HUMAN SUBJECT IDENTIFICATION AND INDOOR POSITIONING VIA PASSIVE SPECTRUM MONITORING

by

HUAIZHENG MU

A thesis submitted in partial fulfillment of the
requirements for the degree of

ELECTRICAL AND COMPUTER ENGINEERING

2021

Oakland University
Rochester, Michigan

Thesis Advisory Committee:

Jia Li, Ph.D., Chair
Manohar Das, Ph.D.
Khalid Mirza, Ph.D.

To my mother and father

ACKNOWLEDGMENTS

The two years of graduate study at the Oakland University is an unforgettable time for me. The knowledge I acquired and the ability to solve problems is vital to my life.

I would like to express my gratefulness to those who helped me to pursue my academic and research goals. First, I would like to extend my sincere gratitude to my mentor, Dr. Jia Li for her careful guidance of my graduate life in the past two years, which greatly improved my understanding of academics and taught me a lot of knowledge and specific research skills. Dr. Li has a wealth of knowledge in the field of signal processing, machine learning, and mathematics. Every discussion has benefited me a lot. In addition, I am deeply moved by Dr. Li's academic rigorous attitude and patient guidance towards every student.

I would like to acknowledge my thesis committee members Dr. Manohar Das and Dr. Khalid Mirza for their suggestions during my master study and research.

Last my thanks would go to my beloved family for their loving considerations and great confidence in me all through these years. I also owe my sincere gratitude to my friends and my classmates who gave me their help and time in listening to me and helping me work out my problems during the difficult course of the thesis.

This research is supported by AFOSR grant FA9550-18-1-0287.

Huaizheng Mu

ABSTRACT

HUMAN SUBJECT IDENTIFICATION AND INDOOR LOCALIZATION VIA PASSIVE SPECTRUM MONITORING

by

HUAIZHENG MU

Adviser: Professor Jia Li, Ph.D.

Human subject identification and indoor positioning have great development and wide applications. The existing identification and indoor positioning methods have their specific limitations. This thesis proposes human subject identification and indoor positioning approaches, which utilize passive radio frequency (P-RF) signal as a biometrics modality to achieve identification and indoor positioning. It is verified that different human subjects could generate different spectrum signatures, and human subject at different locations can disturb the spectrum differently. And these spectrum characteristics can be distinguished and estimated by machine learning (ML) algorithms to realize human subject's identification and indoor positioning. The passive spectrum of the frequency bands that are sensitive to human occupancy are acquired via software-defined radio (SDR). Different machine learning algorithms have been investigated to attain the best identification and positioning accuracies. Experimental results from seven volunteers indicate the classification accuracy is higher than 94% using k-nearest neighbors (KNN) algorithms. The positioning experiments show a positioning error as low as 0.8m among all tested scenarios using Gaussian process regression (GPR).

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
ABSTRACT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
CHAPTER ONE INTRODUCTION	1
1.1 Human Sensing	1
1.2 Human Subject Identification	2
1.3 Human Subject Indoor Positioning	4
1.4 Thesis Contributions	6
1.5 Thesis Outline	7
CHAPTER TWO LITERATURE REVIEW	9
2.1 Passive Radio Spectrum Applications	9
2.2 Existing Identification Techniques	14
2.3 Existing Indoor Localization Techniques	16
2.4 Software Defined Radio	17
2.5 Machine Learning Algorithms	18
2.6 Discussions	19

CHAPTER THREE
HUMAN SUBJECT IDENTIFICATION 22

3.1 Overview 22

3.2 Data Acquisition 23

3.3 Data Pre-processing 28

3.4 Classification and Identification 29

CHAPTER FOUR
HUMAN SUBJECT INDOOR POSITIONING 31

4.1 Overview 31

4.2 Data Acquisition and Data Pre-processing 32

4.3 Classification and Estimation 34

CHAPTER FIVE
RESULTS 39

5.1 Evaluation Method 39

5.2 Identification Results 39

5.3 Indoor Positioning Results 46

CHAPTER SIX
CONCLUSION 53

6.1 Overview of Thesis 53

6.2 Future Directions 54

REFERENCES 56

LIST OF TABLES

Table 1 The frequency range of each band. 10

Table 2 The information of seven volunteers. 25

Table 3 The information of the experiment setup. 27

Table 4 Human subjects' identification accuracy for different machine learning models in different environments. 39

Table 5 The identification of human subjects carrying a cell phone accuracy for different machine learning models in different environments. 42

Table 6 Residuals at different positions for different numbers of devices. 49

LIST OF FIGURES

Figure 1	The whole framework of our research.	6
Figure 2	Human subject identification system.	23
Figure 3	The RTL2832U device.	24
Figure 4	The distributions of the devices and human occupying locations.	26
Figure 5	The power spectrums collected by the 5 SDRs for human subject 1.	28
Figure 6	Human subject indoor positioning system.	30
Figure 7	The distributions of the devices and human occupying locations.	32
Figure 8	The power spectrum collected by the 6 SDRs when location 1 is occupied by a human subject.	34
Figure 9	The confusion matrix of KNN in the lobby.	40
Figure 10	The confusion matrix of KNN in the living room	41
Figure 11	The confusion matrixes of using KNN to classify different human subjects in the lobby.	43
Figure 12	The confusion matrixes of using KNN to classify different human subjects in the living room.	44
Figure 13	The accuracy for different number of SDRs.	46
Figure 14	The distributions of the training positions and testing positions.	47
Figure 15	The human occupancy position estimation results.	48

LIST OF ABBREVIATIONS

RF	Radio frequency
P-RF	Passive radio frequency
PHSID	Passive human subject identification
PHSIR	Human subject identification with radio frequency
PIR	Passive Infrared
ML	Machine learning
KNN	K-nearest neighbors
GPR	Gaussian process regression
ITU	International telecommunication union
ID	Identification
EEG	Electroencephalogram
SVM	Support vector machines
GPS	Global positioning system
RFID	Radio frequency identification
WLAN	Wireless local area network
SDR	Software-defined radio
DVB-T	Digital video broadcast technology
TV	Television

CHAPTER ONE

INTRODUCTION

1.1 Human Sensing

In recent years, modality recognition, artificial intelligence, and machine vision have become hot topics. With the development of these technologies, human sensing has brought great convenience to our lives. Human sensing (e.g., human presence detection, counting, localization, tracking, and identification) plays a pivotal role in people's daily life. In the automotive field, the human-sensing-based applications are most commonly applied to either the surrounding pedestrians' detections, or the driver's identification and authentication. Meanwhile, these technologies are commonly used in indoor navigations, access control, and alarm devices. Therefore, human sensing becomes an indispensable part of our lives.

The realization of human sensing is mainly based on observable human body characteristics and specific human properties. For example, fingerprint recognition and face recognition are developed by utilizing observable human body characteristics. There are some human sensing methods, such as voice, signature, and signal recognition, which are based on the specific properties of the target objects. In addition, two types of sensors are studied in human sensing with respect to the mechanism of how they receive the signals from the objects: sensors. Electric field sensors, ultrasound, doppler-shift sensors, and lasers can be classified as active sensors, while passive sensors are known as motion sensors, pressure-sensitive sensors, inertial sensors, and passive infrared (PIR) sensors.

Our previous research found that the radio frequency is very sensitive to the human subject. It is verified in our previous research That human occupancy has a strong influence on the passive RF spectrum [1]. This method is achieved by utilizing the passive spectrum in the environment, and there is no need to add any new signal sources. Therefore, this method does not need to occupy the radio frequency band.

Based on the sensitivity of the passive RF signals to the human subject, we propose two hypotheses in this thesis. First, different human subjects could generate different spectrum signatures. Second, the occupation of human subject at different indoor locations will generate different RF signatures on the passive spectrum. Based on these two hypotheses, two applications of passive RF spectrum monitoring approaches are proposed: human subject identification and indoor positioning.

1.2 Human Subject Identification

Identity authentication is an important part of information security which relies on physiological recognition and human classification towards individual identification. Traditional authentication methods mainly rely on various certificates, pins, and passwords. However, certificates are easy to lose, and passwords are very liable to forget and crack [2][3]. Therefore, biometric identification is becoming more and more popular. Biometric identification is to use human biological characteristics for classification towards identity authentication. These technologies mainly utilize the inherent physiological or behavioral characteristics to verify personal identity. Physiological biometric technology utilizes the human body structure and shape, which includes fingerprint recognition [4][5][6], face recognition [7][8][9], and iris recognition [10][11][12]. The behavioral characteristics refer to behavior patterns of a human subject such as signature [13], voice [14], and gait [15][16]. Biometric identification is a reliable, convenient, and fast method to provide authentication.

Currently, identification (ID) technologies are based on fingerprint or face recognition widely applied in our daily life such as account login, online payments, and access control. These technologies have their own strengths and limitations. For example, fingerprint recognition can identify humans with high accuracy, but this technology requires the users to be close to the ID devices and cannot realize ID without touching the device. Face, iris, and gait recognition rely on a camera to capture the biometrics, which can only work in a well-lit environment. In addition, some researchers utilized Electroencephalogram (EEG) to identify human subjects [17]. Although the ID accuracy of EEG is very high, it requires contact sensors using complex equipment to receive the signals. Hence, the EEG technology is relatively complicated and costly. Therefore, it is important to develop a user-friendly and low-cost solution to overcome these limitations.

Passive Radio Frequency (P-RF) signals are available almost anywhere except in extreme environments. P-RF signals have multiple benefits such as less contamination of spectrum, low interception, and reduced power requirement. In previous research, human occupancy can alter the passive spectrum [1][18]. According to these characteristics, P-RF signals can be utilized to detect human presence. Based on the previous research, we propose a new hypothesis, which is that different human subjects could generate different spectral signatures, and a system of human ID can be developed for P-RF spectral monitoring. The method neither requires the touch of any device nor depends on the lighting condition of the environment. Software-defined radio (SDR) acquires the P-RF spectrum in the frequency bands that are sensitive to human occupancy.

Statistical machine learning (ML) algorithms are widely used and can classify the sample spectrums associated with different human subjects, including decision trees, support vector machines (SVM), k-nearest neighbors (KNN), and random forest. The trained ML algorithms can be employed to achieve human subject identification.

1.3 Human Subject Indoor Positioning

The development of mobile computing devices and the advancement of wireless network technology have made positioning technology a hot research topic. Positioning technology has been applied in different areas ranging from national defense to security in our daily lives. For instance, military navigation services such as intelligence collection, explosion monitoring, and emergency communications all depend on the accurate positioning of human operators. Applications such as private navigation, monitoring, rescue, and transportation are in huge demand. Therefore, positioning technologies have become an essential function in our life.

The Global Positioning System (GPS) is a mature positioning system based on satellite signals. However, it is not suitable for indoor positioning. First, the materials of the building can appreciably attenuate the satellite signals. Second, there are many electronic devices in indoor environments. The presence of these signals interferes with the transmission of satellite signals. This will greatly lower the accuracy of GPS-based positioning in an indoor environment. Therefore, positioning technologies other than GPS are required for indoor positioning.

Existing indoor positioning methods are mainly based on labeling techniques and active signals. The signal transmitters should be placed in the indoor environment, and users need to obtain the signal strength of the signal receivers and the information about the signal source to locate the human subject. The mainly indoor positioning technologies via active signal include infrared, ultrasonic, and WIFI positioning. Labeling technique such as RFID positioning is to add tags to the human subject and to acquire data from the tags to achieve the target's position. Although the existing methods can achieve human subject indoor positioning with high accuracy, the limitations of these methods cannot be ignored. The existence of obstacles greatly affects the accuracy of infrared positioning. Ultrasonic positioning is limited by the positioning range. In addition, when the location of the signal transmitter is unknown, or the human subject does not carry a tag, positioning becomes very difficult. Therefore, we propose a method that is based on the passive RF spectrum, which requires neither knowing the origin of the signal source nor tagging the human target to achieve indoor positioning.

Based on the previous research, human subject occupancy can alter the passive spectrum. We proposed another new hypothesis, human subject's occupations at different indoor locations can generate different RF signatures on the passive spectrum. Our indoor positioning method can classify and estimate human subject locations by detecting and processing the variations of the scanned spectrum.

Two ML algorithms are used to achieve indoor positioning. A decision tree is used to classify a human subject's position on a grid. Then Gaussian process regression is employed to estimate the coordinates of the human subject. This process provides a higher resolution of indoor localization than the decision tree algorithm [19].

1.4 Thesis Contributions

This paper presents passive human subject identification (PHSIR) and indoor positioning with radio frequency classification approaches utilizing a new biometric sensing modality. A passive spectrum monitoring method is developed to collect human subject signature samples, which are used to train machine learning algorithms to classify human subjects and estimate their indoor positions. The whole framework of our research is shown in Figure 1.

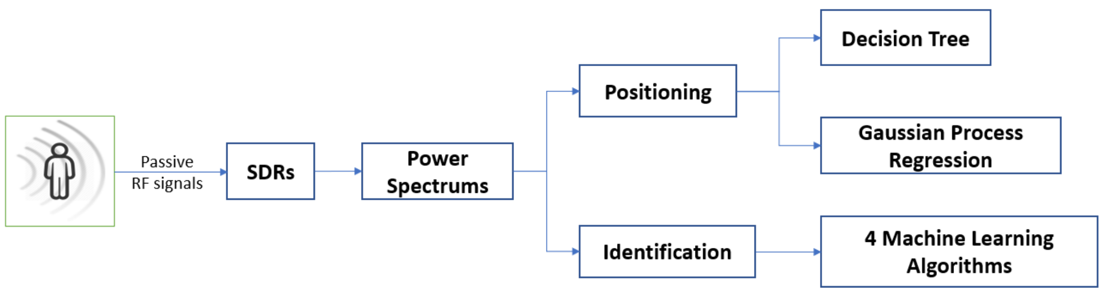

Figure 1. The whole framework of our research.

A passive spectrum monitoring approach is developed for human sensing, which does not need to take up crowded frequency bands so using passive spectrums can save energy. In addition, radiation of RF signals lead to electromagnetic pollution and harmful to people's health. Therefore, passive sensing techniques are preferred overactive sensing techniques if they can get the comparable results. The contributions of this research are as follows.

In the area of human subject identification:

- The hypothesis that different humans generate different P-RF spectrums signatures is verified.
- Four ML algorithms were applied to classify the passive power spectrum for human subjects' identification. The classification accuracy is over 94%.

In the human subject indoor localization part:

- The hypothesis that human subject at different locations generates different signatures on the passive RF on the spectrum is verified.
- A decision tree is trained to roughly classify a subject's position on a grid. And the Gaussian process regression (GPR) method is employed to estimate the indoor location, and the average error is lower than 0.8 m.

1.5 Thesis Outline

The remaining chapters of this thesis are organized as follows. Chapter Two reviews the passive radio spectrum monitoring applications, existing identification and indoor positioning techniques in the literature, several ML algorithms used in our human subject identification and indoor positioning methods, explaining the software-defined radio to support human occupancy authentication and indoor localization, as well as discussing the feature of our methods. Chapter Three presents the proposed identification technique, which includes using SDR to collect raw data for each human subject, preprocessing the raw data, and utilizing four ML algorithms to classify different human subjects. Chapter Four focuses on our indoor positioning technique. Chapter Five shows the experimental results. In the human subject identification experiment, the identification results of the two different experiments are compared. And the influence of human subjects carrying a cell phone is considered. In the human indoor positioning experiment, the effect of using a different number of SDRs on the accuracy of the human occupying positions classification is discussed. Based on the known human occupying position coordinates, the estimated human subject's position is calculated. Chapter Six discusses the conclusions and future directions of our research.

CHAPTER TWO

LITERATURE REVIEW

2.1 Passive Radio Spectrum Applications

Radio technology transmits signals through radio waves. The principle of radio technology is based on the theory of electromagnetic waves. Radio can be transmitted in an air or vacuum environment. Based on this phenomenon, radio can realize information transmission. Information can be loaded on radio waves through modulation. When the electric wave propagates through space and reaches the receiving, the information can be extracted from the current through demodulation, and the purpose of information transmission can be achieved.

Radio has a wide range of applications. The earliest application of radio was the telegraph. The use of the Morse telegraph can realize the transmission of information between ship and land station. With the rapid development of wireless communication technology, the number of available frequency bands has greatly increased. Radio applications are becoming more and more extensive. Nowadays, the main applications of radio include wireless communications, broadcasting, radar, communication satellites, navigation systems, and networks.

Radio waves are generated by the rapid vibration of the magnetic fields. The speed of vibration is the frequency of the wave, and different frequency bands can be used to transmit various information. Radio spectrum is electromagnetic waves with frequencies between 3Hz and 300GHz. The International Telecommunication Union

(ITU) divides the radio spectrum into 12 frequency bands. The frequency range of each band is shown in Table 1.

Table 1. The frequency range of each band.

Band Name (Abbreviation)	**Frequency**
Extremely low frequency (ELF)	3-30Hz
Super low frequency (SLF)	30-300Hz
Ultra-low frequency (ULF)	300-3000Hz
Very low frequency (VLF)	3-30KHz
Low frequency (LF)	30-300KHz
Medium frequency (MF)	300-3000KHz
High frequency (HF)	3-30MHz
Very high frequency (VHF)	30-300MHz
Ultra-high frequency (UHF)	300-3000MHz
Super high frequency (SHF)	3-30GHz
Extremely high frequency (EHF)	30-300GHz
Tremendously high frequency (THF)	300-3000GHz

Currently, with the rapid development of radio and the widespread application of communication equipment, the radio frequency (RF) spectrum becomes a very important resource of the country. Many countries have very strict control over the use of the RF spectrum. The U.S. government owns nearly 60% of the radio spectrum. Most of the spectrum owned by the government is from 300 MHz to 3 GHz frequency band [20]. Due to the limited radio spectrum resources, the reasonable use of spectrum monitoring becomes crucial. If the usage and changes of the spectrum can be recorded and analyzed, the spectrum resources can be effectively utilized, and many applications can be realized by using the characteristics of the spectrum. In order to ensure the normal transmission of information, maintain the transmission of radio waves, and effectively utilize the limited spectrum resources, the establishment and improvement of radio monitoring methods have great significance to the implementation of spectrum management.

Radio signals mainly include active and passive signals. The active signal is the detection signal sent out by the transmitter actively, and then the devices receive the signal, finally, the change of the measured signal is detected through the conversion circuit. The passive signal means that it does not actively transmit the signal, just using the receiver detects the change of the external measured signal, and then the data is monitored by the sensitive element and the conversion circuit. Radar, Wi-Fi, and ultrasonic technologies are mainly realized by using active signals. The main applications of passive signals include passive infrared signals, passive radar, and radio frequency identification (RFID) technology. In our research, we just utilize the characteristics of the passive RF spectrums to the environment and artificial intelligence method to achieve human subject identification and indoor positioning. Passive signals include many advantages. First, the passive signal has no specific signal source or signal transmitter, so our methods do not need to take up any spectrum's resources. Second, passive signals are safer than active signals because passive signals do not transmit signals but passively receive signals. Therefore, electromagnetic signals cannot be used to capture, track and attack passive signals. Third, passive signals are more friendly than active signals because passive signals are relatively easy to deploy and low maintenance costs. Based on the above-mentioned advantages, if the indoor positioning and identification of the human subject can be realized by using passive signals, this will greatly broaden the application of passive signals.

Radar in the general sense refers to an active radar, which is a traditional radar that radiates electromagnetic waves to illuminate targets by itself for detection, positioning, and tracking. The electromagnetic signals emitted by active radars are easily captured, so many researchers have begun to study new radar systems that do not emit electromagnetic waves by themselves. This kind of radar system that uses non-cooperative external radiation sources for detection and positioning is a passive radar. Passive radar does not use the transmitter to emit energy but receives microwave energy reflected by warm objects or other sources to detect targets. A passive radar system has an antenna and a very sensitive receiving device. The ability of passive radar to identify targets mainly depends on the surface temperature difference between the targets and the reflection coefficient of the target, the incident angle between the antenna beam and the target, wireless polarization and beamwidth, and the minimum detectable electrons of the receiver. Nowadays, passive radar has great development and wide application. Passive radar systems can achieve rapid detection of targets [21]. Passive radar adds sub-array to enhance anti-jamming capability and realizes multi-target positioning [22]. Passive radars utilize signal transmissions to achieve target localization and navigation [23]. Some searchers use satellite illumination to realize multi-band passive radar imaging [24]. However, due to the reliance on third-party transmitters, the operator cannot actively control the transmitter. When the effective radiation power of transmitters is low, the signal between the target and the receiver is blocked, or the signal between the receiver and the transmitter is blocked, the passive radar cannot be used normally.

Passive infrared technology adopts the passive infrared method, and this system mainly consists of the optical system and infrared sensor. The detector does not emit any energy, but passively receives and detects infrared radiation from the environment. Once infrared radiation is sensed, it is focused by the optical system to cause the infrared sensor to generate a change in the electrical signal. The PIR technology has widely developments. Some researchers using PIR sensors developed an intelligent energy-saving system. This system uses PIR sensors to sense the presence of students in the classroom to automatically turn on or off the fans, air conditioners, and lights [25]. In addition, a progressive and hybrid target detection algorithm can be used to improve the results of two PIR processing algorithms [26]. PIR also can be used to achieve human indoor localization [27]. Other researchers in our laboratory used deep learning to process PIR data to achieve human biometric authentication and stationary detection, and his method can accurately classify for room occupancy, the number of occupants, the approximate location of the human targets [28] [29].

RFID is a communication technology that can identify specific targets and read and write related data through radio signals without the need to establish mechanical or optical contact between the identification system and specific targets. The principle of RFID is that after the electronic tag enters the reader area, the electronic tag can accept the microwave signal transmitted by the radio frequency identification reader. Rely on the energy obtained by the induced current, the product information stored in the tag can be read. RFID has a wide range of applications. Typical applications include animal chips, car chip anti-theft devices, access control, parking lot control, and automated production lines. The main limitation of RFID technology is that RFID electronic tag information is easily read and maliciously revise. There are many developments. UHF-RFID passive tags were tagged on the drone to locate its position [30]. And RFID also can be used for train localization based on passive tags [31]. Some researchers used an RFID tag, which is fabricated from electro-textile materials and integrated it into clothing. This method is designed for body movement-based human-technology interaction [32]. RFID can be utilized to develop gate control systems based on a Convolutional Neural Network [33]. RFID technology can also be applied to monitor and control the posture of the cane to help vision-impaired people [34].

2.2 Existing Identification Techniques

The current human classification, recognition, and identification technologies mainly depend on physiological characteristics, which include fingerprint, face, and iris recognition, and behavioral characteristics of movement are based on the signature, voice, and gait patterns. These technologies are discussed as follows.

Fingerprint recognition mainly includes two technologies, capacitive fingerprint recognition, and optical fingerprint recognition. The principle of capacitive fingerprint recognition is to integrate pressure sensing, capacitive sensing, thermal sensing, and other sensors into an integrated chip. When the fingerprint is pressed on the surface of the chip, the internal capacitive sensor will generate peaks and valleys in the fingerprint. The charge or temperature difference is used to form a fingerprint image, which is matched with the fingerprint library to complete the fingerprint classification which is referred to as identification (ID) in the biometric literature [35]. The optical fingerprint recognition is when a finger touches the screen, the screen emits light to illuminate the finger area, and the reflected light that illuminates the fingerprint returns to the sensor on the screen to achieve the ID [6]. The advantages of fingerprint recognition technology are low price, small size, and easy integration. However, the disadvantage is this method requires the user to touch the receiver. If the hand is injured or the glove is worn, the recognition cannot be reliably achieved.

Face recognition is a popular biometric technology. Face recognition technology first needs to find the position of the face in the picture. When a face is found in the picture, the facial features will be marked. Then feature points of the human face such as eyes, nose, and mouth are extracted. Finally, facial matching is realized by comparing the face features [36]. Facial recognition is used widely, but it cannot be used in dim light or when the face is obscured. For instance, currently wearing masks in public places due to COVID 19 greatly restricts the efficacy of facial recognition.

Voice recognition is a technology that recognizes human subjects by analyzing the physical characteristics of the user's voice. Although voice recognition technology has been used in some products, they are not very convenient to use, because human voices are highly variable, and signatures require training [14]. In addition, compared to other biometrics, voice recognition is more complicated, and which is inconvenient in some conditions.

Gait recognition is to use the camera to collect the image sequences of a human walking (or other movements), extracts feature from the joints, and compares the dynamic feature patterns with the stored data for behavior classification. Human subjects can be recognized in long-distance or low video quality [37]. Some other researchers developed a method, which is to capture the gait information by using a Wi-Fi signal [38]. However, the limitation of gait-based analysis needs to capture the dynamic information of the human subjects. If using a camera to capture the dynamic information, it will also depend on the quality of the imaging.

2.3 Existing Indoor Positioning Techniques

The current positioning technologies include infrared positioning, ultrasonic positioning, WIFI positioning, and radio frequency identification (RFID) positioning. These methods currently used in indoor positioning are discussed below.

Infrared positioning is used to locate targets by receiving the modulated infrared signal sent by optical sensors installed indoors [39]. This technology has high indoor positioning accuracy. Since light cannot penetrate obstacles, infrared rays can only travel within the line of sight and are easily interfered with by other light [40]. The localization performance by using the infrared positioning technique cannot perform well when objects exist on the site. Besides, to efficiently receive infrared, a number of antennas must be installed on the experimental site, which leads to higher costs.

Ultrasonic positioning is widely used for positioning, and the reflection measurement value is used to calculate the distance to the reference node based on the time difference between the transmitted ultrasonic wave and the response echo of the reference node[41]. Authors of [42] who introduced Active Bat are the pioneers of the ultrasonic positioning system. Intensive deployment of a large number of ultrasonic receiving devices is required to achieve a positioning accuracy with a minimum error of 3cm. Although the ultrasonic positioning technique can be used in a non-line of sight circumstance with high positioning accuracy and small errors, the cost of the devices is high. Moreover, the transmission attenuation of ultrasonic signals cannot be avoided, so the effective positioning range of this technique is somehow limited.

WIFI positioning as a Wireless Local Area Network (WLAN) is composed of wireless routers, and wireless access points, which can realize positioning, monitoring, and tracking tasks in complex environments [43]. Signal propagation models can be used to locate the receiving mobile device. The highest accuracy is between 1 m and 2 m [44]. However, the disadvantage of WIFI positioning is that this technique highly depends on wireless routers and access points, and these WIFI devices must stay online while working. Therefore, this method has certain limitations.

RFID positioning uses radio frequency signals and signal strength to detect positions. In [45], an aggregation algorithm is developed to locate targets in three-dimensional space. The hardware tags in the system are distributed in a network without a signal source, and the distance between tags is characterized by the strength of the signals detected by the tags. This method heavily depends on the tags on the target, which limits the application of this technique.

2.4 Software Defined Radio

RTL-SDR is a software-defined radio (SDR) based on digital video broadcast technology (DVB-T) television (TV) tuners with RTL2832U chips. which is used to collect passive RF signals in our research. SDR can realize many operations by developing software. For instance, an algorithm can be modified to make the SDR scan the specific frequency bands. The SDR is tuned to scan only the frequency bands that are sensitive to human occupancy to improve power efficiency [1]. SDR has been widely used in communications, spectrum monitoring, and RF transmitter identification [46], specifically in improving the power amplifier system and transmitter architecture [47]. SDR can be used for real-time communication [48]. SDR can be utilized to receive the animals' nerve signals [49]. The position of the mobile station can be estimated by using the signal strength received by the SDR [50]. SDR can be used to recognize gestures through Wi-Fi signals [51]. In our previous work, SDR was used to scan the RF signal spectrums to detect and estimate human occupancy. Moreover, the SDR devices used in our experiments are low-cost, compact, and easy to deploy.

2.5 Machine Learning Algorithms

ML algorithms are divided into supervised, and unsupervised learning. In the process of supervised learning training, the algorithm requires specific input and output, and a model is built from the training data set, and the result of the test set has derived a model. Decision trees, support vector machine (SVM), k-nearest neighbors (KNN), and random forest are common methods and have been utilized to identify the human subjects. Decision tree and Gaussian process regression are used to classify and estimate the human subject occupying positions. A decision tree refines a feature set through analysis of meaning features based on the domain of interest, such as for human movements [52]. SVM is a statistical method to align features to categories. For example, gender identification of human faces has general features for normal categories SVM [53]. The KNN algorithm is a method based on local approximation and supports clustering. A KNN classifier shows promise for face classification and electrocardiogram (ECG) biometric 0. The random forest can be regarded as a collection of decision trees. A gait recognition method can uniquely identify humans based on the random forest [54].

Gaussian process regression (GPR), a machine learning model, is a non-parametric model that uses Gaussian process priors to perform regression analysis on data. GPR can provide the posterior of the prediction result, and when the likelihood is normally distributed, the posterior has an analytical form. GPR has been applied in the fields of image processing and automatic control [55]. GPR is very suitable for solving positioning problems[56]. The prediction results obtained by GPR are highly accurate [57]. The advantages of GPR include using only a few training data points for regression to acquire all position results, predicting high-dimensional data, and flexibly using different kernel functions to construct the relationship between the independent variables and the dependent variables [58]. In our research, the independent variables are passive RF spectrums, and the dependent variables are humans occupying positions. We used the GPR model to infer the relationship between the passive RF spectrums and human occupying positions.

2.6 Discussions

Comparing available human classification technologies, there are some differences with the PHSIR method. First, PHISR uses passive signals to recognize and classify human subjects, which is an efficient and friendly method. PSHSIR utilizes the signatures in the passive spectrum for different human subjects to realize the classification, so the new signal transmitters are not needed. It is beneficial for different environments, such as those crowded with various RF signals, which have electromagnetic pollution, and radiation of RF signals can be harmful to people's health. Therefore, passive sensing techniques are preferred overactive sensing techniques if they can achieve high accuracies.

Second, PHSIR is not affected by factors such as environment, light, and temperature. The traditional methods are active sensing as human identification needs the user to touch on the device or using in an environment that sufficient light to achieve. However, PHSIR incorporates passive sensing not requiring a user to touch a sensor, nor wear the receiving devices, but also can be used in a dark or low-light environment.

Finally, cell phones are indispensable electronic devices in our lives. We conducted several experiments to verify the effect of carrying a cell phone for human subject identification accuracy. Our experimentation research found that carrying a cell phone will not affect the accuracy of PHSIR human recognition.

Comparing with the above-mentioned indoor positioning methods, there are some differences with our method. First, the signals utilized in our method are passive RF signals instead of active signals. This is an efficient and environmentally friendly method because a new signal source is not needed in a space. As mentioned above, nowadays, our environment is already crowded with various RF signals, which leads to a lot of electromagnetic pollution. Radiation of RF signals can be harmful to people's health and take up spectrum resources. Therefore, passive sensing techniques are preferred overactive sensing techniques if they can achieve the same level of accuracies.

Second, our method does not require the human subject to carry any tags or receiving devices. This is completely different from the traditional approaches which build a radio map based on received radio signal strength to realize positioning. Our method deploys software-defined radio (SDR) at multiple fixed locations and scans the passive spectrum at those locations. Machine learning methods are applied to map spectrum alterations to human locations. This eliminates the need of tagging the targets or carrying a device, which is beneficial in the applications of monitoring.

Last, RTL-SDR, a software-defined radio, can be tuned to scan only the frequency bands that are sensitive to human occupancy to improve power efficiency [18]. This method can improve the efficiency of machine learning, increase the speed of training models, and enhance the accuracy of the trained model. Moreover, the devices used in our experiments are low-cost, compact, and easy to deploy.

CHAPTER THREE

HUMAN SUBJECT IDENTIFICATION

3.1 Overview

This chapter introduces our passive human subject identification with the radio frequency (PHSIR) method. PHSIR assumes that different human subjects can generate different signatures as passive RF spectrums, because different people have a different heartbeat, body temperature, or body shape, and the subtle variation can be detected by the ML classifiers. Based on these assumptions, utilizing passive RF signals can achieve human identification and disambiguation.

The proposed human subject's identification method includes three stages. The first stage is data acquisition, which applies multiple SDRs to collect passive RF spectrums in the frequency bands that are sensitive to human occupancy in two environments. The second stage is the data pre-processing, which converts the raw data to the power spectrum. The third stage is the analysis of subject signature detection (person, animal, etc.), human recognition (e.g., gait), human classification (e.g., adult) which lead to identification from a known set of requirements (e.g., allowed the student to enter a location. The power spectrums were classified by four ML algorithms to identify human subjects. The details about data acquisition, data pre-processing, and identification of different human subjects are described in the following subsections. The framework of the human subject's identification system via passive spectrums is shown in Figure 2.

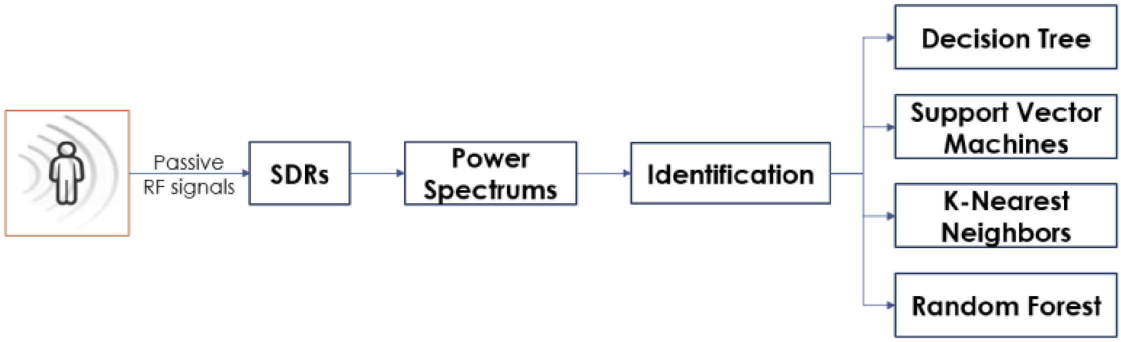

Figure 2. Human subject identification system.

3.2 Data Acquisition

In our experiments, five RTL2832U were applied to scan the selected frequency band to acquire the passive RF spectrum. The device RTL2832U used in our experiment is shown in Figure 3.

In our previous work, the frequency bands sensitive to human occupancy are identified to be around 300MHz to 420MHz. In this experiment, the frequency bands from 300MHz to 420MHz were scanned and the sample step size is 1.2MHz. We conducted the experiments in two different environments. One is the lobby in the Engineering Center building at Oakland University, while the other one is the living room of an apartment, which has more furniture compared to the lobby. Five SDR of the same model were placed around the human subject. Seven volunteers were asked to stand in the center in turn. The information of 7 volunteers is shown in Table 2.

Figure 3. The RTL2832U device.

Table 2. The information of seven volunteers.

Index	Weight (Kg)	Height (Meter)	Age	Sex
1	69	1.83	25	Male
2	66	1.67	30	Male
3	80	1.76	26	Male
4	67	1.75	25	Male
5	65	1.75	30	Male
6	60	1.79	24	Female
7	70	1.74	42	Male

Then seven known human subjects were numbered sequentially from 1 to 7. The distributions of the devices and human occupying locations are shown in Figure 4.

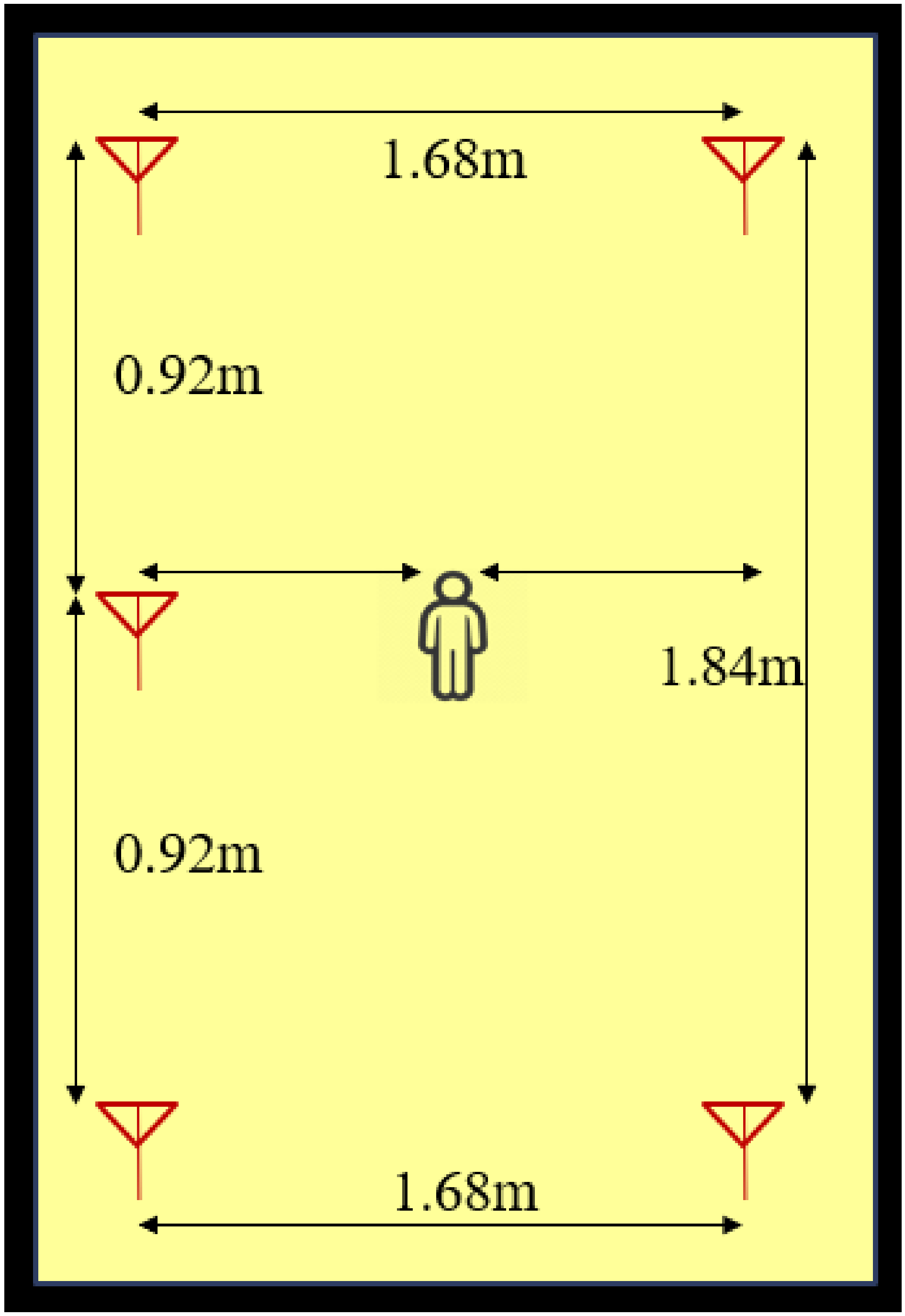

Figure 4. The distributions of the devices and human occupying locations.

The data acquisition was performed with the human subjects occupying the same location in the same posture. A total of 100 frequency bands were scanned in one spectrum sample. Seven human subjects without any electrical devices were asked to stand in the center and then conduct movements. For each human subject,20 samples of RF data were collected in each environment of which some volunteers carried their cell phones. A total of 140 samples were collected in each environment. The information of the experiment setup is listed in Table 3.

Table 3. The information of the experiment setup.

Items	Description
Name of Devices	RTL-SDR
Number of Devices	5
Number of Humans	7
Experimental Site	Living Room Lobby
Frequency Range	300 MHz - 420 MHz
Number of Frequency Bands	100
Scanning Step	1.2 MHz
Sampling Rate	2.4 MHz
Duration	2 milliseconds per frequency band
Spectrum Samples per Human	20

3.3 Data Pre-processing

In our experiment, the input of the ML classifiers is the average power spectrum density, so the average power of each human sensitive frequency bands should be calculated as follow:

$$P(f) = 10 \cdot \frac{\log_{10} \sum_{i=1}^{N} \left(\frac{s_i(f)}{127.5} - 1\right)^2}{N/_2} \tag{3.1}$$

where $P(f)$ is the average power of N samples centered at f, N is the number of samples per frequency band, and $s_i(f)$ represents the value of raw data of the i-th sample received by SDR at frequency f. In our work, the sample rate is 2.4MHz and one collection duration of each frequency band is 2 milliseconds, so N is 4800. The passive power spectrums collected in the lobby by the five SDRs for human subject 1 are shown in Figure 5.

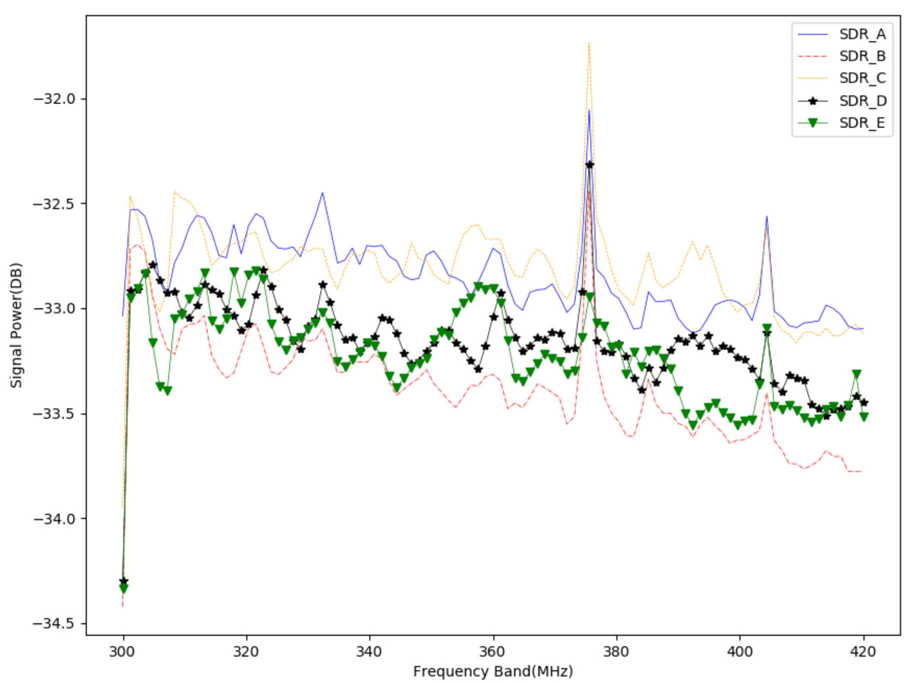

Figure 5. The power spectrums collected by the 5 SDRs for human subject 1.

3.4 Classification and identification

In our experiment, four supervised ML algorithms were used to associate spectrum samples with different human subjects, including decision trees, SVM, KNN, and random forest. The input of the four ML classifiers is the power spectrum density, and the output is the index of the human subjects. A total of 140 samples of the power spectrum density were split into 75% training data set and 25% testing data set. For each experimental environment, four classifiers were trained. Finally, 35 spectrum samples were used to test the classifiers in each environment. As highlighted, since the subjects were known, the scenario represents the opportunity for human subject identification for authentication. Inherent in the analysis is that movement supports detection and recognition, and the ML classifies the subjects. With a known training signature of specific people, the spectrum can be used to not only classify a human subject but as compared to a known signature, ID a specific person who has appropriate access assuming they consented to the P-RF collection to create their profile – as consistent with fingerprint access. The experiment results of the proposed approach will be presented in Chapter 5.

CHAPTER FOUR

HUMAN SUBJECT INDOOR POSITIONING

4.1 Overview

Recent research in our group shows the presence of human subjects can be detected via passive RF signals [1]. The experimental results suggest that humans can cause variations in the passive spectrum. In this work, we find that human subjects at different locations can generate different signatures on the passive spectrum. Utilizing machine learning algorithms to associate these spectrum characteristics with the corresponding human occupancy locations can achieve accurate indoor positioning.

The proposed human positioning method includes 3 steps: data acquisition, data pre-processing, and classification and estimation of human subject position by decision tree and Gaussian process regression. The human subject indoor positioning system is shown in Figure 6.

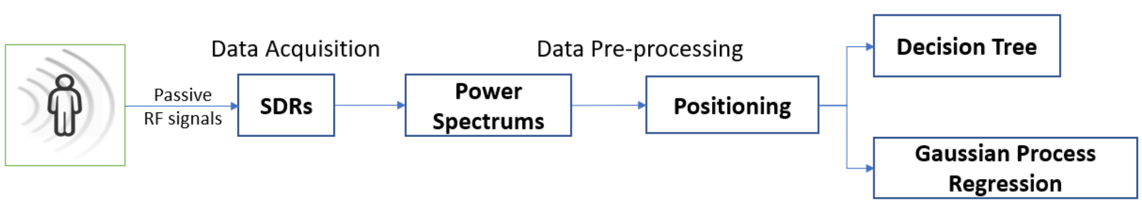

Figure 6. Human subject indoor positioning system.

In the first step, 6 SDRs were deployed to collect spectrum samples at multiple locations simultaneously when a human subject occupies different locations in an indoor environment. In the second step, the pre-process calculates the power spectrum from the collected raw data. In the third step, two machine learning algorithms were used for the classification and estimation of human position.

A decision tree was used to classify the signal power spectrums of humans occupying different positions. This process can verify that humans at different locations can generate different signatures on the passive spectrum. Last, GPR was used to build a model between the signal's spectrums and the location coordinates of the human subject. According to the GPR model and signals spectrums of unknown positions, the positions of the human subject can be estimated. The details about data acquisition, data pre-processing, and classification and estimation of human subject position are described in the following subsections.

4.2 Data Acquisition and Pre-Processing

RTL-SDR was used to collect RF signal data in our experiments. The SDRs can scan the spectrum frequency from 24MHz to 1760MHz. Our previous research shows that the frequency band around 330MHz is the most sensitive to human occupancy. In this work, frequency bands range from 300MHz to 420MHz were scanned. The sampling rate is 2.4MHz. The experimental site was a classroom at Oakland University. The size of the classroom was 10m×12m. Six devices of the same model are placed around the classroom. Their locations are shown in Figure 7. Then, 20 locations evenly distributed in the classroom were selected for human occupancy. The distance between adjacent points is 1.8m. A human subject occupies one location when six SDRs scan the spectrum simultaneously. The distributions of the devices and points are shown in Figure 7.

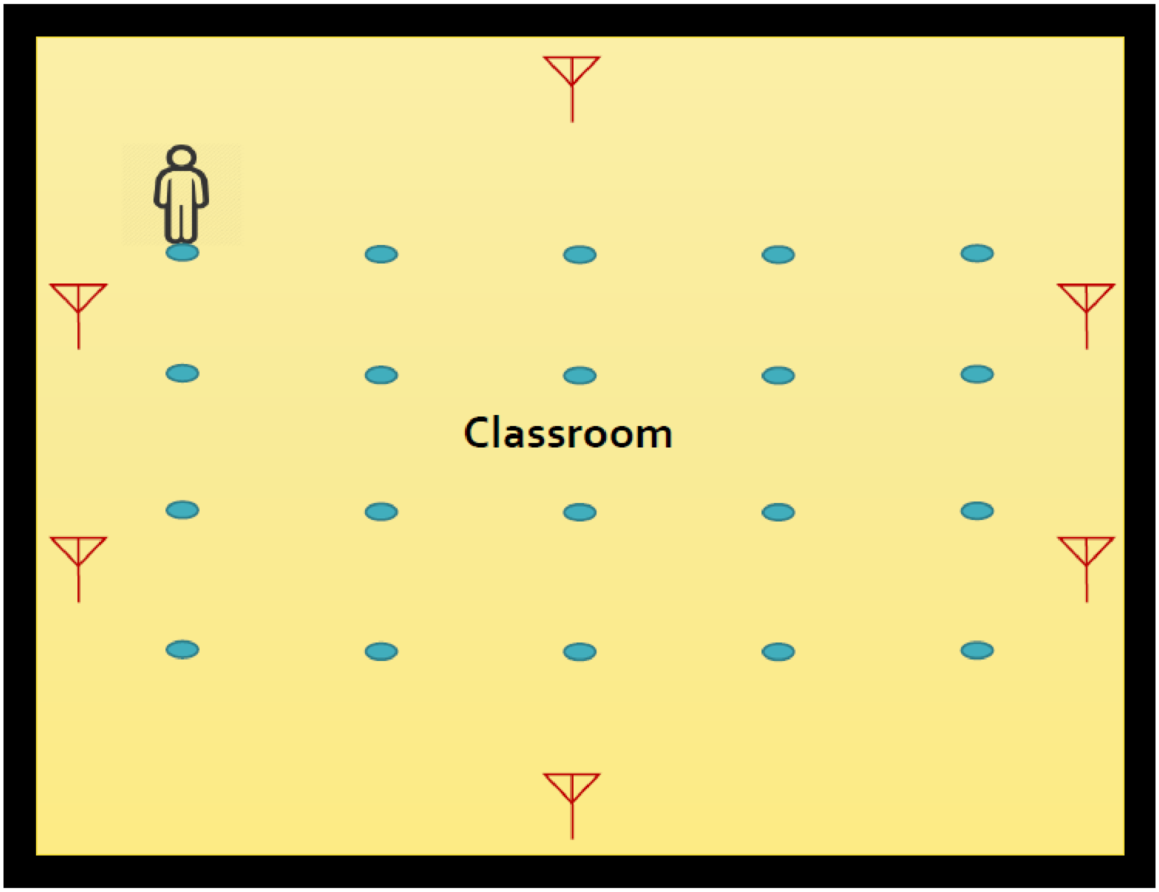

Figure 7. The distributions of the devices and human occupying locations.

The 20 locations are numbered sequentially from the upper left to lower right one row after another. For example, the points in the first row are numbered 1 to 5 from left to right, and so on. The data collected with different human occupancy locations are used differently in the decision tree and GPR algorithms, which will be discussed later in detail. The coordinates of each occupancy location are recorded to facilitate the training of GPR.

During the experiment, the SDRs should not only be away from the metal properly but also be dispersed as much as possible, and only one human subject was inside the classroom. The spectrums for each location of human occupation were collected from 8 pm to 10 pm. We collected 10 spectrum samples when the human subject occupied a specific location. Each sample contains 100 frequency bands.

4.3 Data Pre-processing and Data Pre-processing

In our experiment, the input of the decision tree and GPR is the power spectrum density. The average power of each frequency band is calculated using function (3.1). The difference from the Chapter 3 is that this experiment uses 6 SDRs, and 10 samples were collected at each human occupying. The power spectrum density collected by the 6 SDRs when location 1 is occupied by a human subject is shown in Figure 8.

The decision tree is a classical machine learning approach. Each internal node in the decision tree model represents the judgment of the attribute, each branch represents the output of a judgment result, and each leaf node represents a classification result [59]. A decision tree can classify data sets layer by layer according to feature values. The advantage of the decision tree is that the computational complexity is not high, and the classified results can be presented intuitively.

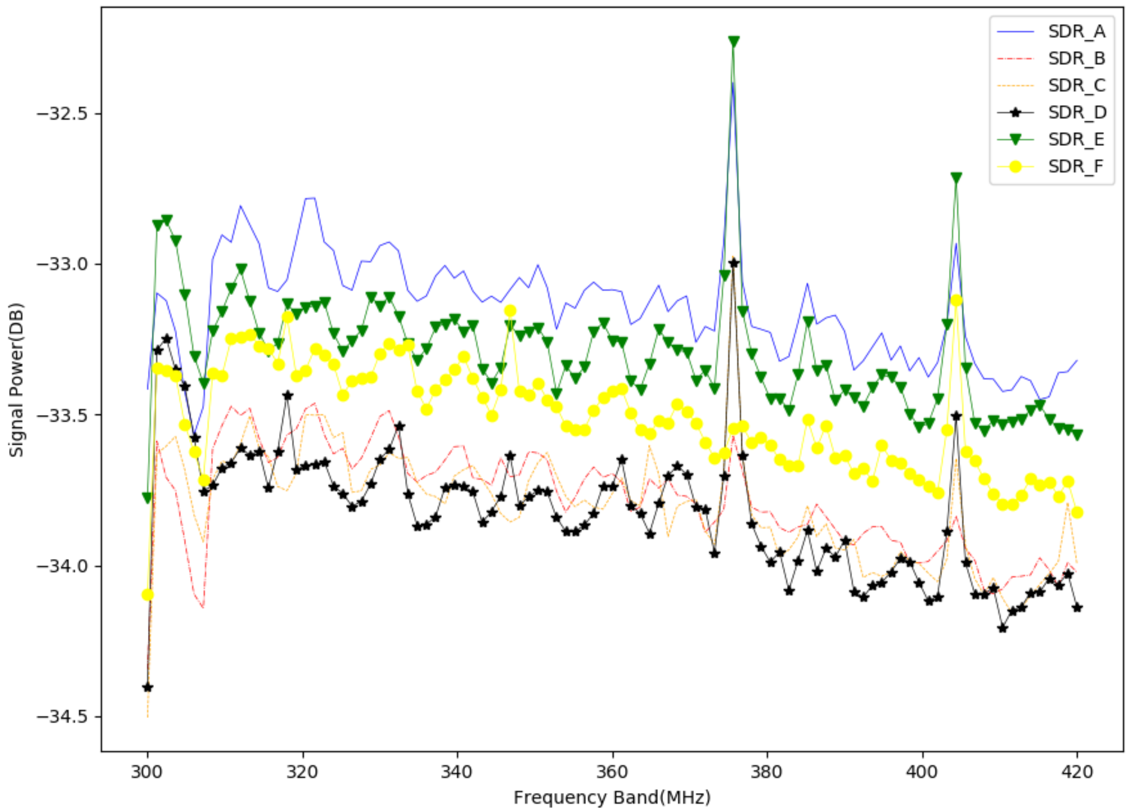

Figure 8. The power spectrum collected by the 6 SDRs when location 1 is occupied by a human subject.

In this study, it is very important to verify whether a human subject at different locations can generate different RF signatures on the spectrum. A decision tree method is applied to test the hypothesis. If the spectrums obtained when different positions were occupied by human subjects can be classified by a decision tree with high accuracy, it will be strong evidence that the hypothesis is true. There are 20 positions tested in our experiments. The signal spectrums obtained at each human occupying position are used as the input of the decision tree. The output of the decision tree is the index of the position.

The process of using GPR to solve the problem is introduced as follows. The first step is to choose the appropriate mean function and kernel function. The kernel function in GPR is the covariance function. The second step is to calculate the kernel matrix of the training samples. The third step is to calculate the kernel vectors of the points to be predicted and all points of the training samples. Finally, the prediction values are obtained through the function of means and covariances. The GPR method is applied in our research as follows.

First, the training set D for GPR is as follows:

$$D = (x, y) = \{(x_i, y_i)\}_{i=1,\dots N} \tag{4.2}$$

where the N-dimensional inputs vector x represents the stacked spectrum of multi-sensors collected by SDRs when the i-th position is occupied by human subjects. The output y_i is the coordinates of the location occupied by the human subject. GPR only depends on the mean $m(x_i)$ and covariance $k(x_i, y_i)_{i,j=1,\dots N}$.

$$m(x_i) = E[f(x_i)] \tag{4.3}$$

$$k(x_i, y_i)_{i,j=1,\dots N} = E\left[\left(f(x_i) - m(x_i)\right)\left(f\left(x_j\right) - m\left(x_j\right)\right)\right] \tag{4.4}$$

The GPR model is to infer the relationship $f: x \rightarrow y$, which maps a spectrum x to a position y. The conditional distribution of y is determined when the x_i is given. The function of as follows:

$$f(x_i) \sim GP\left(m(x_i), k\left(x_i, x_j\right)\right) \tag{4.5}$$

$$y = f(x) + w \tag{4.6}$$

where $w \sim N(0, \sigma_i^2)$ is the measurement noise. $m(x_i)$ is the mean. $k(x_i, y_i)_{i,j=1,\ldots N}$ is the covariance. Usually, when the data is preprocessed, the mean value is 0. Then the prior distribution of y can be expressed as follows:

$$y \sim N(0, K(x, x) + \sigma_i^2) \tag{4.7}$$

In function (4.7), $K(x, x) = K_N = \{k(x_i, x_j)\}_{i,j=1,\ldots N}$ is a $N \times N$ covariance matrix. The joint distribution of training set (x, y) and test set (x_*, y_*) can be expressed as follows:

$$\begin{bmatrix} y \\ y_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} K(x, x) + \sigma_N^2 I_N & K(x, x_*) \\ K(x_*, x) & k(x_*, x_*) \end{bmatrix}\right) \tag{4.8}$$

where $K(x, x_*) = K(x_*, x)^T$ is a $N \times 1$ covariance matrix between the test set x_* and the input x of the training set. $k(x_*, x_*)$ is the covariance matrix of test point x_*. I_N is the N-dimensional identity matrix. In our work, the squared exponential covariance is chosen to be the kernel function. The function of the kernel is presented as follows:

$$k(x_i, x_j) = \sigma^2 \exp\left(-\frac{d(x_i, x_j)^2}{2l^2}\right) \tag{4.9}$$

In this function, σ is constant. $d(x_i, x_j)^2$ represents the Euclidean distance. l is the length scale of the kernel function. For fear that the trained model is overfitting or underfitting, the value of σ needs to be determined carefully. l needs to be given an initial value and boundary. Then the posterior distribution of $f(x_*)$ can be calculated using the functions below:

$$P(f(x_*)|x_*) = N(\mu_*(x_*), \sigma_*^2(x_*)) \tag{4.10}$$

$$\mu_*(x_*) = K(x_*, x)[K(x, x) + \sigma_N^2 I_N] f(x) \tag{4.11}$$

$$\sigma_*^2(x_*) = k(x_*, x_*) - K(x_*, x) K(x, x) + \sigma_N^2 {I_N}^{-1} K(x, x_*) \tag{4.12}$$

$P(f(x_*)|x_*)$ is the conditional probability of $f(x_*)$. $\mu_*(x_*)$ is the mean of prediction corresponding to x_*, which is the coordinate of position $f(x_*)$. $\sigma_*^2(x_*)$ is the variance of the predicted value. Thus, the accuracy of positioning is calculated with the Euclidean distance between the original and predicted positions. The experiment results of the proposed approach will be presented in Chapter 5.

CHAPTER FIVE

RESULTS

5.1 Evaluation Method

The F1 score is commonly used to evaluate the accuracy of the model so in our experiment, the F1 score was calculated to obtain the identification accuracy. The function of the F1 score is as follows:

$$accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{5.1}$$

$$precision = \frac{TP}{TP + FP} \tag{5.2}$$

$$recall = \frac{TP}{TP + FN} \tag{5.3}$$

$$F1 = \frac{2 \cdot \ precision \cdot \ recall}{precision + recall} \tag{5.4}$$

In these functions, TP means true positive. FP is a false positive. FN represents false negatives. TN means true negative. The FI score can only be between 0 and 1. When the value of the F1 score is close to 1, the accuracy of the model is high, and the classification result is better. When the value of the F1 score is close to 0, the accuracy of the model is low, and the classification result is worse.

5.1 Human Subject Identification Results

This part presents the experiments to demonstrate the proposed human identification method and compares the identification accuracies using different ML algorithms in two environments. In addition, the impact of human subjects carrying a cell phone is provided as a sensitivity analysis. The details are described in the following.

We compared the human subjects' ID accuracy of different ML models in different environments. The results are shown in Table 4.

Table 4. Human subjects' identification accuracy for different machine learning models in different environments.

	Decision tree	**SVM**	**KNN**	**Random forest**
Lobby	74.3%	62.8%	**94.3%**	74.3%
Living room	71.4%	82.8%	**94.3%**	74.3%

The results show that decision trees and random forest achieved the same ID accuracy in different environments, and the accuracy in the living room is higher than the accuracy in the lobby when using the SVM to identify human subjects. KNN got the highest accuracy in two different scenarios with 94.3%. And the confusion matrixes of KNN in the lobby and living room are shown in Figure 9 and Figure 10.

In Figure 9, X-axis represents the actual human subject label, and Y-axis represents the predicted human subject label. The results show that for human subject two, 19 samples were classified into human subject two, one sample was predicted to be human subject one. For human subject three, two samples were classified into human subject one, and one sample was classified into human subject two. And one sample of human subject six was incorrectly predicted.

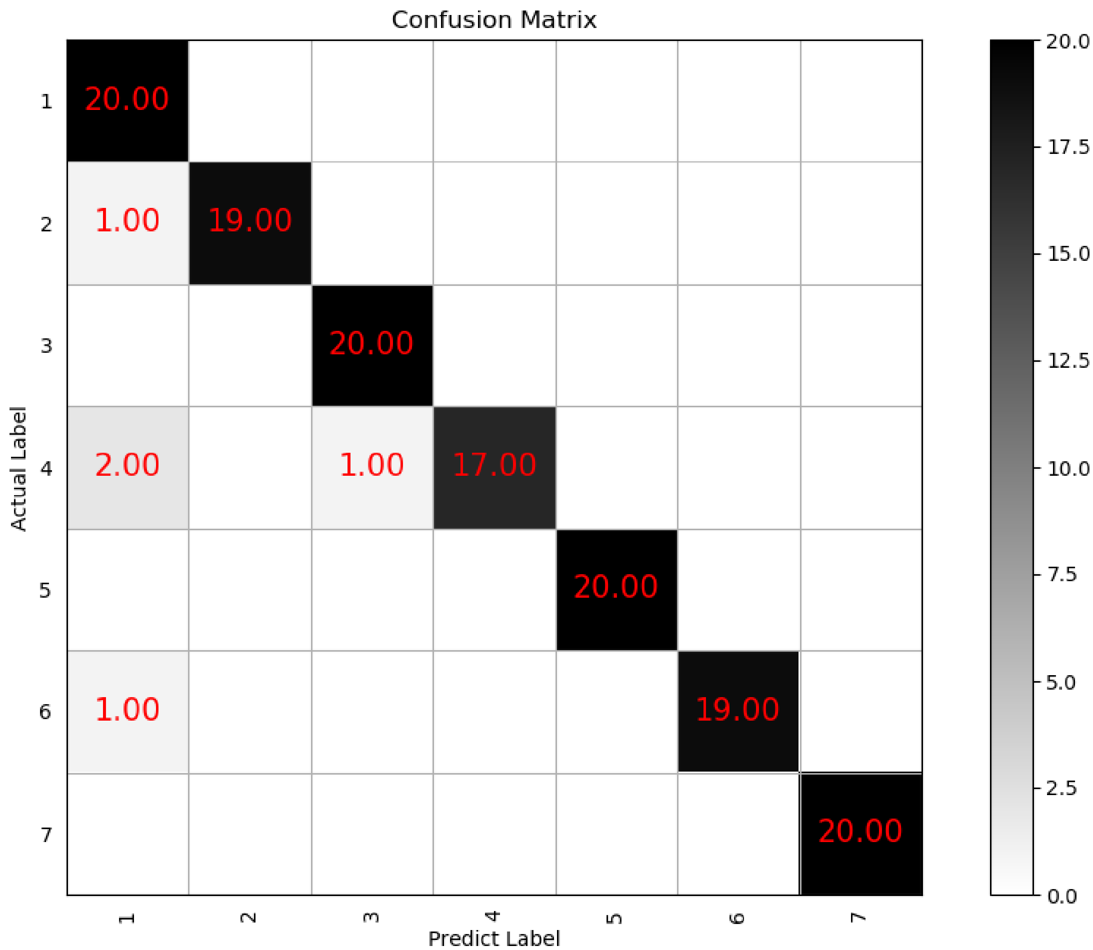

Figure 9. The confusion matrix of KNN in the lobby.

In Figure 10, we can see that almost every human subjects' data can be accurately classified by KNN, but the human subject two and five have a little fuzzy. One sample of human subject two and two samples of human subject four was classified into human subject five.

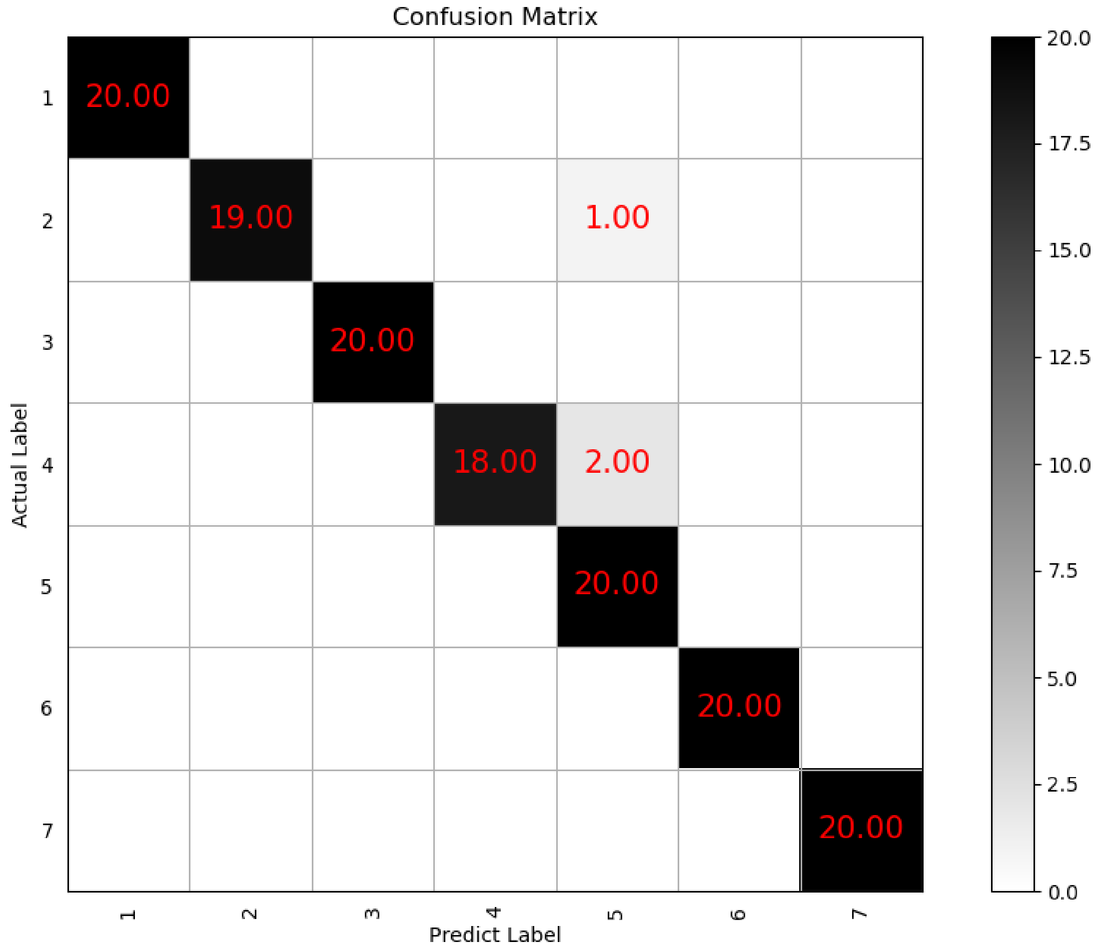

Figure 10. The confusion matrix of KNN in the living room.

These results not only verified different human subjects can generate different signatures on passive RF spectrums but also demonstrated spectrum monitoring can be used to identify different human subjects. The signatures generated by different human subjects on the spectrum are more suitable to be distinguished by KNN.

In order to consider the impact of spectrum congestion, the subjects were asked to carry a cell phone. The data were collected in two environments when human subjects carried their cell phone that can be compared without a phone. The identification results are shown in Table 5.

Table 5. The identification of human subjects carrying a cell phone accuracy for different machine learning models in different environments.

	Decision tree	**SVM**	**KNN**	**Random forest**
Lobby	68.5%	51.4%	**97.1%**	82.8%
Living room	88.5%	88.5%	**97.1%**	96.2%

In Table 5, using KNN achieved the highest ID accuracy - over 97%. This result means carrying the cell does not reduce the identification accuracy; and in fact, increased the performance. And the confusion matrixes of using KNN to classify different human subjects in the lobby and living room are shown in Figure 11 and Figure 12.

In Figure 11, only one sample of human subject seven was classified into human subject one, but other human subjects' classification accuracy achieves 100%.

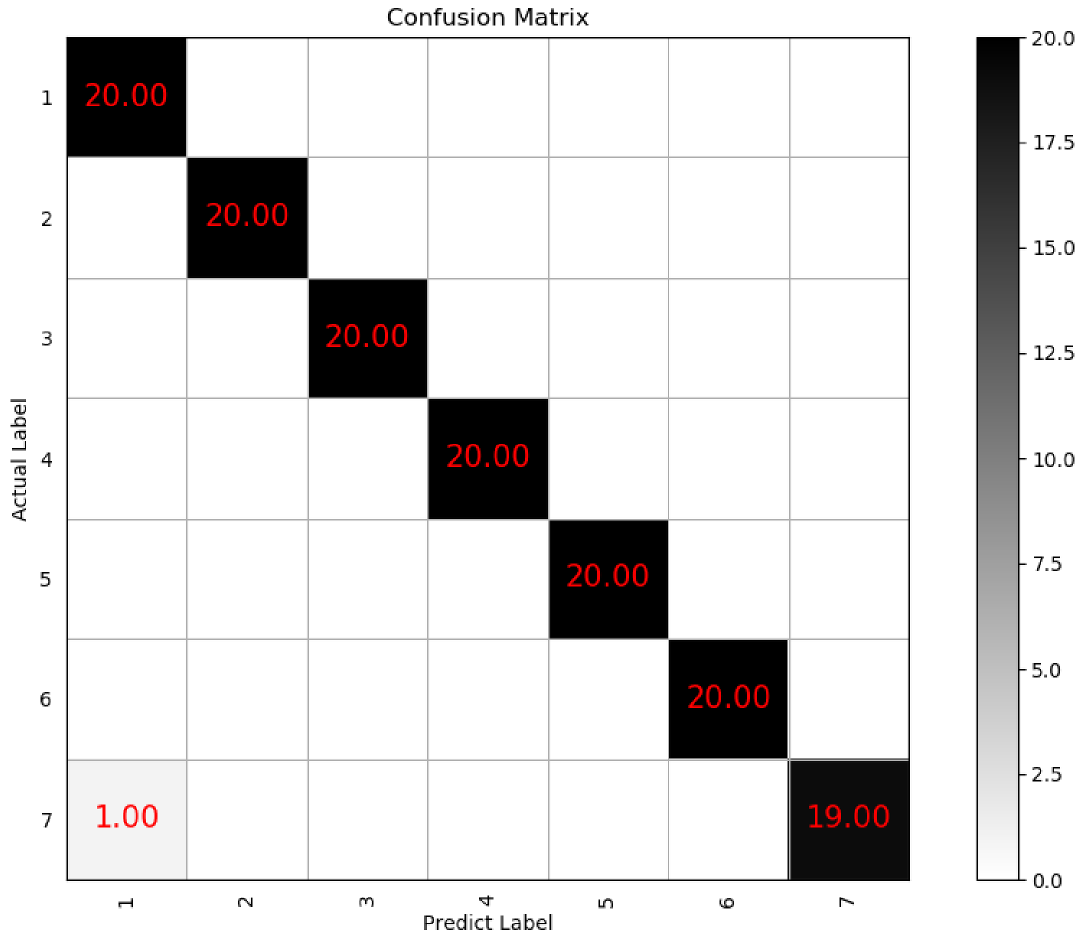

Figure 11. The confusion matrixes of using KNN to classify different human subjects in the lobby.

In Figure 12, human subjects one and three have one sample data fuzzy each other. The classification results of any other human subject are very good.

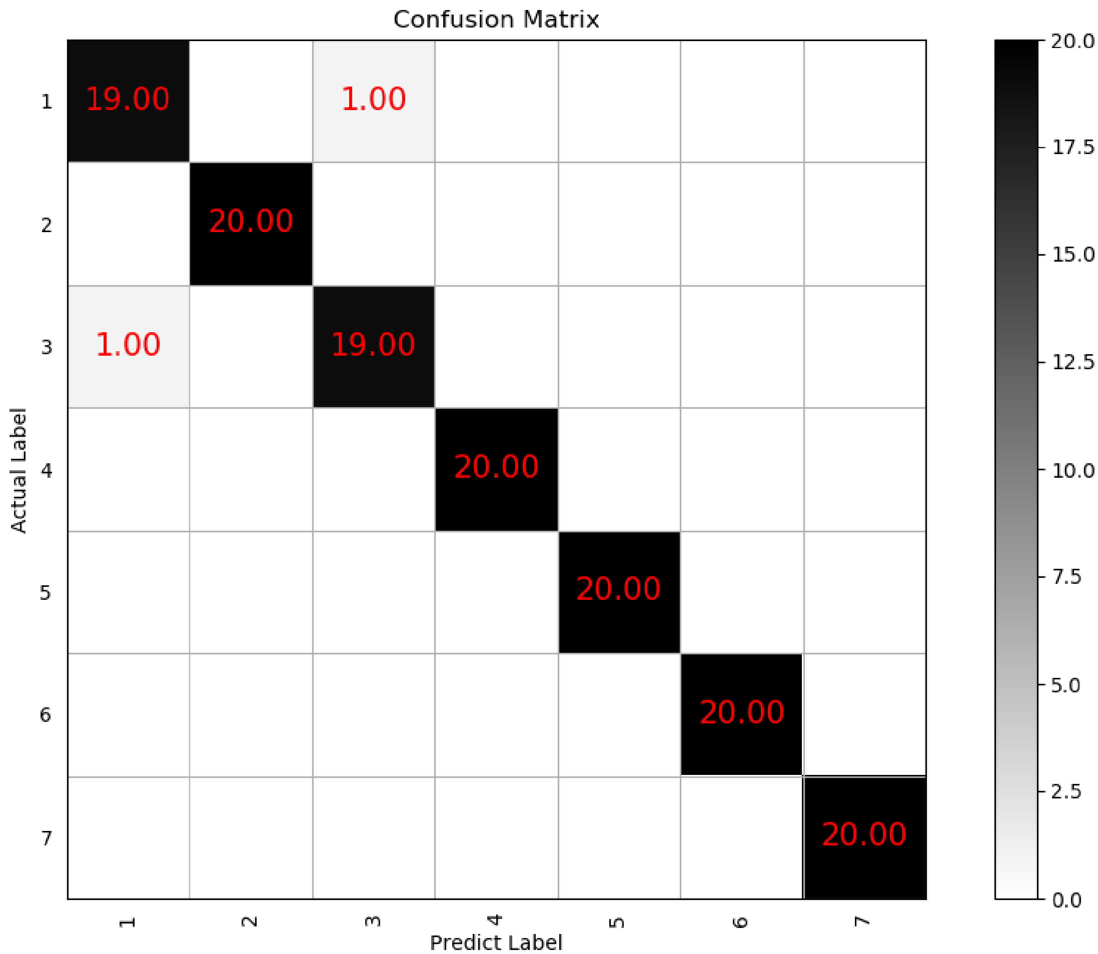

Figure 12. The confusion matrixes of using KNN to classify different human subjects in the living room.

Compared to the accuracy in the lobby, all four ML methods have achieved equal or higher accuracy in the living room. The considered reason is that the living room's area size is smaller than the lobby's area size, and there are more items in the living room, so the spectrum can be more sensitive to the human subject occupancy. The signatures generated by different human subjects can be more easily detected. Hence, these results demonstrate that the PHSIR method is suitable for identifying human subjects in a closed environment. More enhanced experiments will afford analysis for the sensitivity and robustness of the technique.

5.3 Indoor Positioning Results

This section presents the experiments to demonstrate the proposed indoor positioning method and compares the positioning accuracies when using a different number of SDRs.

A decision tree is first applied to validate the hypothesis that human subjects occupying different locations can cause variations of the passive spectrum. The passive RF spectrums at 20 human occupying positions are the input of the decision tree. The indexes of 20 positions from 1 to 20 are used as the output of the decision tree.

The cross-validation method is utilized to verify the accuracy of the classification model and obtain the accuracy by calculating the F1 score. If the value of the F1 score is high, which indicates that there is a strong correlation between signal spectrums and human position. However, if the value of the F1 score is low, this means that it is difficult to build an accurate model between the spectrums and the location coordinates of the human subject.

We compared the impact of using different numbers of SDRs to obtain data for classification results. The classification accuracy is shown in Figure 13.

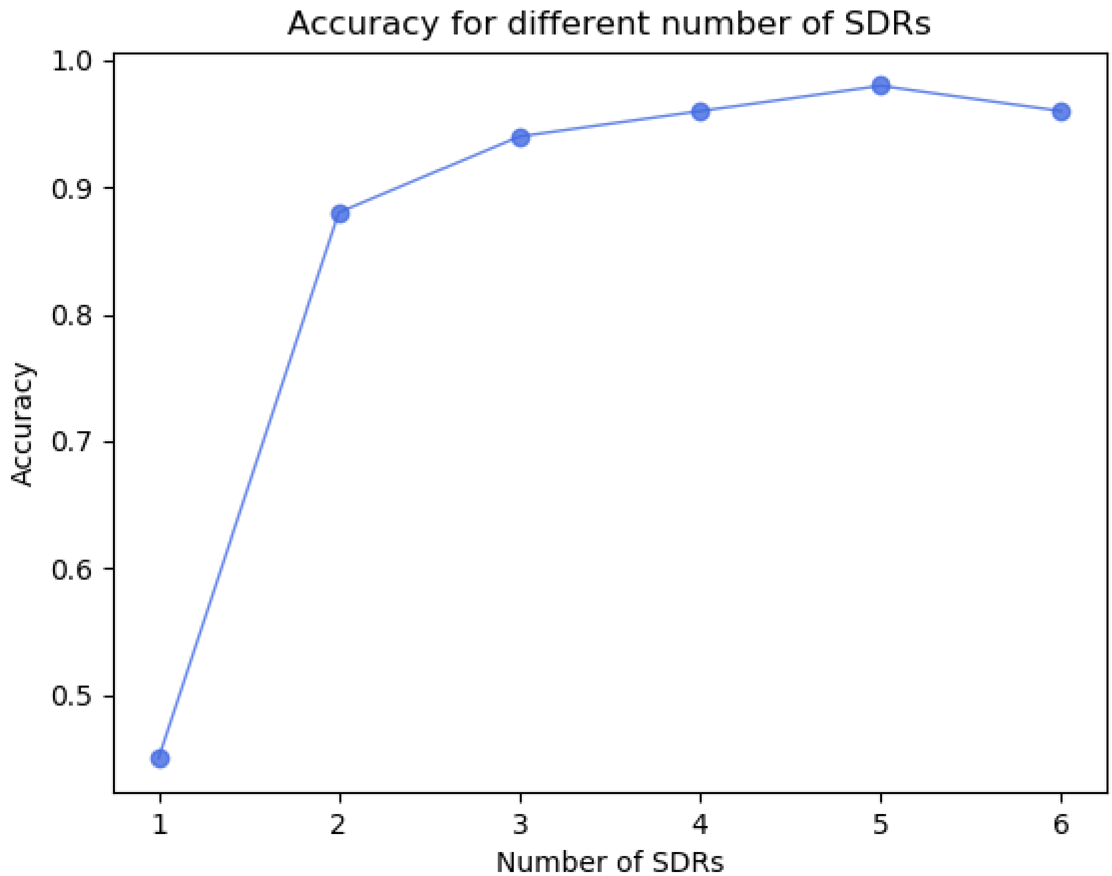

Figure 13. The accuracy for different number of SDRs

The results show that signal spectrums can be classified by human position, and the accuracy of classification is very high. We found that when more than three devices were used for classification, the accuracy was stabilized. When using spectrums collected by one SDR, the accuracy dropped significantly. When using spectrums collected by multiple devices to train the model, the accuracy increased. These results verify that human occupancy can alter the passive RF spectrum, and humans at different locations can generate different signatures on the spectrum. It also shows that using five SDRs produces the most accurate result. Therefore, this spectrum's characteristics can be utilized to achieve human positioning classification. Next, we used 4, 5, and 6 SDRs respectively to estimate the human positions.

GPR was used to predict the position of test points. We selected 16 positions for GPR training and 4 positions for testing. The distributions of training and testing positions are shown in Figure 14. The dots represent the positions of training sets, and the stars represent the positions of testing sets.

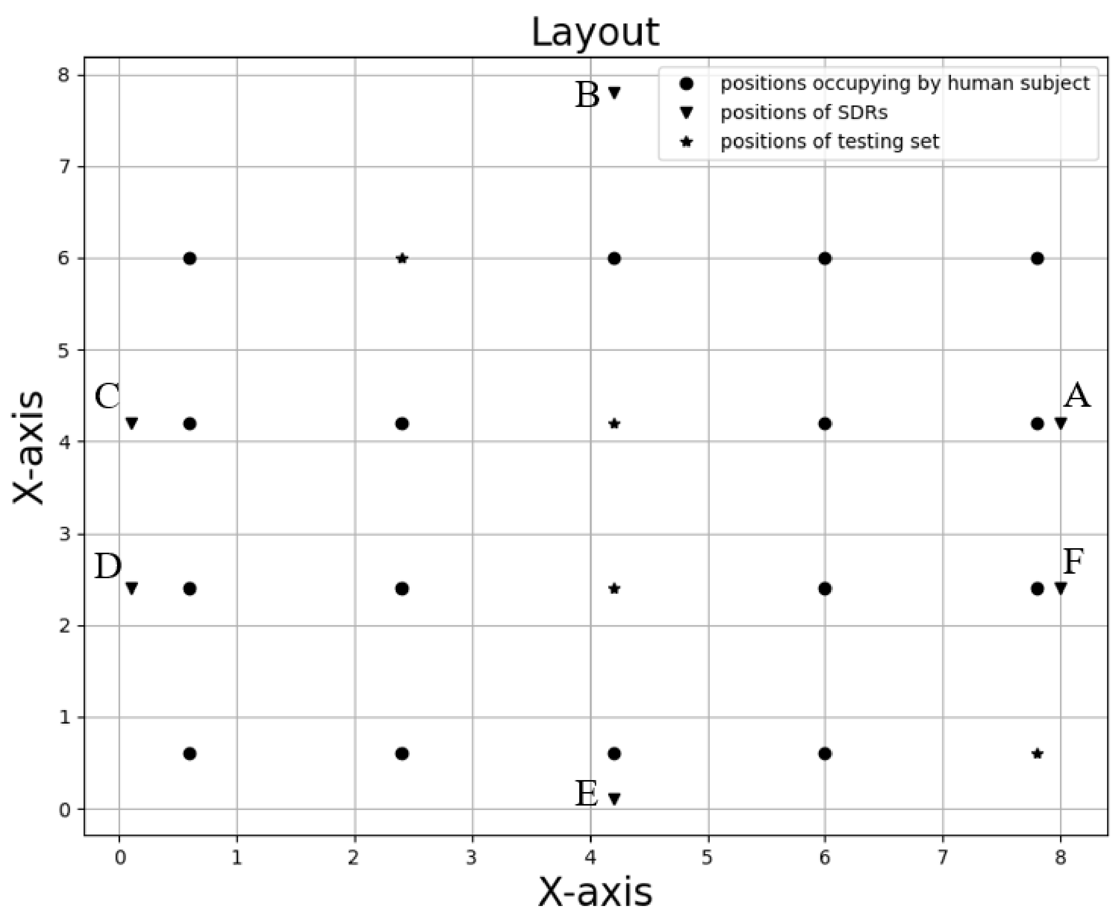

Figure 14. The distributions of the training positions and testing positions.

We compared the accuracy of using different numbers of SDRs to obtain spectrums for prediction results of testing human subject positions. The test results of each human occupying position are shown in Figure 15.

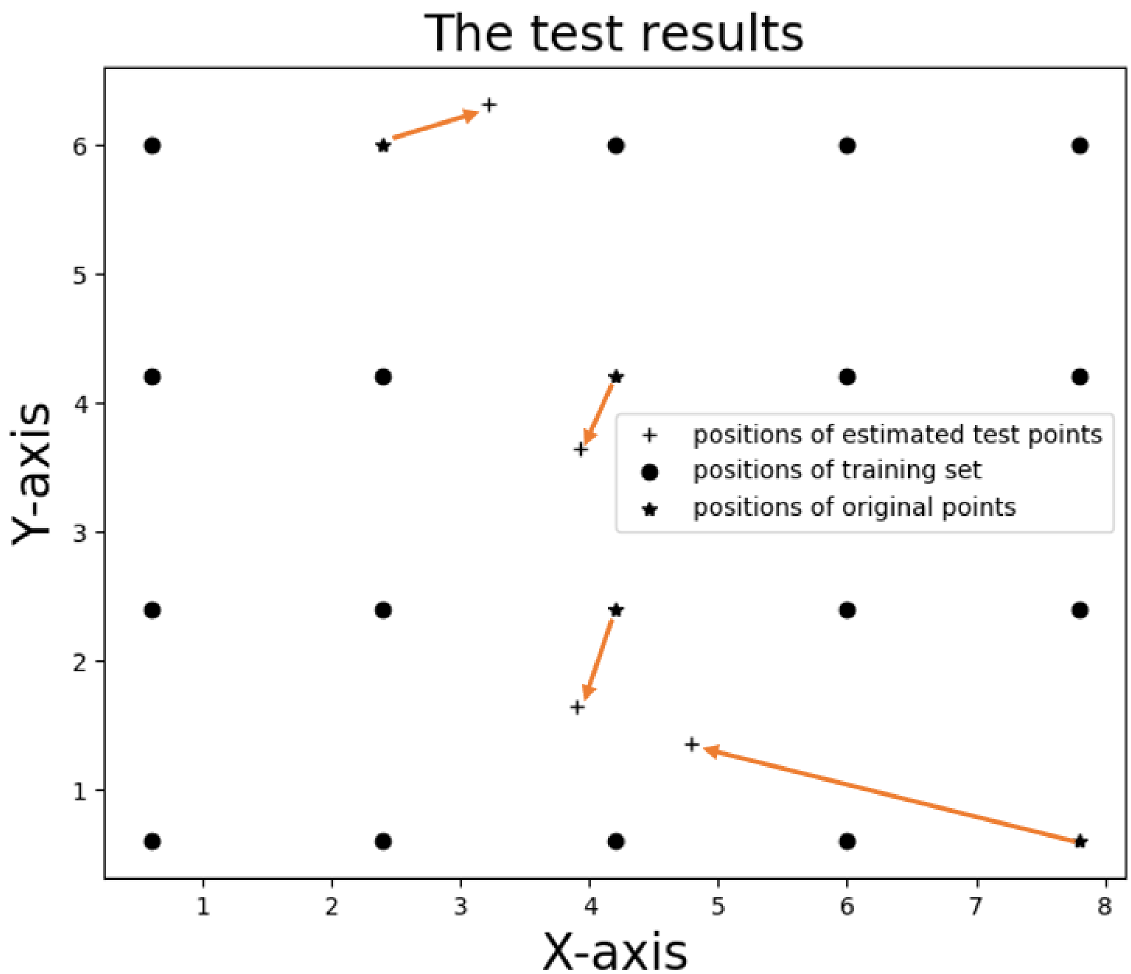

Figure 15. The human occupancy position estimation results.

In Figure 15, the dots indicate the positions of the training set. The stars indicate the original position of the test point. The crosses represent the predicted positions based on the spectrum obtained on the stars point through the GPR model. A set of crosses and stars connected by each arrow indicates each set of original and predicted positions. The results of position 1, 2, and 3 show good accuracy, while position 4 has the worst result.

In order to obtain more detailed information and compare the effects of different numbers of SDRs, we calculated the average residuals of different test points after modeling different numbers of SDRs. The residuals were calculated using Euclidean distance between cross points and star points. The results are shown in Table 6.

Table 6. Residuals at different positions for different numbers of devices.

Position	**6 devices**	**5 devices**	**4 devices**
1	0.837m	0.835m	0.826m
2	0.618m	0.694m	0.775m
3	0.803m	0.823m	0.852m
4	3.106m	3.108m	3.131m

In Table 6, the residuals calculated from the data received by six devices, five devices, and four devices are separately presented. The layout of the 6 devices is shown in Figure 14. The values of the second column are calculated by utilizing all six devices.

The values of the third column are the average residuals of removing one device from devices A, C, D, and F. The values of the last column are the average residuals from the combination of device B, E, and A/C, D/F.

This table shows the residuals of different numbers of SDRs at different test positions. The prediction results for positions 1, 2, and 3 are fairly accurate. The average value of the residual is around 0.8m. However, the result of the position 4 prediction is very poor. In our experiments, when the SDRs were placed in the central area of the four boundaries, the variation of spectrum caused by the human subject at the corners is the least compared to other positions of the human subject. The behind reason this needs to be explored in future research.

The residuals of using 4, 5, and 6 devices did not have extreme change. When we used six devices, the residual results were only slightly better than using 5 and 4 devices. It shows that the residual error of positioning can be stabilized when using 4 devices.

CHAPTER SIX

CONCLUSION

6.1 Overview of Thesis

In this thesis, a passive method was developed from which results verify that different human occupancy can generate different signatures on the spectrum. In two different environments, the human subjects' ID accuracy is 94%. Human subject carrying the cell phone does not reduce the ID accuracy and was shown to increase the accuracy is 97% in two different environments. PHSIR includes ML methods for classification and can be used as a good secondary identification method when known subject signatures are available. For example, PHSIR can be used in environments that require certification to enter such as offices, laboratories, and meeting spaces. PHSIR scans the passive spectrum for the human subjects who have permission to enter and then the signatures on the spectrums for authorized human subjects are saved in the database. The human subjects who match the signature on the spectrums in the database can pass or enter through identification.

Second, we verify that human occupancy can alter the passive RF spectrum, and humans at different locations can generate different signatures on the spectrum. The passive RF spectrum can be utilized to achieve the positioning of the human subject in an indoor environment. A decision tree can classify the positions of the human on a grid with 98% accuracy using 5 SDRs. Gaussian process regression is applied to construct a model to map the passive RF spectrums and the coordinates of human locations. Using the passive RF spectrums when an unknown position is occupied by the human subject, the model can accurately estimate the human subject's location. The residual error shows that the positioning error is less than 0.8m.

6.2 Future Directions

Our system utilized passive RF signal to develop human subject detection approach. In our work, we just identified different human subjects when they keep the same posture, so there are some restrictions on the applications of the method. Our short-term goals can try to achieve the distinction between different human subjects when they take dynamic postures. In addition, our indoor positioning method just can locate one human subject. In the future, we will focus on the multiple human subjects indoor positioning. There is still a long way to go.

Our research further verified that the passive radio frequency spectrum is sensitive to the human subject. Based on this characteristic, more human subject detection applications can be developed in the future. For example, the passive RF spectrums can be utilized to count the multiple human subjects, to measure the speed of the human subject's movement, and to tracking the movement of the human subject [60].

Many passive signal technologies such as passive infrared and passive radar have the same properties as passive RF signals. These systems do not have specific transmitters, which just utilized receivers to passively receive signals. These technologies are also widely used to identify, locate, and track targets. Machine learning and deep learning are also often used to achieve these applications. But these technologies have not widely used in human subject detections. In the future, human subject detection solutions, that utilize other passive signal technologies can be developed. My classmate in our laboratory has developed some research, that utilizes passive infrared signals to detect the human subject.

Single sensor modality has specific limitations, so multi-sensor fusion technology can provide a good solution to achieve recognition and detection of targets. Nowadays, multi-sensor fusion technology is universally used in the autonomous driving field. By using the cooperation of radar, lidar, and camera to realize road condition recognition in different environments. Multi-sensor fusion technology can also be applied to passive recognition technology. The application of this technology in human subject detection can not only improve the accuracy but also overcome the interference of the environment. Our laboratory will also conduct in-depth research on the application of this technology in the future. The research into this field would be worth of time.

REFERENCES

[1] J. Liu, H. Mu, A. Vakil, B. Ewing, X. Shen, E. Blasch and J. Li, "Human Occupancy Detection via Passive Cognitive Radio," *Sensors*, vol. 20, no. 15, p. 4248, Jul. 2020.

[2] S. Prabhakar, S. Pankanti and A. K. Jain, "Biometric recognition: security and privacy concerns," *in IEEE Security & Privacy*, 1(2): 33-42, March-April 2003.

[3] T. Ko, "Multimodal biometric identification for large user population using fingerprint, face and iris recognition," *34th Applied Imagery and Pattern Recognition Workshop (AIPR'05)*, Washington, DC, USA, 2005, pp. 6 pp.-223.

[4] D. A. Dharmawan and M. Y. Mustar, "Deep Fingerprint Classification in a Low-cost Environment," *International Conference on Information Technology and Electrical Engineering (ICITEE)*, 2020, pp. 297-301.

[5] M. S. Al-Ani, T. N. Muhamad, H. A. Muhamad and A. A. Nuri, "Effective fingerprint recognition approach based on double fingerprint thumb," *2017 International Conference on Current Research in Computer Science and Information Technology (ICCIT)*, Slemani, 2017, pp. 75-80.

[6] E. Sano, T. Maeda, T. Nakamura, M. Shikai, K. Sakata, M. Matsushita, K. Sasakawa "Fingerprint Authentication Using Optical Characteristics in a Finger," *2006 SICE-ICASE International Joint Conference*, 2006, pp. 1774-1777.

[7] L. Fei, W. Yajie, Q. Hongkun and W. Linlin, "Gender Identification Using SVM Based on Human Face Images," *International Conference on Virtual Reality and Visualization*, 2014, pp. 444-448.

[8] B. Knyazev, R. Shvetsov, N. Efremova and A. Kuharenko, "Leveraging Large Face Recognition Data for Emotion Classification," *2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, Xi'an, China, 2018, pp. 692-696.

[9] C. Liu, "The development trend of evaluating face-recognition technology," *2014 International Conference on Mechatronics and Control (ICMC)*, Jinzhou, China, 2014, pp. 1540-1544.

[10] M. D. Amiri, Wafa Barkhoda, Habibollah Danyali and F. A. Tab, "Affine invariant Iris identification using angular and radial partitioning," *2009 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Ajman, United Arab Emirates, 2009, pp. 574-578.

[11] V. G. Garagad and N. C. Iyer, "A novel technique of Iris Identification for biometric systems," *International Conference on Advances in Computing Communications and Informatics (ICACCI)*, 2014, pp. 973-978.

[12] D. M. Daniel and B. Monica, "Person authentication technique using human iris recognition," *2010 9th International Symposium on Electronics and Telecommunications, Timisoara*, Romania, 2010, pp. 265-268.

[13] M. Wang, X. Lin, L. Wang, M. Wen and H. Zan, "Handwriting Signature Identification Based on Improved Adaptive Median Filtering Algorithm," *International Conference on Computer Science and Applications (CSA)*, 2015, pp. 55-59.

[14] S. Kinkiri and S. Keates, "Speaker Identification: Variations of a Human voice," *International Conference on Advances in Computing and Communication Engineering (ICACCE)*, 2020.

[15] V. S. Papanastasiou, R. P. Trommel, R. I. A. Harmanny and A. Yarovoy, "Deep Learning-based identification of human gait by radar micro-Doppler measurements," *European Radar Conference (EuRAD)*, 2021, pp. 49-52.

[16] K. Kurita, "Human Identification from Walking Signal Based on Measurement of Current Generated by Electrostatic Induction," *2011 International Conference on Biometrics and Kansei Engineering, Takamatsu*, Japan, 2011, pp. 232-237.

[17] Y. Di, X. An, S. Liu, F. He and D. Ming, "Using Convolutional Neural Networks for Identification Based on EEG Signals," *10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2018, pp. 119-122.

[18] J. Liu, A. Vakil, R. Ewing, X. Shen and J. Li, "Human Presence Detection via Deep Learning of Passive Radio Frequency Data," *IEEE National Aerospace and Electronics Conference (NAECON)*, 2019, pp. 296-301.

[19] H. Mu, J. Liu, R. Ewing and J. Li, "Human Indoor Positioning via Passive Spectrum Monitoring," *2021 55th Annual Conference on Information Sciences and Systems (CISS)*, 2021, pp. 1-6.

[20] P. Amirshahi, P. Ransom, S. Grippando and I. Navarro, "Comparison of Centralized and Distributed Spectrum Monitoring Methods for Enabling Spectrum Sharing Between Weather Satellites and Terrestrial Networks," *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2019, pp. 1-8.

[21] Y. Liu, X. Wan and X. Sun, "GPU parallel acceleration of target detection in passive radar system," *2016 CIE International Conference on Radar (RADAR)*, 2016, pp. 1-4.

[22] J. H. Huang, M. N. Barr, J. L. Garry and G. E. Smith, "Subarray processing for passive radar localization," *2017 IEEE Radar Conference (RadarConf)*, 2017, pp. 0248-0252.

[23] V. Navrátil, J. L. Garry, A. O'Brien and G. E. Smith, "Utilization of terrestrial navigation signals for passive radar," *2017 IEEE Radar Conference (RadarConf)*, 2017, pp. 0825-0829.

[24] D. Cristallini, I. Pisciottano and H. Kuschel, "Multi-Band Passive Radar Imaging Using Satellite Illumination," *2018 International Conference on Radar (RADAR)*, 2018, pp. 1-6.

[25] C. Twumasi, K. A. Dotche, W. Banuenumah and F. Sekyere, "Energy saving system using a PIR sensor for classroom monitoring," *2017 IEEE PES PowerAfrica*, 2017, pp. 347-351.

[26] S. G. Hong, N. S. Kim and I. Lee, "Progressive and hybrid target detection with PIR sensors," *The 18th IEEE International Symposium on Consumer Electronics (ISCE 2014)*, 2014, pp. 1-2.

[27] K. Lai, B. Ku and C. Wen, "Using cooperative PIR sensing for human indoor localization," *2018 27th Wireless and Optical Communication Conference (WOCC)*, 2018, pp. 1-5.

[28] J. Andrews, A. Vakil and J. Li, "Biometric Authentication and Stationary Detection of Human Subjects by Deep Learning of Passive Infrared (PIR) Sensor Data," *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2020, pp. 1-6.

[29] J. Andrews, M. Kowsika, A. Vakil and J. Li, "A Motion Induced Passive Infrared (PIR) Sensor for Stationary Human Occupancy Detection," *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 2020, pp. 1295-1304.

[30] A. Buffi, P. Nepa and R. Cioni, "SARFID on drone: Drone-based UHF-RFID tag localization," *2017 IEEE International Conference on RFID Technology & Application (RFID-TA)*, 2017, pp. 40-44.

[31] A. Buffi and P. Nepa, "An RFID-based technique for train localization with passive tags," *2017 IEEE International Conference on RFID (RFID)*, 2017, pp. 155-160.

[32] H. He, X. Chen, L. Ukkonen and J. Virkki, "Clothing-Integrated Passive RFID Strain Sensor Platform for Body Movement-Based Controlling," *2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 2019, pp. 236-239.

[33] G. Álvarez-Narciandi, A. Motroni, M. R. Pino, A. Buffi and P. Nepa, "A UHF-RFID gate control system based on a Convolutional Neural Network," *2019 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 2019, pp. 353-356.

[34] W. Li, W. Zhang and K. W. Tam, "Wireless control for Vision Impaired based on posture effect of UHF RFID cane," *2016 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, 2016, pp. 93-96.

[35] S. Heo, J. Song, K. Park, E. Choi, S. M. Kim and F. Bien, "A low-offset, low-noise, fully differential receiver with a differential signaling method for fingerprint mutual capacitive touch screen," *International SoC Design Conference (ISOCC)*, 2017, pp. 166-167.

[36] A. Matin, F. Mahmud and M. T. B. Shawkat, "Recognition of an individual using the unique features of human face," *IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2016, pp. 57-60.

[37] P. Jaychand Upadhyay, P. T. Gonsalves, R. Paranjpe, H. Purohit and R. Joshi, "Biometric Identification Using Gait Analysis by Deep Learning," *2020 IEEE International Conference for Innovation in Technology (INOCON)*, Bangluru, India, 2020, pp. 1-4.

[38] D. Roy, T. Mukherjee, M. Chatterjee, E. Blasch, E. Pasiliao, "RFAL: Adversarial Learning for RF Transmitter Identification and Classification," *in IEEE Transactions on Cognitive Communications and Networking*, 6(2): 783-801, 2020.

[39] R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J. F. Jr, "Evolution of Indoor Positioning Technologies: A Survey," *Journal of Sensors*, p. 23.

[40] A. Alarifi, A. Al-Salman, M. Alsaleh, A. Alnafessah, S. Al-Hadhrami, M. Al-Ammar, H. Al-Khalifa, "Ultra Wideband Indoor Positioning Technologies: Analysis and Recent Advances," *Sensors*, vol. 16, no. 5, p. 707, May 2016.

[41] H. Zuo, N. Zhang, J. Xu, J. Juan Shan, and J. Jiao, "The Study Research on Indoor Positioning Technique," *presented at the 2015 4th National Conference on Electrical*, Electronics and Computer Engineering, Xi'an, China, 2016.

[42] O. J. Woodman and R. K. Harle, "Concurrent scheduling in the Active Bat location system," *in 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, Mannheim, Germany, Mar. 2010, pp. 431–437.

[43] S. Gao and S. Prasad, “Employing Spatial Analysis in Indoor Positioning and Tracking Using Wi-Fi Access Points,” *In Proceedings of the Eighth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness (ISA '16). Association for Computing Machinery,* New York, NY, USA, 27–34.

[44] Z. Wu, K. Fu, E. Jedari, S. R. Shuvra, R. Rashidzadeh and M. Saif, "A Fast and Resource Efficient Method for Indoor Positioning Using Received Signal Strength," *in IEEE Transactions on Vehicular Technology*, vol. 65, no. 12, pp. 9747-9758, Dec. 2016.

[45] P. Jankowski-Mihulowicz and M. Weglarski, “Definition, Characteristics and Determining Parameters of Antennas in Terms of Synthesizing the Interrogation Zone in RFID Systems,” *in Radio Frequency Identification*, P. C. Crepaldi and T. C. Pimenta, Eds. In Tech, 2017.

[46] J. Zhang, B. Wei, W. Hu and S. S. Kanhere, "WiFi-ID: Human Identification Using WiFi Signal," *International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2016, pp. 75-82.

[47] F. M. Ghannouchi, "Power amplifier and transmitter architectures for software defined radio systems," *in IEEE Circuits and Systems Magazine*, vol. 10, no. 4, pp. 56-63, Fourthquarter 2010.

[48] R. Danymol, T. Ajitha and R. Gandhiraj, "Real-time communication system design using RTL-SDR and Raspberry Pi," *International Conference on Advanced Computing and Communication Systems,* 2013.

[49] J. Yaoyao, L. Byunghun, K. Fanpeng, K. Zhaoping, M. Connolly, B. Mahmoudi, M. Ghovanloo, "A Software-Defined Radio Receiver for Wireless Recording from Freely Behaving Animals," *IEEE Transactions on Biomedical Circuits and Systems*, 13(6): 1645-1654, Dec. 2019.

[50] K.Vasudeva, B.S.Ciftler, A. Altamar, I.Guvenc, "An experimental study on RSS-based wireless localization with software defined radio," *In Proceedings of the WAMICON*; 2014.

[51] T. Zhang, T. Song, D. Chen, T. Zhang, J. Zhuang, “A Wifi-Based Gesture Recognition System Using Software-Defined Radio," *IEEE Access 2019*, 7, 604 131102–131113.

[52] S. Archasantisuk and T. Aoyagi, "The human movement identification using the radio signal strength in WBAN," *International Symposium on Medical Information and Communication Technology (ISMICT)*, 2015, pp. 59-63.

[53] Y. Xu, J. Y. Wang, B. X. Cao and J. Yang, "Multi sensors based ultrasonic human face identification: Experiment and analysis," *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2012, pp. 257-261.

[54] M. N. Alam Nipu, S. Talukder, M. S. Islam and A. Chakrabarty, "Human Identification Using WIFI Signal," *Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 2018, pp. 300-304.

[55] C. Rasmussen and C. Williams, “Gaussian processes in machine learning” *in MIT Press*, 2006, Chapter 2, 4-5, 7-8, pp.7-30, 79-128, 151-185.

[56] A. Bekkali, T. Masuo, T. Tominaga, N. Nakamoto and H. Ban, "Gaussian processes for learning-based indoor localization," *2011 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, Xi'an, 2011, pp. 1-6.

[57] S. Yiu and K. Yang, "Gaussian Process Assisted Fingerprinting Localization," *in IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 683-690, Oct. 2016.

[58] B. Ferris, D. Haehnel and D. Fox, “Gaussian processes for signal strength-based location estimation”, *in Proc. of Robotics Science and Systems*, 2006.

[59] D. Sánchez-Rodríguez, P. Hernández-Morera, J. Quinteiro, and I. Alonso-González, “A Low Complexity System Based on Multiple Weighted Decision Trees for Indoor Localization,” *Sensors*, vol. 15, no. 6, pp. 14809–14829, Jun. 2015.

[60] E. Blasch, A. J. Aved, “Physics-Based and Human-derived Information Fusion Video Activity Analysis,” *Int’l. Conf. on Information Fusion*, 2018.