal strength can help to reduce the positioning error.

2. **Enhancing the data collection procedure, in terms of cost and time.** The data-driven nature of this approach requires a large amount of a priori information as data support [30]. Current data collection procedures are expensive, with regards to both time and cost [29]. Cost reduction is crucial for the widespread use of this approach. Only methods whose hardware is cheap, that are not time consuming and computationally inexpensive find space in the market [29].

Personal contributions. During my internship, I tried to face both challenges. The former is discussed in Chapter 3. The second, on which I dedicated more time, is discussed in Chapters 4 and 5. Chapter 6 concludes the thesis.

3 Analysis: positioning accuracy of the current Wi-Fi fingerprint IPS algorithms

I started my internship activity studying the indoor positioning systems currently available on the market and in the literature. I found interesting and challenging the idea of re-implementing and comparing some among the current positioning algorithm described in the state of the art. This chapter presents the results and conclusions of this study. In particular, Section 3.1 presents an overview of the current positioning algorithm, and motivates my decision to re-implement one of the recent milestone articles about indoor positioning. Section 3.2 describes the dataset used for comparison and the experimental setup. Section 3.3 shows results and conclusions.

3.1 Comparison of the state-of-art Indoor Positioning Algorithms

3.1.1 Deep Neural Network with SAE

The article *"Low-Effort Place Recognition with WiFi Fingerprints Using Deep Learning"* [31], published by the researchers Michal Nowicki and Jan Wietrzykowski during the *International Conference Automation* in 2017, is important in the literature about Indoor Positioning Systems because it adopts the knowledge developed in machine learning and neural networks to develop an effective low-cost indoor positioning algorithm.

Motivation: why re-implementing a state-of-art positioning algorithm based on NN

Using neural networks in the implementation of indoor fingerprinting positioning algorithm is important because:

1. Provide satisfactory (best?) results in terms of accuracy performance. Using stacked autoencoders allows to efficiently reduce the feature space in order to achieve robust and precise classification [31], [32], [33];
2. Significantly lower the work-force burden of the localization system design. Neural Networks provide a global location recognition solution from WiFi data on a sparse map of scans, and at a significantly reduced effort for manual tuning [31].

NN Implementation

The authors propose a classification DNN with autoencoders to predict floor and building based on a single query scan. Stacked autoencoders (SAE) are parts of the deep network used to reduce the dimensionality of the input data by learning the reduced representation of the original data during unsupervised training [34]. The SAE is learned during unsupervised training and the goal is to train the pair encoder-decoder to achieve the same information at the output as it was provided as input. The input to SAE are signal strengths detected in a scan with one RSS-value for each network visible in the training database. The output of decoder is the reconstructed input from reduced representation. The used SAE (blue) with additional decoder part (red) is presented in Figure 3.
When the unsupervised learning of weights of SAE is finished, the decoder part of the network is disconnected and typical fully-connected layers of a deep network are connected to the output of the encoder, the classifier (as presented in Figure 4). Dropout is employed between hidden layers of the classifier, which randomly drops connections between layers during training to force the network to learn redundant representation and thus achieve better generalization and avoid overfitting. The

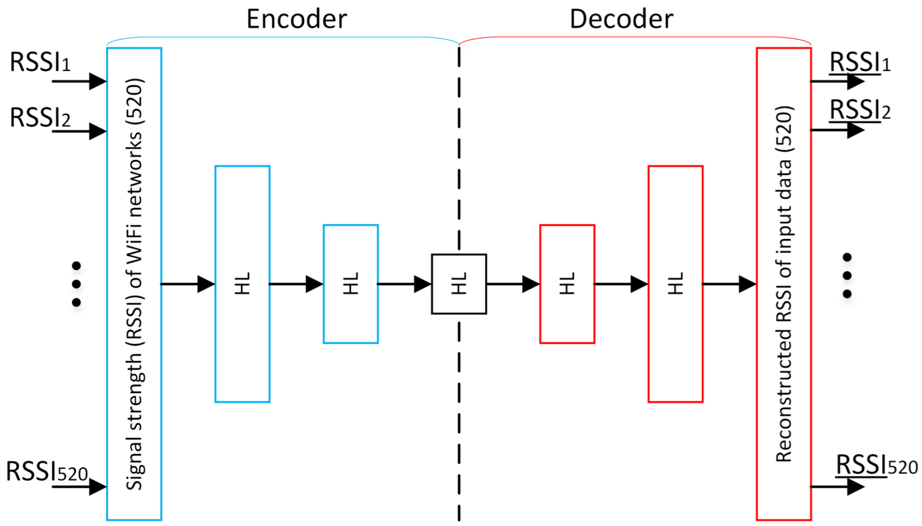


Figure 3: Stacked autoencoder (SAE) used in DNN to determine floor and building. The input to SAE are RSS-values detected in one scan. The output of decoder is the reconstructed input from reduced representation. The HL stands for hidden layer and the numbers in parentheses represent the number of neurons in the layer [31].

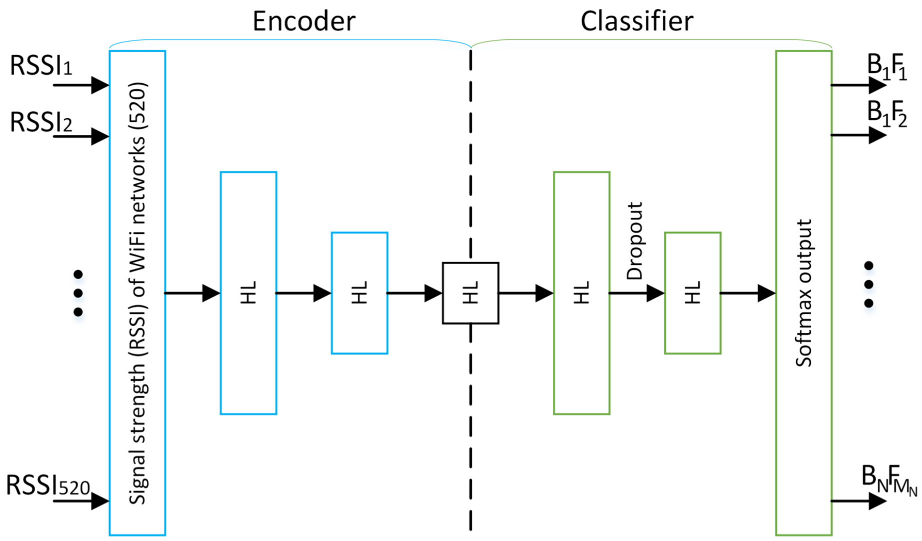


Figure 4: Architecture of DNN with SAE used to classify building and floor based on provided input WiFi scan. The already pre-trained encoder part is connected to classifier. The numbers in parentheses represent the number of neurons in the layer [31].

final output layer is a softmax layer that outputs the probabilities of a current sample belonging to analyzed classes.

3.1.2 Other classification algorithms employed for indoor positioning

Since the estimation of the target device can be obtained as the solution of a classification problem, classification algorithms that are well-known in the literature of machine learning can be employed for localization purposes [29].

K-Nearest Neighbors K-Nearest Neighbor (K-NN) [35], [36] is a distance based classifier that classifies instances based on their similarity. The unknown tuple in K-NN, i.e. the query scan QS, is assigned to most common reference point among its K-nearest neighbors. To do that, the algorithm computes the distances between the query point and all the reference points stored in the fingerprinting database. The distance values are then sorted in ascending order and only the top K distances are considered. An estimate of the position of the query point, in terms of room where the user is, is finally returned as the mode of the K-closest reference points, i.e., the mode of the K-nearest neighbors of the query point [37], [38], [39].

Random Forest Classifier Random forests are an example of ensemble learner built on decision trees [40]. A decision tree is a hierarchical structure including decision (non-terminal) nodes, branches and leaf (terminal) nodes that represent attributes (features), conditions and classes, respectively. Entropy or information gain can be used to create nodes in a decision tree. Each non-terminal node that contains a condition is used to determine which branch to follow from that node. If the condition is true then algorithm will follow one branch otherwise it will follow the other one. When the algorithm reaches a leaf node, then the label stored in the leaf returned as a class. An ensemble method is used based on the notion that multiple over-fitting estimators can be combined to reduce the effect of over-fitting in decision trees. Random forest makes use of an ensemble of parallel estimators, each of which over-fits the data, and averages the results to find a better classification [41], [42].

Support Vector Machine SVM [43], [44] is a classification technique that uses statistical learning theory. It looks for a linear optimal hyper plane so that the margin of separation between the classes is maximized. In practice, most data are not linearly separable; thus, to make the separation feasible, transformation is performed using a Kernel function. Data are mapped into a high dimensional input space, and SVM constructs an optimal separating hyperplane in this space [45].

Naïve Bayes Classifier A Naïve Bayes classifier [46] based on Bayes Theorem is a supervised learning algorithm. It calculates probability of each attribute in the data assuming that they are equally important and independent of each other. This assumption is called as class conditional independence. It is robust to noisy data, easy to build, shows good accuracy and high speed when applied to large databases and sometimes performs more complicated classification models. Hence, it is widely used in classification tasks [47].

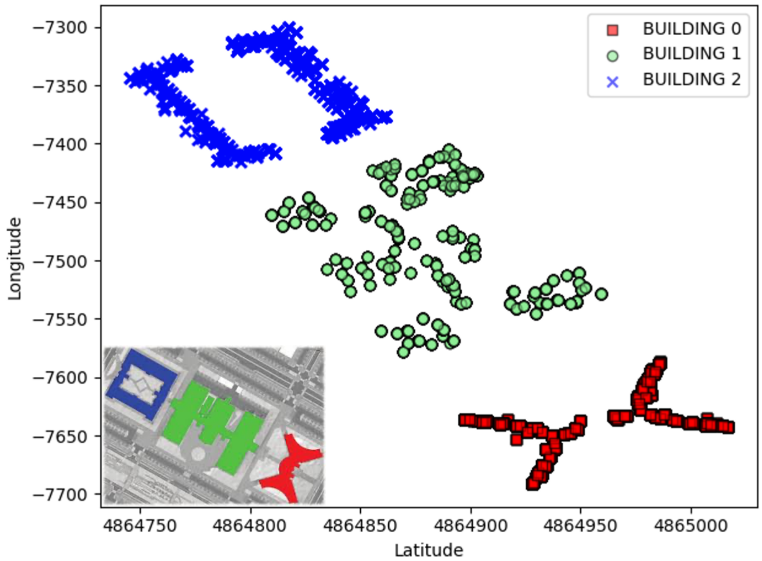


Figure 5: Scatter of data points contained in the UJIIndoorLoc dataset and shape comparison with a satellite photo. Bottom-left-corner: map of the Tx Buildings in the UJI Riu Sec Campus. Red refers to TI building, green corresponds to TD building and blue stands for TC building [48].

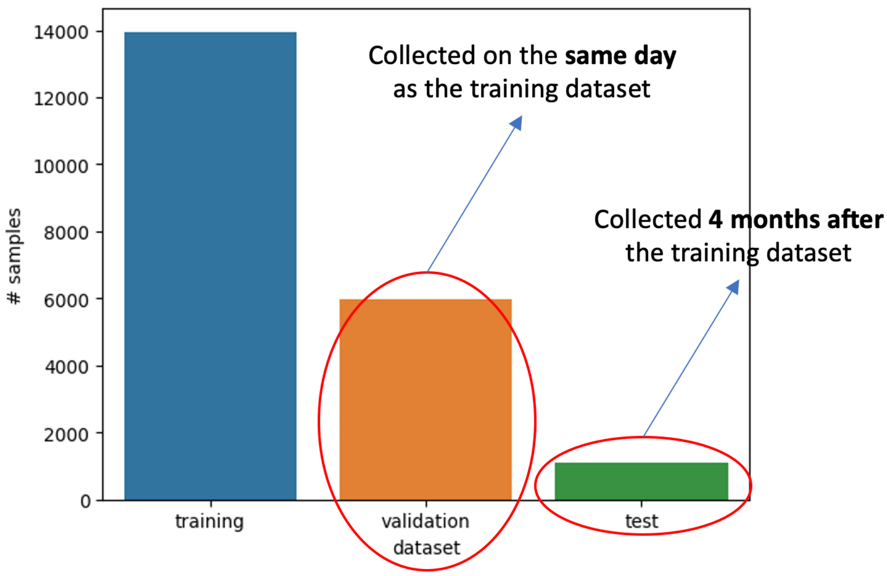


Figure 6: Frequency distribution of the number of samples split among training, validation and test set. The validation set contains samples collected on the same day as the training set. Samples in the test set have been collected 4 months later than the ones collected in the training set.

3.2 Performance evaluation

3.2.1 The UJIIndoorLoc dataset

To evaluate the implemented algorithm and allow comparison with traditional classification methods, I decided to use a publicly available dataset that contains Wi-Fi RSS-readings measured by smartphones and labeled positions. The UJIIndoorLoc dataset [48] contains Wi-Fi measurements used during EvAAL competition at IPIN Conference 2015 and is made publicly available. Positions are labelled from 3 buildings with 4 or more floors of the University of Jaume I in Spain at the area of almost 110000 square meters. The registered data comes from 25 different Android devices and was recorded with the help of 20 users. Figure 5 contains a scatter plot of the positions of all the samples contained in the UJIIndoorLoc dataset. The shape of the measurement points resembles the shape of the real buildings, depicted from a satellite photo in the bottom-left corner.
Each scan in the database contains 529 attributes. At the area of operation, 520 different APs were discovered and therefore the first 520 attributes inform about the received signal strength from those APs. The signal strengths vary from -104 dBm in a case of poor reception to almost 0 dBm. When AP is not available, the value of 100 is provided. The remaining 9 attributes contain information about longitude and latitude of measurement, floor number, building ID, space ID, relative position, user ID, phone ID and the timestamp of the measurement.
Figure 6 represents the frequency distribution of the measurements contained in the UJIIndoorLoc dataset. The dataset consists in total of 21048 Wi-Fi scans. The training set contains 13955 samples. In the dataset, there are two validation sets. The first, that I called validation dataset, contains 5982 scans collected on the same day as the scans in the training set. The second validation set, that I call test set, contains 1111 scans collected 4 months after the ones in the training set. Of the 520 APs detected, 312 are in common between validation and test set. 98 APs disappears after 4 months, and 55 new APs are detected in the test set. Therefore, the localization on data from UJIIndoorLoc is challenging, but the obtained results can be used to estimate the real-life performance of the system. The test set allows to evaluate how the localization algorithm performs when the fingerprints in the database are subject to time-aging.

3.2.2 Experimental results

I compared the indoor positioning algorithms presented in Section 3.1 on the common dataset described in Section 3.2.1. The experimental results are shown in Figure 7.
The implementation with neural networks and stacked autoencorders yields around 99% of correct recognitions on validation dataset, that contains samples collected on the same day as the training. This result is consistent with other classification algorithms such as K-Nearest Neighbors and Random Forest Classifier. However, with respect with the other algorithms, the proposed solution allowed to increase the correct recognition rate on testing samples collected 4 months later from 88% up to 93%. The recognition test performed on the testing samples yields 93% correct building and floor recognitions. The K-Nearest Neighbors approach, that finds the most similar scan in training samples based on Euclidean distance, resulted in only 88% correct recognitions. Random Forest Classifier and Support Vector Machine achieved results that are similar to the K-NN approach. Finally, Naive Bayes classifier only obtained around 86% of correct building and floor recognitions on validation test, and 81% on the test set.
Comparing to results obtain during EvAAL competition [33], the competing teams achieved results from around 87% to 96% percent, but using information from inertial sensors that are unavailable online. Nevertheless, my implementation achieves similar results, with lower cost in terms of data collection and computational effort.

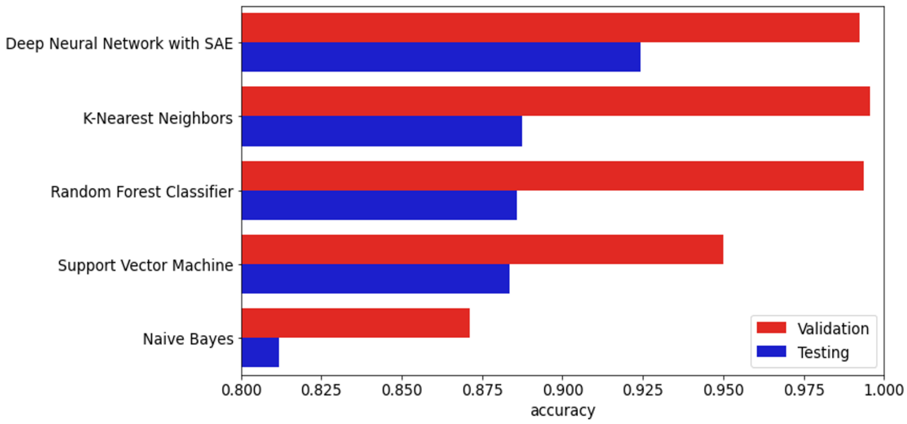


Figure 7: Comparison of correct recognition ratios for different classification algorithms evaluated in building and floor classification problem obtained for validation samples (collected on the same day as the training set) and testing samples (collected four months after the training set).

Team name	**Floor prediction**
MOSAIC	93.86%
HFTS	96.25%
RTLS@UM	93.74%
ICSL	86,93%
My implementation	***92.44%***

Figure 8: Comparison to the best results achieved by the four teams that participated in the EvAAL competition at International Conference on Indoor Positioning and Indoor Navigation (IPIN) 2015. The competing teams achieved floor hit rates in percentage of correct estimates from around 86% to 96%. Data is taken from: Moreira, Adriano, et al. "Wi-Fi fingerprinting in the real world-RTLS@ UM at the EvAAL competition." 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, 2015 [33].

3.3 Preliminary conclusions

Experimental results shows that my implementation with DNN and autoencoders achieves comparable results to state-of-the-art systems that require more data as input, such as measures from inertial sensors, and take more time to carefully develop and tune [33]. When measurements are fresh, e.g., collected on the same day as the samples in the database, the proposed implementation yields an accuracy that is nearly one. Compared to other implementations, the presented approach also perform better when the measurements in the fingerprinting database suffer from time-aging. The accuracy obtained when training the network on samples that are 4 months older than the query scan is about 93%. Time-aging inevitably occurs in everyday life, therefore is crucial for the positioning algorithm to still yield acceptable performance with time.

In conclusion, Neural Networks allow Wi-Fi localization algorithm to achieve very high accuracy with relatively low-effort. However, all positioning algorithms perform best when the samples in the training set are many and fresh. What is the cost to collect a lot of fresh samples?

4 Analysis: data collection cost in the current Indoor Positioning Algorithms

Positioning algorithms based on neural networks and, in general, all positioning algorithms, require a lot of fresh measurement data. Is then important to give an answer to the following questions:

1. How many samples do we need to collect in the data collection phase?
   Does more data means higher accuracy? If yes, in which cases?

2. What is the price to pay for data collection?
   Is data collection feasible with the modern technology?

The second part of my internship addressed these problems.

4.1 How many samples are needed to obtained the desired accuracy?

One major issue of fingerprint-based localization is that it is difficult to determine how much data is needed to obtain the desired accuracy during the offline phase [1]. Does the sampling of more fingerprint data provide better results?
Fortunately, the large size of the UJIIndoorLoc dataset [48] allows to address this problem. I set up an experiment that consists in down-sampling the initial dataset. Using the deep neural network implementation, I evaluated the accuracy of the localization system on the UJIIndoorLoc dataset while gradually reducing the size of the dataset. Figures 9 and 10 show the results in case of validation samples collected on the same day as the ones in the training set, and test samples collected four months after, respectively. The figures present on the x-axis the number of samples per class considered; on the y-axis, the accuracy performance. Starting from the initial dataset that considers around 1000 samples per class, I down-sampled the number of scans per building and floor to 500, 200, 100, 50, 20 and 10.
The results show that for large sizes of the dataset, i.e., considering 200 or more scans per location, the accuracy of the positioning algorithm reaches a plateau. The maximum accuracy for validation samples collected on the same day as the training is over 99%. When the test samples are collected four months later than the training samples, the plateau is at around 93%. The performance has a drop as soon as the number of samples per class considered goes under 200. For example, in the validation plot, the accuracy drops to around 70%, 50%, 48% when 50, 20, and 10 scans per location are considered. In case of the test plot, when samples in the fingerprinting database experience a time-aging of four months, the accuracy drops to 62%, 49%, 31%, respectively.
The conclusion of this study is that, in order to achieve a satisfactory performance in terms of accuracy, the number of scans needed for each location must be at least 200, or more. The answer to the initial question is affirmative: yes, up to 200 samples, more data measured and stored in the fingerprinting database means higher accuracy. Building a sufficiently large fingerprint database is vital to achieve high localization performance.

4.2 Data collection cost

During the data collection phase, to obtain an acceptable level of accuracy more than 200 scans per reference point are required. Up to 200 samples, more data means higher accuracy.
What is the price to pay for data collection? The data collection process is time-consuming and user labor-intensive. User labor-intensive in the sense that the user is manually demanded to label each reference point he visits. Time-consuming since Android OS impose strict limitations in terms

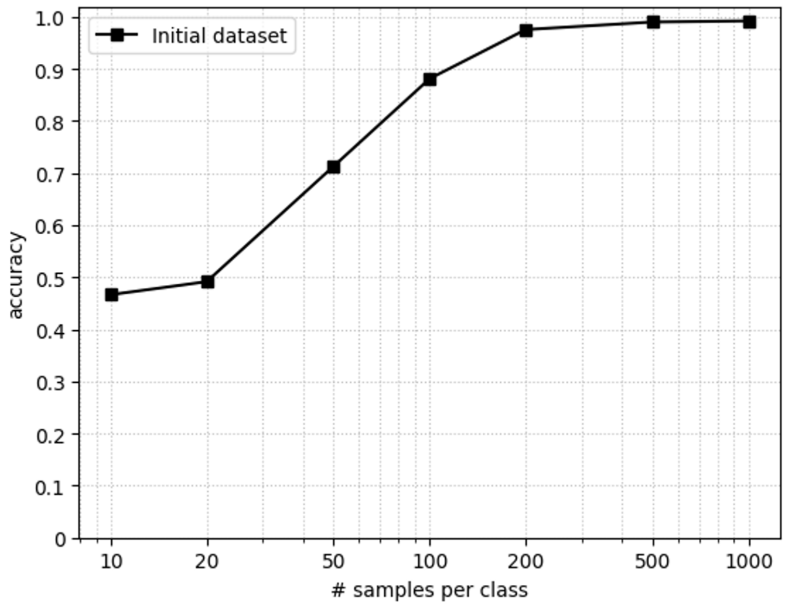


Figure 9: Localization accuracy for different sizes of the fingerprinting database on validation dataset collected on the same day as the training dataset.

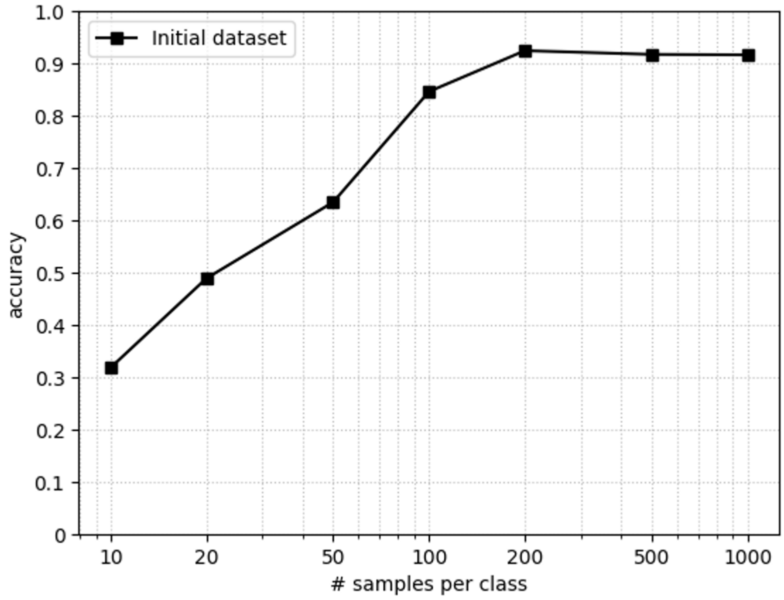


Figure 10: Localization accuracy for different sizes of the fingerprinting database on test dataset collected four months after the training dataset.

of frequency of scans. The official Android developers website specifies that throttling is imposed to the system call *WifiManager.startScan()* because of security issues related to the danger of remotely measuring multiples scans on the same device, without the user being aware. Quoting the Android developers website [49]:

"The following limitations apply to the frequency of scans using *WifiManager.startScan()*.

Android 8.0 and 8.1:

Each background app can scan one time in a 30-minute period.

Android 9:

Each foreground app can scan four times in a 2-minute period.
All background apps combined can scan one time in a 30-minute period.

Android 10 and higher:

The same throttling limits from Android 9 apply".

Is data collection feasible with the modern smartphones technology? Supposing to use Android 9 and requiring the user to use the app always in foreground, the allowed number of scans is two per minute. Considering that at least 200 scans are required for each location, the user would spend 100 minutes for each location in the indoor environment. To make things worse, asking the user to spend 100 minutes per location with the app in foreground is not convenient. If as system designers we want to measure scans in background, the frequency of scans allowed by Android is one time in 30 minutes, that is, 2 times per hour. Since 200 scans per location are required, the estimated time for the user to fill the fingerprinting database for one reference point is 100 hours.
In my opinion, this excessive time to perform fingerprint data collection affects the popularity and applications of fingerprint-based localization. On the basis of these considerations, the goal of my internship has become the evaluation and eventually proposition of technique to reduce the data collection cost.

4.3 How literature faces the problem of high data collection cost?

To reduce the cost of data collection, several methods have been proposed in the literature. In the state-of-the-art, there are basically two schools of thoughts.

(1) Use modelling to choose in a "clever" way which sample collect

This family of techniques aims at finding alternative methods to construct the fingerprint database with low human effort. The idea is to choose in a "clever" way which sample collect thus reducing the labeling cost of fingerprint collection.
In [50], the authors present a semi-supervised manifold learning technique for building a fingerprint database from partially labeled data, where only a small portion of the signal strength measurements must be marked with the corresponding identifier of the environment. Semi-supervised learning and active learning techniques allow to reduce the human effort for labeling the samples. Jun et al. [51] present the design, implementation and evaluation of AP-Sequence: a fingerprint-based localization system that achieves extremely low overhead in fingerprint map construction and maintenance.

A novel fingerprint collection technique is proposed in [52] that detects WiFi APs to form WiFi fingerprints from the signals collected by ZigBee interfaces. The authors of [53] propose a fingerprint-based device-free localization system named iUpdater to significantly reduce the labor cost and increase the accuracy. It is able to accurately update the whole database with RSS measurements at a small number of reference locations, thus reducing the human labor cost. In [54], the authors propose AcMu, an automatic and continuous radio map self-updating service for wireless indoor localization that exploits the static behaviors of mobile devices. They use stationary mobile devices as reference points to collect real-time RSS samples. In [55], the authors propose a novel method to construct a comprehensive fingerprint database by using the radio propagation model. The authors in [56] propose a Gaussian process regression for fingerprint-based localization that uses realistic and virtual indoor dynamic measurement data.

(2) Use modelling to recover fingerprints based on existing partial fingerprints

This second family of techniques aims at recover new fingerprints on the base of existing partial fingerprints. Methods as fingerprint interpolation and gradient based fingerprint generation are used, that do not require the user to scan for new fingerprints.
In [57] and [58], the authors propose a method based on compressive sensing to recover absent fingerprints. Their approach shows the hidden structure and redundancy characteristics of fingerprints in a merging matrix. The method in [59] leverages a more stable RSSI gradient to build a gradient-based fingerprint map by comparing the absolute RSSI values at nearby positions. In [60], an FM-based indoor localization system that does not require proactive site profiling is presented to construct the fingerprint database based solely on an estimate of indoor RSS distribution. Milioris et al. [61] use the Matrix completion framework to build complete training maps from a part of the reference fingerprints by learning the relevant fingerprint structure. The authors of [62] propose the Enriched Training Database (ETD), which is a web-service that enables the management and storage of training fingerprints and includes additional enriching functionality. The user can automatically generate virtual fingerprints based on propagation modeling of the virtual training points through the enriching functionality. The same authors also propose a new method to acquire training fingerprint locations that eliminates the burden of manually defining training points and covers areas with insufficient density to train fingerprints.

4.4 Disadvantages of the existing solutions

Despite many researchers proposed techniques for the construction of the fingerprint database to reduce human effort, no general solution exists in the literature.
In my opinion, the problems of the methods proposed in the state-of-the-art are:

1. The first family of techniques reduces the labeling cost, but still requires a huge human effort for the data collection in terms of time. These methods still assume the availability of a considerable number of samples (that would be unlabeled rather that labeled), reducing the user labor without actually reducing the time needed for data collection. Moreover, the presence of unlabeled samples strongly reduces the number of reference points in the database, and this dramatically decrease the accuracy performance.

2. The second family of techniques shows a very interesting approach. These methods try to generate new samples that integrate and enlarge the radio-map. However, they usually relies on propagation models or make use of mobile trajectories whose collection is not feasible with commercial off-the-shelf hardware.

5 Data Augmentation techniques for indoor localization

To truly expand the number of fingerprints in the database while reducing the human effort, I propose three methods that really increase the amount of training data collected at each reference point with no human effort and very low time expense.
For the design of this solutions, I took inspiration from the data augmentation techniques already developed by researchers for images. My work consisted in adapting image data augmentation existing techniques to the physics of the problem of collecting RSS-fingerprints.
Image data augmentation is a topic that been extensively treated by researchers in the last years [63]. Limited data is a major obstacle in applying deep learning models like deep convolutional neural networks. Often, imbalanced classes can be an additional hindrance: while there may be sufficient data for some classes, equally important, under-sampled classes may suffer from poor class-specific accuracy. There are many ways to address complications associated with limited data in machine learning. Image augmentation is a data-space solution to the problem of limited data [64]. It consists in useful techniques that can increase the size of the training set without acquiring new images. The idea is simple: duplicate images with some kind of variation so the model can learn from more examples. Ideally, we can augment the image in a way that preserves the features key to making predictions, but rearranges the pixels enough that it adds some noise. Data augmentation for images increases the diversity of the training set. A very simple example of data augmentation for images is image rotation. This phenomenon is intuitive. If the model learns from more examples of a given class, it is more likely to correctly predict that class. Data Augmentation encompasses a suite of techniques that enhance the size and quality of training dataset such that better Deep Learning models can be built using them. The image augmentation algorithms developed in the literature include geometric transformations, color space augmentations, kernel filters, mixing images, random erasing, feature space augmentation, adversarial training, generative adversarial networks, neural style transfer, and meta-learning [63]. But we must be careful: augmentation can be counterproductive if it produces images very dissimilar to what the model will be tested on, so this process must be executed with care.
In the next three sections, the three proposed techniques for data augmentation of Wi-Fi fingerprints will be presented. Finally, experiments and results will be provided.

5.1 WF-LERP: Wi-Fi Fingerprints generated through Linear Interpolation

Wi-Fi Fingerprintings generated through Linear Interpolation (WF-LERP) uses interpolation methods to reduce the time needed for radio map creation. The idea is to measure RSS-values in a few positions and then calculate the rest of the fingerprints by linear interpolation. Interpolation is a process of estimation of missing values from set of known values. Therefore, interpolation was taken into account as suitable solution to estimate fingerprints at additional reference points. The advantage of interpolation models, with respect to propagation models, is that they do not require the knowledge of the position of APs and obstacles, such as walls and furniture. As it is known, walls and other obstacles cause RSS fluctuation in the same position. Due to this, propagation models have poor performance in indoor environments. Interpolation methods can be used to overcome this problem. The principle behind WF-LERP is based on estimation of fingerprint between a pair of reference points. Interpolated fingerprints contain only record of APs appeared in both reference points. Final RSS has been calculated as weighted average:

$$RSS_{linear} = w_1 RSS_1 + w_2 RSS_2, \tag{6}$$

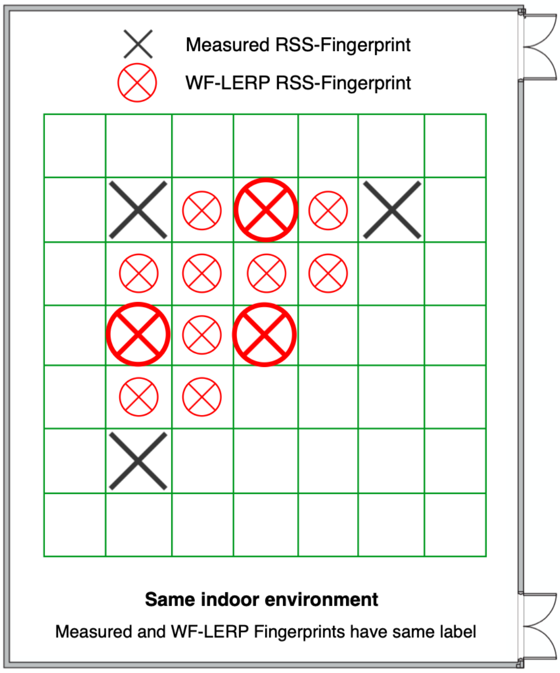


Figure 11: Illustration of the WF-LERP RSS-Fingerprint generation process. Starting from a few samples collected in the measurement area (black crosses), the database is enlarged with virtually generated samples (red crosses).

where RSS_{linear} represents the new generated RSS through WF-LERP, w_1 and w_2 denotes weights, and RSS_1, RSS_2 represents RSS-values measured at the two reference points.
Let's consider for example the indoor environment in Figure 11. In the environment, only three fingerprints in three reference points are known (black crosses). With the interpolation model, new fingerprints in unexplored positions are estimated. If we set $w_1 = w_2 = 0.5$, for example, the fingerprints indicated with big red crosses are generated. Opportunely adjusting the weights, more and more fingerprints can be generated. All these fingerprints will have the same label, since they belong to the same indoor environment, with no human effort for the user.

5.2 WF-DCGAN: Wi-Fi Fingerprints generated through Generative Adversarial Networks

Wi-Fi Fingerprint generated through Deep Convolutional Generative Adversarial Networks (WF-DCGAN) is a method to increase the amount of Wi-Fi Fingerprints collected at each location based on Deep Convolutional Generative Adversarial Networks (DCGANs). The method can be summarized in the following three steps:

1. Transform RSS data into feature vectors, and feature vectors into feature images;
2. Generate similar feature images with GANs;
3. Merge initial fingerprints with generated feature images.

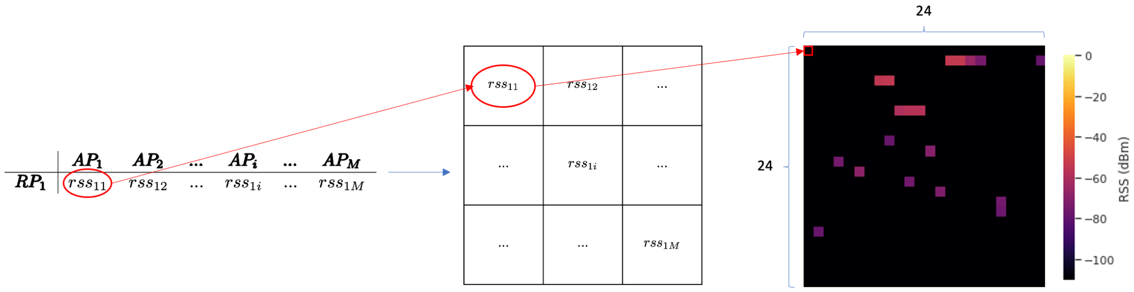


Figure 12: Mapping a feature vector into a feature image.

Map feature vectors into feature images

The idea behind WF-DCGAN is borrowed from image augmentation theory [63]. The first step is mapping RSS-scans into feature vectors, and then feature vectors into feature images.
The mapping process is explained in Figure 12. On the left, a Wi-Fi scan (vector of length M, where M is the number of APs detected) is mapped into an matrix of M pixels, with shape $\sqrt{M} \cdot \sqrt{M}$. Each pixel represents a RSS measurement, whose value is in the range $[0, -110]$ dBm. On the right, the matrix is represented as an image of shape 24×24 pixels. The right image in Figure 12 shows a real Wi-Fi fingerprint measured at the ground floor of the first building.

Generate similar feature images with GANs

Second step is to generate feature images that are similar to the ones stored in the fingerprinting database, using a Generative Adversarial Network (GAN) [65].
GAN is inspired by two-player game theory. The two players in the GAN model are the generative model G and discriminative model D [65]. G captures the distribution of sample data to generate a sample similar to the training data with added noise that obeys a certain distribution (uniform, in this case). Using this approach, the generated samples approximate real samples taken at the same location. D is a binary classifier that estimates the probability that a sample comes from the training data. If the sample comes from the real training data, D outputs a high probability; otherwise it outputs a small probability.
During training, one side is fixed, and the network weights at the other side are updated and alternately iterated. During the training process, both sides attempt to optimize their networks, forming a rivalry that continues until the two parties reach a dynamic balance (Nash Equilibrium). G restores the distribution of the training data and creates samples increasingly similar to the real data until D can no longer discriminate the results at an accuracy above 50%, at which point the discriminator and generator have reached Nash equilibrium. D and G play the following two-player minimax game with the value function $V(D, G)$ as follows:

$$\min_G \max_D V(D,G) = E_{x\sim P_{Data}(x)}\left\{\log D(x)\right\} + E_{z\sim P_Z(z)}\left\{\log\left(1 - D(G(z))\right)\right\} \tag{7}$$

where x is a sample taken from real data, z is a sample generated by the generator, $P_{Data}(x)$ is the distribution of the real data, $P_Z(z)$ is the distribution of the generator, $G(z)$ is the output of the generator, and $D(x)$ is the output of the discriminator.

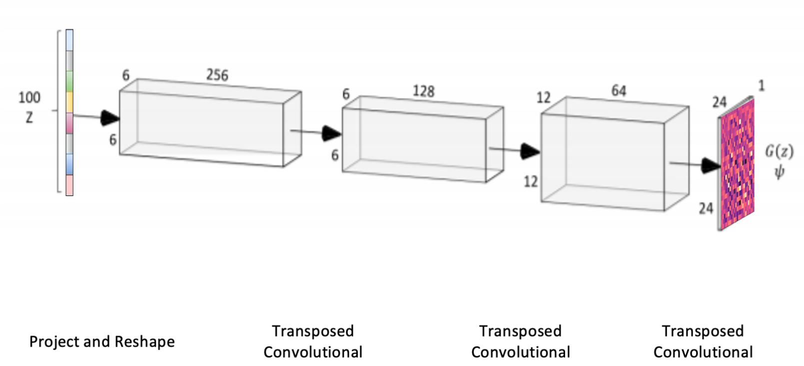


Figure 13: Generative model, implemented by a deconvolutional network.

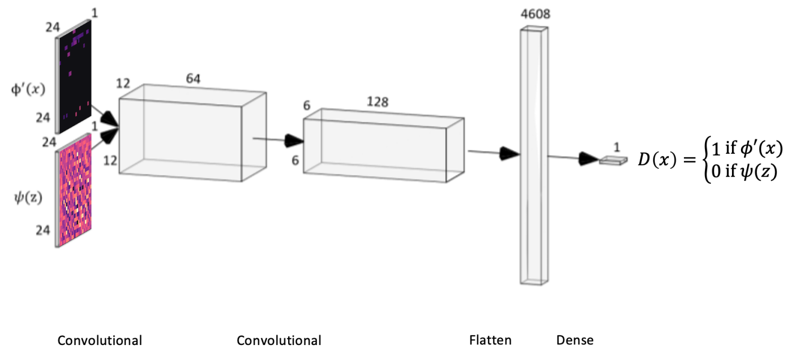


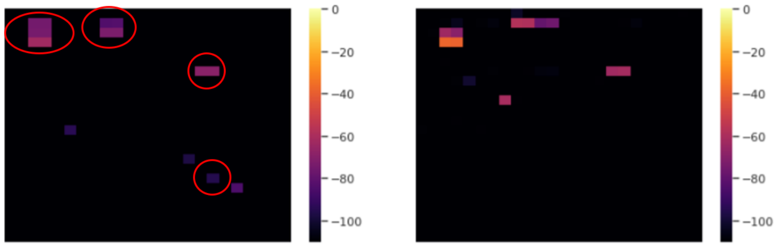
Figure 14: Discriminative model, implemented by a convolutional network.

DCGAN combines a CNN using supervised learning with a GAN using unsupervised learning [1]. I propose to extend the fingerprint database by using DCGAN to generate additional amplitude feature maps.

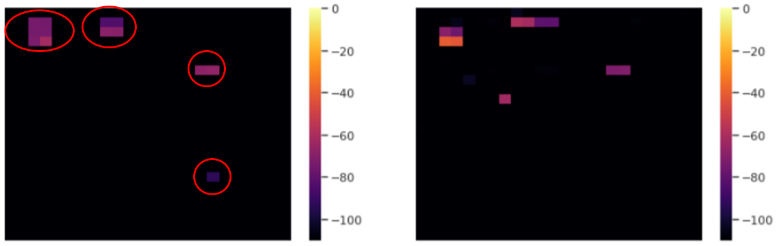
The network structure of the WF-DCGAN model is depicted in Figures 13 and 14. In Figure 13, G is the generative model, implemented by a deconvolutional network; Z is the signal, which has a uniform distribution; "transposed convolutional" represents the deconvolutional layer in the CNN model; Ψ denotes the feature maps generated by the deconvolutional layers in the generator. In Figure 14, D is the discriminative model, implemented by a convolutional network. Φ' are feature maps from the training set; Ψ denotes the generated samples; "convolutional" represents the convolution layer in the CNN model; $D(x)$ indicates the probability that the input sample is taken from the real training set (corresponding to x), or is generated by $G(z)$ as input.

Merge initial fingerprints with generated feature images

Once similar feature images have been generated through the DCGAN networks, they need to be mapped again into feature vectors and then into RSS readings. Finally, the generated RSS readings are merged into the initial fingerprint database, resulting in an expanded database. Figures 15 and 16 show examples of comparison between real and generated images representing Wi-Fi fingerprints measured in the same environments. From the figures, it is possible to observe that a certain degree

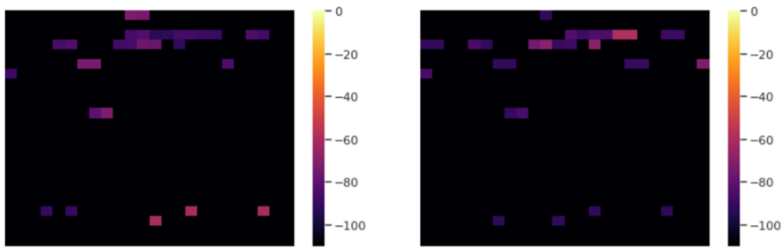


(a) Real images representing Wi-Fi fingerprints from class B0F0

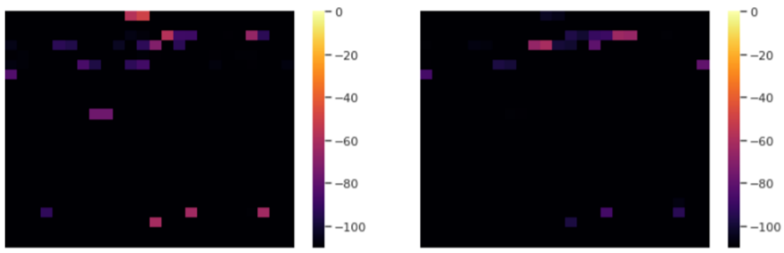


(b) Generated images representing Wi-Fi fingerprints from class B0F0

Figure 15: Comparison between real (a) and generated (b) images representing Wi-Fi fingerprints measured in the same environment, i.e., ground floor of first building.



(a) Real images representing Wi-Fi fingerprints from class B2F4



(b) Generated images representing Wi-Fi fingerprints from class B2F4

Figure 16: Comparison between real (a) and generated (b) images representing Wi-Fi fingerprints measured in the same environment, i.e., top floor of third building.

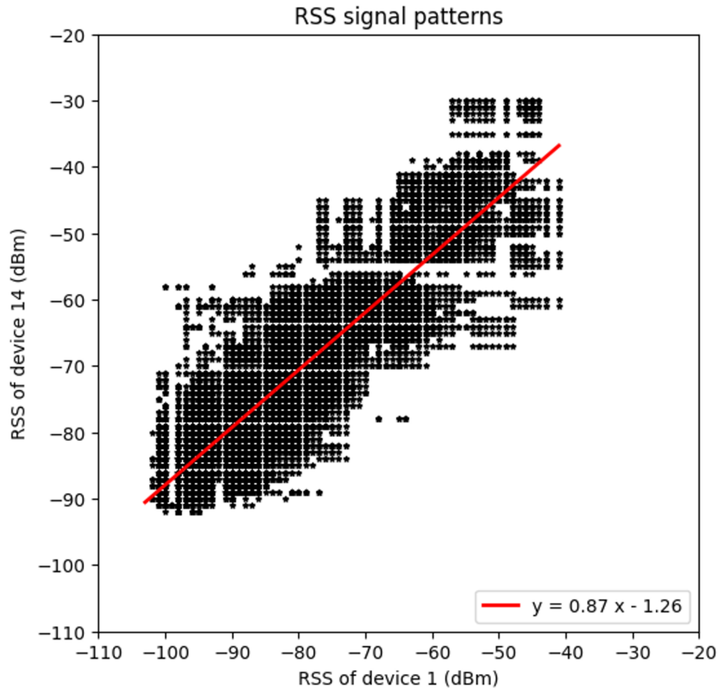


Figure 17: When multiple devices contribute in the fingerprint collection process, a RSS-reading error of ±10 dB is introduced.

of similarity exists between the real fingerprints and the ones generated through WF-DCGAN. However, a certain degree of diversity is added in the database that helps the model to better recognize the different classes.

5.3 WF-RND: Wi-Fi Fingerprints generated adding Uniform Random Noise

The third and last method proposed to reduce the time and human effort needed for radio map creation is once again borrowed from image augmentation techniques. A very simple example of data augmentation for images are geometric transformations, such as image rotation [63]. The idea is very intuitive: geometric transformations can be applied on the image, introducing diversity. The model will learn from different examples of the same class, thus resulting in an improved ability to recognize patterns.

A similar idea can be applied to RSS measurement arrays, with the difference that geometric transformations are not possible on arrays. However, it is always possible to perturb the values with a random noise. In case of RSS measurements, I have observed that RSS readings in the same location taken from different devices show difference in RSS values of around ±10 dB. Figure 17, that I obtained using the UJIIndoorLoc dataset, confirms the observation.

The WF-RND method takes into account this consideration, adding a uniform random noise in the range $(-10, 10)$ dB to all fingerprints in the database, thus generating new fingerprints. This method will be used as comparison to evaluate how well WF-LERP and WF-DCGAN perform compared to a random generation of samples.

$$S_{i+1} = \left[rss_{i,1} + U(-10,10), \cdots, rss_{i,j} + U(-10,10), \cdots, rss_{i,M} + U(-10,10)\right]. \quad (8)$$

5.4 Experiments and Results

Experiments are conducted to evaluate the performance of the proposed fingerprint database extension methods.

Experiment Methodology

The experiment is conducted using the UJIIndoorLoc dataset [48], described in Section 3.2.1. The dataset has been chosen since it is the largest collection of RSS measurements publicly available in the Internet. It contains samples from three building, each of them with four floors or more. Test RSS-values are collected four months after the ones used for training and validation. This characteristic enables studies on the localization performance when the Wi-Fi fingerprints suffer from time-aging. After four months, of the 465 APs measured, 98 APs were missing and 55 new APs were detected.
As common benchmark to evaluate the performance of the proposed fingerprint database extension methods, the deep neural network implementation with stacked auto-encoders [31] has been chosen as localization algorithm. Compared with other positioning algorithms, the neural network implementation achieves best accuracy performance both when the fingerprints are fresh, and when a four-months aging is considered.

Localization performance

WF-LERP. Figures 18 and 19 show the localization performance in terms of accuracy when the WF-LERP extension method is used, for different sizes of the training set. Once again, I down-sampled the training set by reducing the number of samples per class considered. The following figures present on the x-axis the number of samples per class that have been given as input to the deep neural network positioning algorithm; on the y-axis, the accuracy performance when the same positioning algorithm is applied.
Figure 18 shows the localization performance obtained with the WF-LERP fingerprint augmentation method, when the RSS-measurements in the validation dataset are collected on the same day as the RSS-scans in the training dataset. The black line shows a comparison with the performance achieved by the initial dataset, when no data augmentation technique is implemented. When more than 200 samples per class are considered, all the implementations converge to an unitary accuracy. In all the other cases, the WF-LERP data augmentation technique by far outperforms the situation when no data augmentation method is used. When only 10 samples per class are considered, the accuracy performance increases from 0.48 - when no data augmentation is implemented - to 1.0 - when the initial dataset is enlarged through WF-LERP by a factor 15. The accuracy performance increases of the 52%. Figure 19 considers the situation with the Wi-Fi fingerprints suffering a four-months time-aging. After four months, the plateau moved from 100% to approximately 92%. This result is expected, since in the four months the indoor environment has changed. Of the 465 APs measured, 98 APs have disappeared and 55 new APs have been detected. The interesting consideration is that, when only 10 samples per location are collected, the accuracy increases from 0.31 to 0.79, with a raise of the 48%.

WF-DCGAN. Figures 20 and 21 show the accuracy performance when the WF-DCGAN fingerprint augmentation method is applied, for different sizes of the training set.
In Figure 20, the case of fresh fingerprints is considered. When the number of samples per class is set to 200 or more, all the implementations yields an accuracy close to 100%. However, when less that 200 samples are considered, the WF-DCGAN data augmentation method outperforms the initial situation. In particular, in the case of 10 samples per class, the accuracy increases from 0.48 - no data augmentation - to 0.99 - augmentation of factor 20, with a growth of the 51%. The performance achieved is very similar to the one obtained using WF-LERP. Figure 21 represents the case with four-months time-aging. The plateau is still at around 0.92. However, when only 10 samples per location are collected, the accuracy increases from 0.31 to 0.63, of the 32% only.

WF-RND. The WF-RND data augmentation method has been implemented mainly with the purpose to be a benchmark for the comparison of the WF-LERP and WF-DCGAN techiques. Despite it only adds uniform random noise to the already existing fingerprints, the WF-RND method alone still outperforms the case when no data augmentation is used.
Figure 22 shows the accuracy performance achieved by the WF-RND method in case of fresh fingerprints. When only 10 scans per location are taken, a data augmentation of factor 30 brings the accuracy from 0.48 to 0.99, with an increase of the 51%. A factor 15 augmentation reaches at maximum accuracy 0.91, with a growth in the performance of the 43% with respect to the case when no augmentation is considered. In Figure 23, the WF-RND data augmentation method with a four-months fingerprint aging is represented. Still, with only 10 samples per location, the accuracy grows of the 45%.

Fresh fingerprints. Figures 24 and 25 compare the three fingerprint database extension methods, respectively WF-LERP, WF-DCGAN, and WF-RND, with the case when no database extension is applied, for the same augmentation factor. In the case when fingerprints are fresh, i.e., collected on same day as the ones used for training (Figure 24), WF-LERP performs best. In all cases, both WF-LERP and WF-DCGAN outperform WF-RND, showing that the fingerprints generated through the proposed methods add more information than a random augmentation. In some situations, e.g., for augmentation factors 2, 2.5, and 5, the accuracy performance obtained with WF-DCGAN can outperform the one obtained with WF-LERP.

Four-months old fingerprints. In the situation when fingerprints suffer a four-months aging (Figure 25), again WF-LERP perform best. WF-DCGAN, instead, only outperforms WF-RND when the augmentation factor is low (below 5). For high values of augmentation factor (greater than 5), the accuracy obtained with WF-RND is higher that the one achieved by DCGAN. This means that WF-DCGAN can generate samples that are very close to the real ones only when the fingerprints are fresh. Differently from WF-LERP, WF-DCGAN is not always able to generate samples that endure with time over a four-months time-aging.

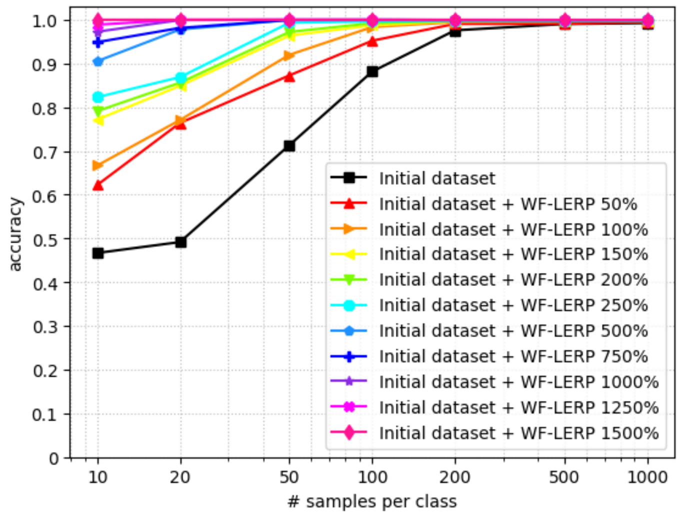


Figure 18: Accuracy performance achieved by WF-LERP fingerprint generation method for different sizes of training set, in presence of fresh fingerprints (i.e., the fingerprints in validation set have been collected on the same day as the ones used for training).

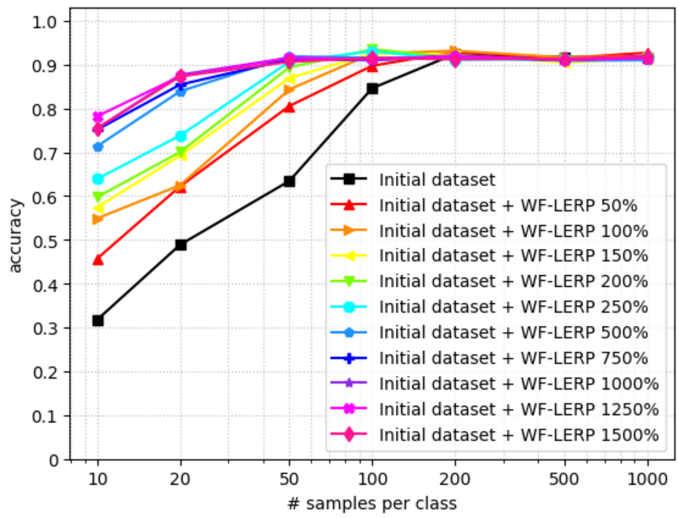


Figure 19: Accuracy performance achieved by WF-LERP fingerprint generation method for different sizes of training set, in presence of fingerprints that suffer from time-aging (i.e., the fingerprints in validation set have been collected 4 months later than the ones used for training).

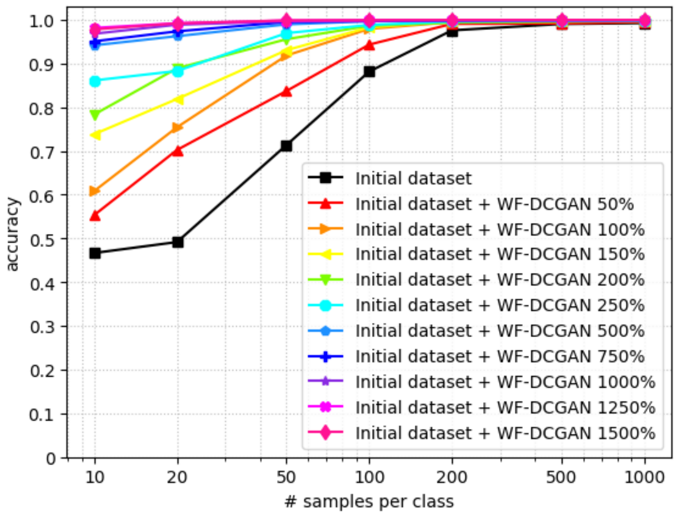


Figure 20: Accuracy performance achieved by WF-DCGAN fingerprint generation method for different sizes of training set, in presence of fresh fingerprints (i.e., the fingerprints in validation set have been collected on the same day as the ones used for training).

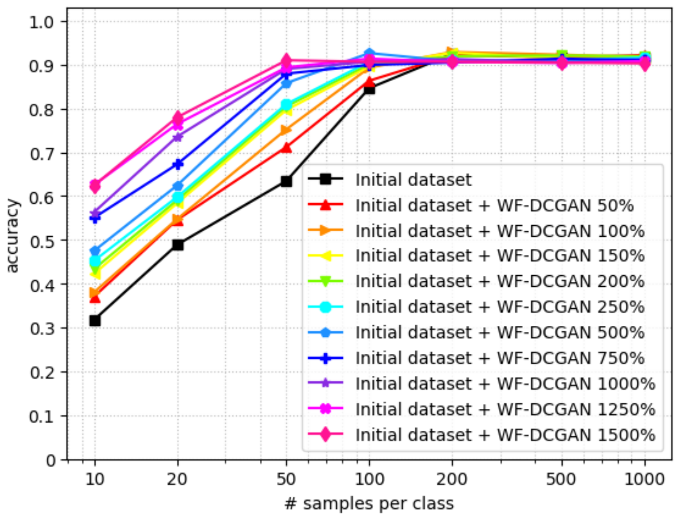


Figure 21: Accuracy performance achieved by WF-DCGAN fingerprint generation method for different sizes of training set, in presence of fingerprints that suffer from time-aging (i.e., the fingerprints in validation set have been collected 4 months later than the ones used for training).

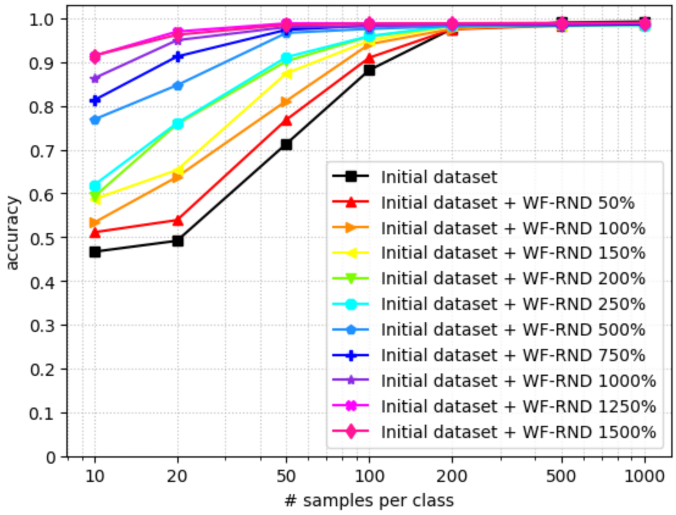


Figure 22: Accuracy performance achieved by WF-RND fingerprint generation method for different sizes of training set, in presence of fresh fingerprints (i.e., the fingerprints in validation set have been collected on the same day as the ones used for training).

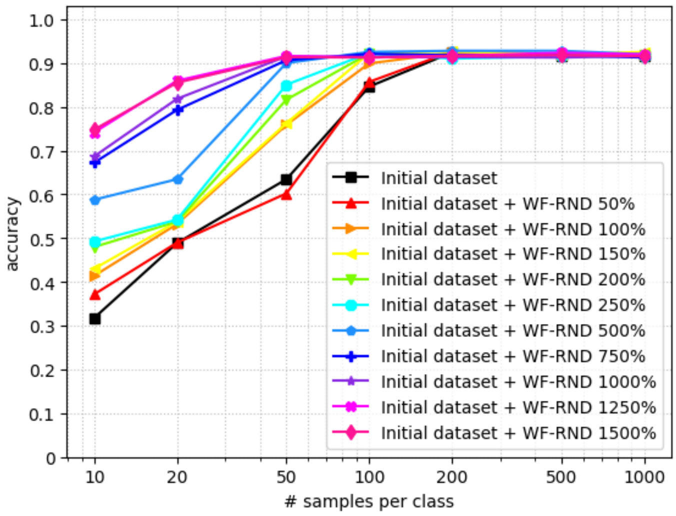


Figure 23: Accuracy performance achieved by WF-RND fingerprint generation method for different sizes of training set, in presence of fingerprints that suffer from time-aging (i.e., the fingerprints in validation set have been collected 4 months later than the ones used for training).

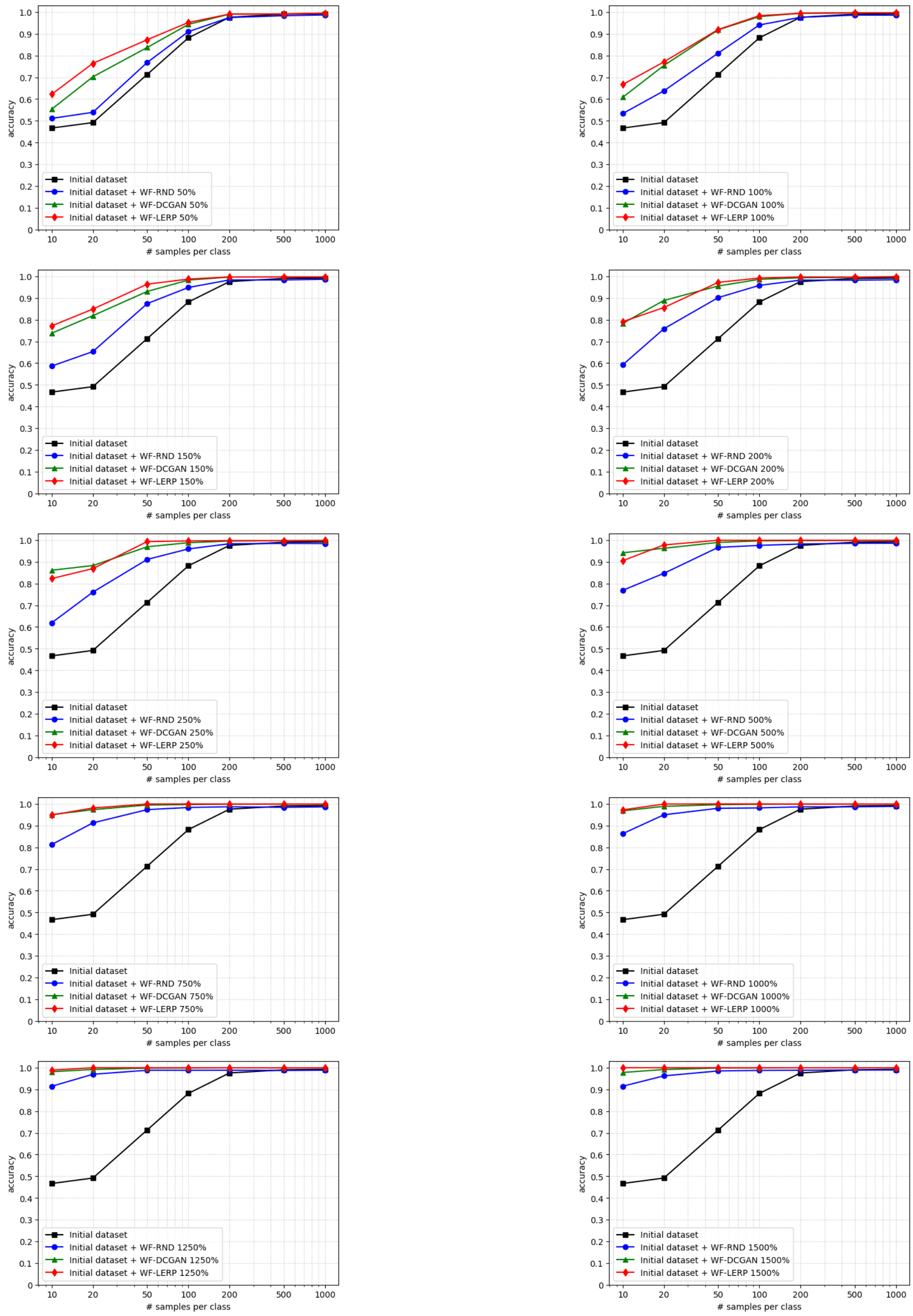


Figure 24: Accuracy performance achieved by the three fingerprint database extension methods proposed, respectively WF-LERP, WF-DCGAN, and WF-RND, and by the initial dataset, for the same augmentation factor, when the fingerprints are fresh.

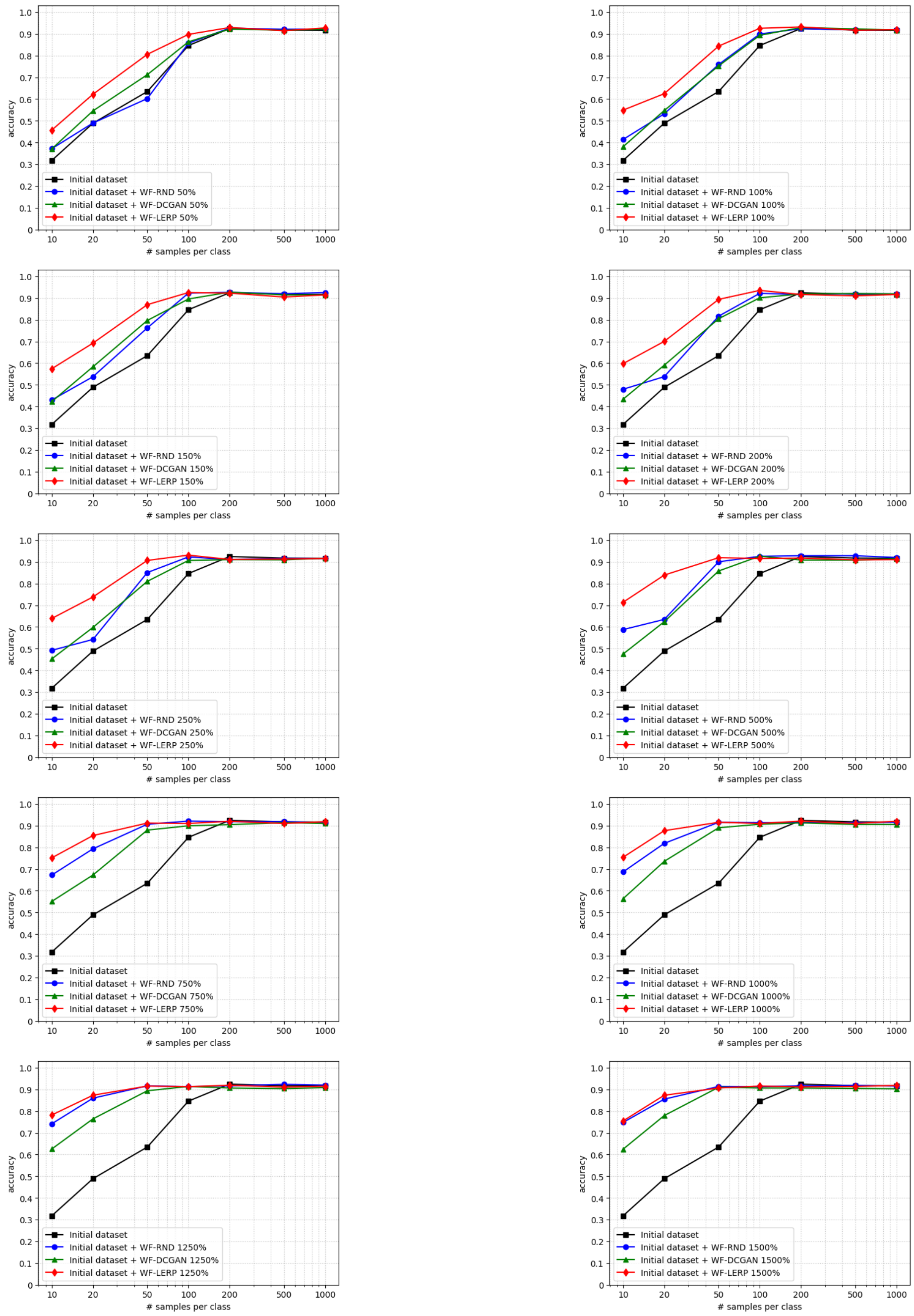


Figure 25: Accuracy performance achieved by the three fingerprint database extension methods proposed, respectively WF-LERP, WF-DCGAN, and WF-RND, and by the initial dataset, for the same augmentation factor, when the fingerprints suffers a 4-months time-aging.

6 Conclusions and Future Work

In my internship, I addressed the problem of reducing the cost and human effort of the fingerprint collection phase in Wi-Fi indoor positioning systems.

Time-schedule of the internship

I worked on this research project for approximately six months, from March, 16-th to August, 31-st. I scheduled the tasks and activities in the following way:

1. I dedicated months 1 and 2 on the acknowledgement of the state of the art about Indoor Positioning Systems. However, this activity took more time than expected, so I continued it until month 4 of my internship;

2. I thought that a better understanding of the state of the art was possible with the re-implementation of some among the milestone papers about indoor positioning. Hence, in parallel with the study of the literature, I worked on re-implementing the most famous Wi-Fi fingerprinting techniques;

3. In month 3, I worked on the analysis of the problems of current data-driven fingerprinting techniques. I identified the high cost of the data collection process as a major problem in the state-of-the-art fingerprinting algorithms, and provided rational motivations about my thesis;

4. In months 4 and 5, I took inspiration from the image augmentation field, and I worked on the design of three techniques for fingerprints generation, with the aim of enlarging the radio map at very low cost for the user. After the implementation, time has been spent on the experimental setup and on the results computation;

5. I dedicated full month 6 for the preparation of the oral defense, that resulted with the production of a presentation, and for the writing of the final report.

Technical difficulties encountered

At the end of my internship, it is time to take stocks of my first research activity.
My first steps in the research world have been quite unusual: I started my internship in quarantine, and frequented the laboratory only in the last month, with a very reduced number of colleagues. I have encountered some difficulties because of this situation, that I faced with the help of professors of my team. In the end, I consider my first research experience very formative.
At the end of my experience, I can say that I improved my abilities of independent working and problem solving. My internship helped me to open my mind on problems identification and on solutions thinking. Research is not an exact science: no universal solutions exist, and sometimes, after an hard work, results are not always as expected. From this point of view, research is very similar to the real life. This is the reason why, after this research experience, I feel grown as a person.
From the technical point of view, my internship has been challenging because I moved my first steps in the design and implementation of neural networks, using the Python programming language, and the open libraries Keras and TensorFlow. To learn the basics, I studied on books and followed online classes and tutorials.

Conclusions and future work

Fingerprint data augmentation techniques aimed at reducing the data collection cost in Wi-Fi RSS-based Indoor Positioning Systems are effective. Since the new fingerprints are virtually generated from the existing ones, the proposed techniques are able to cut-off the cost of data collection in terms of time and user labor. Among the proposed implementations, WF-LERP method achieves best performance if compared to WF-DCGAN and WF-RND.

Fresh fingerprints. When the scans in the validation database have been collected in the same day as the scans used for the construction of the fingerprinting database, WF-LERP shows an accuracy of 100% in terms of predictions of building and floor where the validation scans have been taken, compared to the initial 48% achieved by the localization algorithm with no data augmentation. To reach the maximum value of accuracy, the initial dataset needed the collection of more than 200 samples per location, requiring a time between 100 minutes and 100 horus. With WF-LERP data augmentation method, the same accuracy performance is achieved with only 10 samples per location, whose collection process requires a time between 5 minutes and 5 hours. Thus, the proposed method reduces the data collection time of a factor 20, or, under the same number of scans per location, increases the accuracy of the positioning algorithm of the 52%. When the fingerprints are fresh, WF-DCGAN performs as well as WF-LERP. Both proposed solution outperform WF-RND. This proves that the proposed methods are able to virtually generate new fingerprints that are similar to the real ones and that can be used to enlarge the fingerprinting database.

Fingerprints that suffer from time-aging. When the scans in the validation database have been collected four months later than the scans used for the construction of the fingerprinting database, WF-LERP still outperforms WF-DCGAN and WF-RND, but is not able to reach the same accuracy as in the case without time-aging of the fingerprinting database. This can be explained considering that during the four months the indoor environment had changed: of the 465 APs measured, 98 APs were missing and 55 new APs were detected. However, WF-LERP is still able to generate fingerprints that are valid after four months, achieving an accuracy of around 80% when only 10 measurements per location are taken, and an accuracy of around 93% when 50 or more scans per location are considered. Compared to the case with no data augmentation, the accuracy increases of the 48%. In presence of time-aging, it may happen that WF-RND outperforms WF-DCGAN. After four months, WF-DCGAN was not able to virtually generate fingerprints that are plausible if compared to the real ones: WF-LERP always outperforms WF-DCGAN. Time-robustness is a very important property for the localization algorithm in real life applications. A positioning algorithm that is time robust does not require the user to continuously update the fingerprinting database. The operation of updating the fingerprinting database can be very annoying from the user point of view and should be demanded as few as possible.

Final considerations In conclusion of this work I can state that, among the fingerprint augmentation method proposed, WF-LERP presents a very low computational cost (as easy as a simple linear interpolation implementation), the best accuracy performance (100% accuracy with only 10 fingerprints collected per location), and the best results in terms of robustness over time (with only 10 samples, the accuracy performance after 4 months is still around 80%). WF-DCGAN performs as well as WF-LERP when the fingerprints are fresh, but shows a performance drop in presence of time-aging, and requires a more complex implementation and parameter tuning. As future work, the challenge is understanding if indoor localization is possible without labels information. These developments would enable new localization features for ElectroSmart [66].

References

[1] Li, Qiyue, et al. "AF-DCGAN: Amplitude feature deep convolutional GAN for fingerprint construction in indoor localization systems." IEEE Transactions on Emerging Topics in Computational Intelligence. IEEE, 2019.

[2] Berkovich, Gennady. "Accurate and reliable real-time indoor positioning on commercial smartphones." 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, 2014.

[3] Farrell, J., and M. Barth. "The global positioning system." New York: McGrawHill Professional, 1999.

[4] Rao, Bharat, and Louis Minakakis. "Evolution of mobile location-based services." Communications of the ACM 46.12, 2003.

[5] Sharp, Ian, and Kegen Yu. Wireless Positioning: Principles and Practice. Springer, 2019.

[6] Bahl, Paramvir, and Venkata N. Padmanabhan. "RADAR: An in-building RF-based user location and tracking system." Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Vol. 2. IEEE, 2000.

[7] Khudhair, Ahmed Azeez, et al. "Wireless indoor localization systems and techniques: survey and comparative study." Indonesian Journal of Electrical Engineering and Computer Science, 2016.

[8] Al Nuaimi, Klaithem, and Hesham Kamel. "A survey of indoor positioning systems and algorithms." 2011 international conference on innovations in information technology. IEEE, 2011.

[9] Mautz, Rainer. "Indoor positioning technologies." PhD Thesis, 2012.

[10] Liu, Hui, et al. "Survey of wireless indoor positioning techniques and systems." IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2007.

[11] Gu, Yanying, Anthony Lo, and Ignas Niemegeers. "A survey of indoor positioning systems for wireless personal networks." IEEE Communications surveys and tutorials, 2009.

[12] Parkinson, Bradford W., et al., eds. Global positioning system: Theory and applications, Volume II. American Institute of Aeronautics and Astronautics, 1996.

[13] Hofmann-Wellenhof, Bernhard, Herbert Lichtenegger, and James Collins. Global positioning system: theory and practice. Springer Science and Business Media, 2012.

[14] Alarifi, Abdulrahman, et al. "Ultra wideband indoor positioning technologies: Analysis and recent advances". Sensors, 2016.

[15] Hightower, Jeffrey, and Gaetano Borriello. "Location sensing techniques." IEEE Computer, 2001.

[16] Svalastog, Mari Saua. Indoor positioning-technologies, services and architectures. PhD thesis, 2007.

[17] Song, Zhenlong, Gangyi Jiang, and Chao Huang. "A survey on indoor positioning technologies." International Conference on Theoretical and Mathematical Foundations of Computer Science. Springer, 2011.

[18] Gorostiza, Ernesto Martín, et al. "Infrared sensor system for mobile-robot positioning in intelligent spaces." Sensors, 2011.

[19] Brassart, Eric, Claude Pegard, and Mustapha Mouaddib. "Localization using infrared beacons." Robotica, 2000.

[20] Aitenbichler, Erwin, and Max Muhlhauser. "An IR local positioning system for smart items and devices." 23rd International Conference on Distributed Computing Systems Workshops. IEEE, 2003.

[21] Beauregard, Stephane, and Harald Haas. "Pedestrian dead reckoning: A basis for personal positioning." Proceedings of the 3rd Workshop on Positioning, Navigation and Communication, 2006.

[22] Ivanov, Rosen. "Indoor navigation system for visually impaired." Proceedings of the 11th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing on International Conference on Computer Systems and Technologies, 2010.

[23] Mautz, Rainer, and Sebastian Tilch. "Survey of optical indoor positioning systems." 2011 international conference on indoor positioning and indoor navigation. IEEE, 2011.

[24] Roshanaei, Mahnaz, and Mina Maleki. "Dynamic-KNN: A novel locating method in WLAN based on Angle of Arrival." 2009 IEEE Symposium on Industrial Electronics and Applications. IEEE, 2009.

[25] Wang, Xin, Zongxin Wang, and Bob O'Dea. "A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation." IEEE Transactions on Vehicular Technology, 2003.

[26] Yamasaki, Ryota, et al. "TDOA location system for IEEE 802.11 b WLAN." IEEE Wireless Communications and Networking Conference, 2005. IEEE, 2005.

[27] Izquierdo, Fernán, et al. "Performance evaluation of a TOA-based trilateration method to locate terminals in WLAN." 2006 1st International Symposium on Wireless Pervasive Computing. IEEE, 2006.

[28] Perez-Vega, Constantino, and José Miguel López Higuera. "A simple and efficient model for indoor path-loss prediction." Measurement science and technology, 1997.

[29] Xia, Shixiong, et al. "Indoor fingerprint positioning based on Wi-Fi: An overview." ISPRS International Journal of Geo-Information, 2017.

[30] Batistić, Luka, and Mladen Tomic. "Overview of indoor positioning system technologies." 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2018.

[31] Nowicki, Michał, and Jan Wietrzykowski. "Low-effort place recognition with WiFi fingerprints using deep learning." International Conference Automation. Springer, 2017.

[32] Luo, Junhai, and Huanbin Gao. "Deep belief networks for fingerprinting indoor localization using ultrawideband technology." International Journal of Distributed Sensor Networks, 2016.

[33] Moreira, Adriano, et al. ”Wi-Fi fingerprinting in the real world-RTLS@ UM at the EvAAL competition.” 2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN). IEEE, 2015.

[34] Vincent, Pascal, et al. ”Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion.” Journal of machine learning research, 2010.

[35] Peterson, Leif E. ”K-nearest neighbor.” Scholarpedia, 2009.

[36] Aha, David W., Dennis Kibler, and Marc K. Albert. ”Instance-based learning algorithms.” Machine learning, 1991.

[37] Li, Dong, Baoxian Zhang, and Cheng Li. ”A feature-scaling-based k-nearest neighbor algorithm for indoor positioning systems.” IEEE Internet of Things Journal, 2015.

[38] Li, Changgeng, Zhengyang Qiu, and Changtong Liu. ”An improved weighted k-nearest neighbor algorithm for indoor positioning.” Wireless Personal Communications, 2017.

[39] Shin, Beomju, et al. ”Enhanced weighted K-nearest neighbor algorithm for indoor Wi-Fi positioning systems.” 2012 8th International Conference on Computing Technology and Information Management (NCM and ICNIT). IEEE, 2012.

[40] Liaw, Andy, and Matthew Wiener. ”Classification and regression by randomForest.” R news, 2002.

[41] Calderoni, Luca, et al. ”Indoor localization in a hospital environment using random forest classifiers.” Expert Systems with Applications, 2015.

[42] Wang, Yanzhao, et al. ”WiFi indoor localization with CSI fingerprinting-based random forest.” Sensors, 2018.

[43] Suykens, Johan AK, and Joos Vandewalle. ”Least squares support vector machine classifiers.” Neural processing letters, 1999.

[44] Cauwenberghs, Gert, and Tomaso Poggio. ”Incremental and decremental support vector machine learning.” Advances in neural information processing systems, 2001.

[45] Xu, He, et al. ”An RFID indoor positioning algorithm based on support vector regression.” Sensors, 2018.

[46] Rish, Irina. ”An empirical study of the naive Bayes classifier.” IJCAI 2001 workshop on empirical methods in artificial intelligence, 2001.

[47] Bozkurt, Sinem, et al. ”A comparative study on machine learning algorithms for indoor positioning.” 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA). IEEE, 2015.

[48] Torres-Sospedra, Joaquín, et al. ”UJIIndoorLoc: A new multi-building and multi-floor database for WLAN fingerprint-based indoor localization problems.” 2014 international conference on indoor positioning and indoor navigation (IPIN). IEEE, 2014.

[49] Android Developers Docs. Https://developer.android.com/guide/topics/connectivity/wifi-scan

[50] Pulkkinen, Teemu, Teemu Roos, and Petri Myllymäki. "Semi-supervised learning for wlan positioning." International Conference on Artificial Neural Networks. Springer, 2011.

[51] Jun, Junghyun, et al. "Low-overhead WiFi fingerprinting." IEEE Transactions on Mobile Computing, 2017.

[52] Niu, Jianwei, et al. "ZIL: An energy-efficient indoor localization system using ZigBee radio to detect WiFi fingerprints." IEEE Journal on Selected Areas in Communications, 2015.

[53] Chang, Liqiong, et al. "iUpdater: Low cost RSS fingerprints updating for device-free localization." 2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2017.

[54] Wu, Chenshu, Zheng Yang, and Chaowei Xiao. "Automatic radio map adaptation for indoor localization using smartphones." IEEE Transactions on Mobile Computing, 2017.

[55] He, Chunrong, et al. "A novel radio map construction method to reduce collection effort for indoor localization." Measurement, 2016.

[56] Cho, Youngsu, et al. "GPR based Wi-Fi radio map construction from real/virtual indoor dynamic surveying data." 2013 13th International Conference on Control, Automation and Systems (ICCAS 2013). IEEE, 2013.

[57] Gu, Zhuan, et al. "Reducing fingerprint collection for indoor localization." Computer Communications, 2016.

[58] Sung, Chang Kyung, et al. "Interference mitigation based on Bayesian compressive sensing for wireless localization systems in unlicensed band." IEEE Transactions on Vehicular Technology, 2017.

[59] Shu, Yuanchao, et al. "Gradient-based fingerprinting for indoor localization and tracking." IEEE Transactions on Industrial Electronics, 2015.

[60] Yoon, Sungro, et al. "ACMI: FM-based indoor localization via autonomous fingerprinting." IEEE Transactions on Mobile Computing, 2015.

[61] Milioris, Dimitris, Milan Bradonjić, and Paul Mühlethaler. "Building complete training maps for indoor location estimation." Conference on Computer Communications. IEEE, 2015.

[62] Lemic, Filip, et al. "Enriched Training Database for improving the WiFi RSSI-based indoor fingerprinting performance." 2016 13th IEEE Annual Consumer Communications and Networking Conference (CCNC). IEEE, 2016.

[63] Shorten, Connor, and Taghi M. Khoshgoftaar. "A survey on image data augmentation for deep learning." Journal of Big Data, 2019.

[64] Kostrikov, Ilya, Denis Yarats, and Rob Fergus. "Image augmentation is all you need: Regularizing deep reinforcement learning from pixels." arXiv preprint, 2020.

[65] Goodfellow, Ian J., et al. "Generative adversarial networks." Advances in neural information processing systems, 2014.

[66] ElectroSmart EMF Detector, https://electrosmart.app