INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

Departamento de Engenharia de Electrónica e Telecomunicações e de Computadores

Study of physiological and structural variability in the acquisition of vital signs with Bio-Radar

Beatriz Lino Soares

Licenciada

Dissertação para obtenção do Grau de Mestre
em Engenharia Eletrónica e Telecomunicações

Orientadores : Professor Doutor Pedro Renato Tavares Pinho
Professor Doutor Daniel Albuquerque

Júri:

Presidente: Professor Doutor Carlos Eduardo de Meneses Ribeiro

Vogais: Professor Doutor Hugo Plácido da Silva
Professor Doutor Pedro Renato Tavares Pinho

dezembro, 2022

INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA

Departamento de Engenharia de Electrónica e Telecomunicações e de Computadores

Study of physiological and structural variability in the acquisition of vital signs with Bio-Radar

Beatriz Lino Soares

Licenciada

Dissertação para obtenção do Grau de Mestre
em Engenharia Eletrónica e Telecomunicações

Orientadores : Professor Doutor Pedro Renato Tavares Pinho
Professor Doutor Daniel Albuquerque

Júri:

Presidente: Professor Doutor Carlos Eduardo de Meneses Ribeiro

Vogais: Professor Doutor Hugo Plácido da Silva
Professor Doutor Pedro Renato Tavares Pinho

dezembro, 2022

Aos meus pais, irmã e namorado

Acknowledgments

First, I would like to thank my supervisor, Professor Pedro Renato Tavares Pinho, and co-supervisors, Professor Daniel Filipe Albuquerque and Ph.D. Student Carolina Teixeira de Sousa Gouveia, for the excellent guidance, indispensable help, and passing on of knowledge. Their continuous support was crucial for the success and conclusion of this dissertation. Once again, thank you for all the knowledge you have transmitted to me and the trust you have given me.

I would like to thank the Instituto Telecomunicações Aveiro for providing access to the software license, material, and equipment for this dissertation's dataset of vital signs acquisition.

I have to thank my parents for the opportunity they gave me to complete this course and fight for my dreams. I am grateful to them and my sister for their emotional support, and for the fact that they believed in me and my abilities, often more than I did, throughout my journey. I would also like to thank my boyfriend for all the time he listened to me, for the advices he gave me, for the support, and for believing in me.

I would like to thank Dr. Teresa Inácio, who has accompanied me throughout all the years, both medically and personally, advising me in all the decisions that I have taken along my journey.

I am very grateful to all my friends and neighbors for mobilizing my home to help me with the acquisition of the necessary data. Without them, I could not had such an extensive dataset in to help me with the acquisition of the necessary data; Thus, the dissertation would not have been completed successfully. Finally, I would like to thank my friend, Jorge Figueiredo, for all the hours dedicated to other college work, which allowed me to give more attention to my dissertation, and for the patience he had during the most stressful hours in these last two years.

Abstract

The monitoring of vital signals is usually carried out by sensors and electrodes. However, it may not be viable or the best solution for people with burn tissues or with more delicate skin, not to mention cases with infectious diseases, where contact should be kept to a minimum. Thus, vital signs monitoring using radar (*Bio-Radar*) has become a hot topic of research and development.

Several studies state that there is variability in vital signs between people. However, in the *Bio-Radar* area, these issues have not been addressed. In this regard, this dissertation intends to verify if it is possible to evaluate the gender, age, Body Mass Index (BMI), and Chest Wall Perimeter (CWP) through the use of radar signals, namely *Bio-Radar*, used to the vital signs acquisition.

In order to achieve this goal, the vital signs of 92 people (46 females and 46 males), aged between 18 and 50 years old were acquired. With this dataset, it was possible to develop a statistical study of relevant characteristics extracted from the signals. Later, three Machine Learning (ML) algorithms, namely Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and *Random Forest*, were trained to identify gender, age, BMI and CWP.

Finally, the relation between the respiratory amplitude and the respiratory rhythm is analyzed.

Keywords: Bio-Radar, Dataset, Gender classification, Age classification, BMI classification, CWP classification, Machine Learning, Physiological variability, Radar signals, Body stature variability, Vital Signs.

Resumo

A monitorização de sinais vitais é maioritariamente realizada com recurso a sensores e elétrodos. Contudo, pode não ser a forma mais viável ou a melhor solução para pessoas com tecidos queimados ou com pele mais delicada, assim como os casos de doenças infecciosas, onde o contacto deve ser mantido a um nível mínimo. Assim, a monitorização de sinais vitais utilizando radar (*Bio-Radar*) tornou-se num tema de investigação e desenvolvimento.

Vários estudos indicam que existe variabilidade nos sinais vitais entre as pessoas. No entanto, na área do *Bio-Radar*, estas questões não têm vindo a ser abordadas. Neste sentido, esta dissertação pretende verificar se é possível avaliar o sexo, idade, Índice de Massa Corporal (IMC), e Perímetro da Caixa torácica (PCT) através da utilização de sinais de radar, nomeadamente *Bio-Radar*, utilizados para a aquisição de sinais vitais.

Para atingir este objectivo, foram adquiridos os sinais vitais de 92 pessoas (46 mulheres e 46 homens), com idades compreendidas entre os 18 e os 50 anos. Com este conjunto de dados, foi possível desenvolver um estudo estatístico das características relevantes extraídas dos sinais. Mais tarde, três algoritmos de Machine Learning (ML), nomeadamente Support Vector Machine (SVM), K-Nearest Neighbor (KNN) e *Random Forest*, foram treinados para identificar o género, idade, IMC e PCT.

Finalmente, é analisada a relação entre a amplitude respiratória e o ritmo respiratório.

Palavras-chave: Bio-Radar, Sinais radar, *Dataset*, Classificação de género, Classificação de idade, Classificação de IMC, Classificação de PCT, *Machine Learning*, Variabilidade do corpo estrutural, Variabilidade fisiológica.

Contents

List of Figures

List of Tables

Acronyms

AE	Absolute Error
BBI	Beat-to-Beat Interval
BMI	Body Mass Index
BPM	Beats Per Minute
BrPM	Breaths Per Minute
CDF	Cumulative Distribution Function
COPD	Chronic Obstructive Pulmonary Disease
CV	Cross-Validation
CW	Continuous Wave
CWP	Chest Wall Perimeter
DSP	Digital Signal Processing
ECG	Electrocardiogram
FFT	Fast Fourier Transform
FMCW	Frequency-Modulated Continuous Wave
FRC	Functional Residual Capacity
HF	High Frequency
HR	Heart Rate
HRV	Heart Rate Variability

IBI Interbreath Interval
IQ In-phase and Quadrature
IQR Inter-Quartile Range
ISM Industrial, Scientific, and Medical

KNN K-Nearest Neighbor

LF Low Frequency

MAE Mean Absolute Error
ML Machine Learning

OC Optimized Cost

PMI Point of Maximal Impulse
PSD Power Spectral Density

RF Radio Frequency
RMSE Root Mean Squared Error
RMSSD Root Mean Square of Successive Differences
RR Respiratory Rate
RRV Respiratory Rate Variability
RSA Respiratory Sinus Arrhythmia
RV Residual Volume
Rx Reception

SDNN Standard Deviation of IBI
SNR Signal-to-Noise Ratio
STD Standard Deviation
SVM Support Vector Machine

TLC Total Lung Capacity
Tx Transmission

USRP Universal Software Radio Peripheral
UWB Ultra Wideband

WBA-SAR Whole-Body-Average Specific Absorption Rate

1

Introduction

Monitoring vital signs is essential not only in hospitalized patients, but it is also crucial for in-home healthcare, i.e., for bedridden patients. Usually, the most common methods based in contact means using electrodes and sensors. However, non-contact vital signs monitoring has been gaining space in the investigation world. One method used in non-contact vital signs monitoring is based in radar systems, namely the Bio-Radar concept [1].

The Bio-Radar concept was presented for the first time in 1975 as a radar solution for healthcare [2]. Thus, the Bio-Radar technology aims to combine the concept of radar and biomedical measurements to detect vital signs such as breathing and heartbeat without using electrodes or sensors [3]. Over time many studies have been developed for various applications in this area, such as emotion recognition [4–7], sleep sensing [8–12], and driver monitoring [13–15].

In the Bio-Radar technology, all techniques resort to electromagnetic waves to monitor a subject's vital signs; therefore, the challenges that emerged in this area are due to electromagnetic interferences, clutter and random motions. Besides these challenges, which are possible to solve with properly hardware or software, there are characteristics, related with the individuals' physiological and body stature variability that affect the breathing [16], and consequently may affect the received signals.

This dissertation purposes a study about the impact of the physiological and body stature variability in the acquisition of vital signs with non-contact, via *Bio-Radar*, namely gender, age, Body Mass Index (BMI), and Chest Wall Perimeter (CWP).

1.1 Motivation

In this area, many experiences have been developed in subjects. The population in these studies is predominantly mixed. However, the obtained results are not analyzed by gender, and the data related to each person as height, weight, and age, are provided for readers' knowledge.

For this reason, this work will be focused on verifying if there are notorious differences in vital signs acquired by the Bio-Radar system related to gender and body stature. Moreover, it will be evaluated if it is possible to identify the subject's gender, age, BMI, and CWP from only the information presented in the radar signal. In other investigation areas, some studies state that there are indeed physiological differences between genders, such as airways, lungs dimensions [17–21], chest wall movements [22], body stature [23–25], and diaphragm dimensions [23]. These physiological differences might have an impact on respiratory function and consequently on signal acquisition.

In addition to the existing gender differences, other studies indicate that BMI may influence the signals acquired by this system since a higher percentage of body fat is associated with lower lung volumes [26]. Also, aging causes a decrease in the expansion capacity of the chest wall. These facts may lead to lower inhaled air in these cases.

The main goal of this work is to verify studies, in other areas, stated about the existence of physiological differences related to genders, and body statures, are noticeable in the acquired signals. In case they are perceptible, it can bring advantages: at the system level, it is known that algorithms will have to be developed that are as generic as possible; at the instrument level, once it has been able to notice these variabilities, through the signal, it will be possible to study ways of identifying people, diseases or other conditions.

Besides the main goal, this work also aims to categorize people by gender, age, BMI and CWP, through the collected respiratory signals. It is intended to perform this categorization by training ML models. For this reason, and due to the scarcity of public datasets with the required characteristics and information, it is necessary to obtain a large as possible dataset.

1.2 Objectives

Knowing all the challenges and motivations of this dissertation, the main objective of this dissertation is to conclude about the possible impact of physiological differences

between genders or body stature in Bio-Radar vital signs acquisition. In this way, the specific objectives of this dissertation are:

- Collect vital signs from a wide group of subjects. This process will be carried out using the Bio-Radar prototype developed in [27]. In this step, it is essential to balance the population of volunteers that will be participating in this study;
- Analyse the data collected in the previous step. This analysis will take into consideration information about each subject, such as height, weight, age, and chest wall perimeter;
- Conclude about the previous data analysis and the possible impact of all the physiological differences;
- Extract features from the radar signal and analyse them statistically;
- Distinguish the gender, age, BMI and CWP group of the participants, through ML, giving as input, the features statistically significant.

1.3 Document Organization

This section describes the dissertation's structure and each chapter's content. The dissertation is divided in the following 6 chapters:

- **Chapter 1** - The current chapter reveals the identified challenges that led to motivations for this study and work. Furthermore, all the objectives established for the success of the dissertation are summarised here;
- **Chapter 2** - This chapter summarises how the respiratory and cardiovascular systems work, their movements and main muscles associated with them. In addition, some physiological differences between genders, aging, BMI are presented at the lungs, chest wall, and diaphragm levels. To close the chapter, there is a section dedicated to the non-contact acquisition of vital signals, where the three main Bio-Radar operating modes are described, as well as the choice of which will be used for the practical part of the dissertation. Also, this section describes the chosen Bio-Radar prototype in detail. To conclude, a survey of the studies developed with the same operating mode is carried out to highlight the innovation presented in this dissertation;
- **Chapter 3** - The preliminary study of the Bio-Radar model and the signal processing is presented in this chapter, beginning with the DC offset estimation and

its removal, and then the determination of the respiratory signal amplitude and RR calculation;

- **Chapter 4** - This chapter is dedicated to the experimental part of the dissertation. First of all, the scenario where the acquisition of the vital signs took place is described, as well as the set-up assembled to perform them. After that, a brief description of the achieved dataset is made. In this chapter it is also described the complete analysis of all the Bio-Radar signal features.

- **Chapter 5** - In this chapter the application of ML are explained in detail. Furthermore, the results of the application of the ML, i.e. the possibility of gender, age, BMI and CWP classification are revealed;

- **Chapter 6** - In this last chapter, conclusions about the developed work are presented, and future improvements are pointed out.

1.4 Contribution from the developed work

Considering the state of the art presented in this dissertation, there are some important contributions:

- Machine learning algorithms were applied using the acquired vital signs and it was possible to identify the subject's gender, age, BMI, and CWP with an accuracy of 76.66%, 71.13%, 72.52%, and 74.61%, respectively.

With the developed work from this dissertation, it was submitted the following papers:

- Beatriz Soares, Carolina Gouveia, Daniel Albuquerque, and Pedro Pinho. "Impact of Body Stature and Physiological Variability in the Acquisition of Vital Signs using Continuous Wave Radar", IEEE Sensors.

It was also published the following paper:

- Beatriz Soares, Carolina Gouveia, Daniel Albuquerque, and Pedro Pinho. "Impact of the Human Variability in the Bio-Radar Signals", 16th Congress of the Portuguese Committee of URSI "Quantum communications: what future?", Lisbon, Portugal 2022.

At last, this project won the best project in the "Social Responsibility" category award, 16th Congress of the Portuguese Committee of URSI "Quantum communications: what future?", Lisbon, Portugal 2022.

2

Vital Signals Acquisition with Radar

2.1 Vital Signals

Throughout the following chapters, core topics of the work herein described are physiological and body stature variability in acquiring vital signs with Bio-Radar. Therefore basic knowledge of the respiratory and cardiovascular systems is indispensable. In the present section, a brief introduction is done regarding the cardiovascular and respiratory systems, focusing on their principles and how they can be detected at the skin surface by the radar.

2.1.1 Cardiovascular System

The cardiovascular system comprises the heart, blood vessels, and blood. Its primary function is transporting nutrients and oxygen-rich blood to all body parts and carrying deoxygenated blood back to the lungs [28]. The heart pumps blood through the lungs to all body tissues. When the heart muscles relax (diastole) and contract (systole), it generates a pressure change that drives blood. The heart is located in the chest cavity, and during the cycles, the size of the heart changes, causing biomechanical activity at the chest wall that can be detected in the skin surface by non-contact sensors [29].

Usually, in healthy patients, the biomechanical activity of the right ventricle cannot be

detected through palpation. The movement that causes a more significant displacement, perceived at the chest wall surface, is from the left ventricle [30]. While contracting, the heart becomes spherical, and its diameter increases, causing more pressure against the chest. Furthermore, it rotates, making the lower frontal part of the left ventricle reach the chest wall front [31]. Thus, in healthy subjects, the impulse caused by the left ventricle can be easier detected through the Point of Maximal Impulse (PMI) (the point where there is a maximal impulse against the chest). This point is located above the anatomical apex, between the fourth and fifth intercostal spaces of the left mid-clavicular line (Figure 2.1) [32].

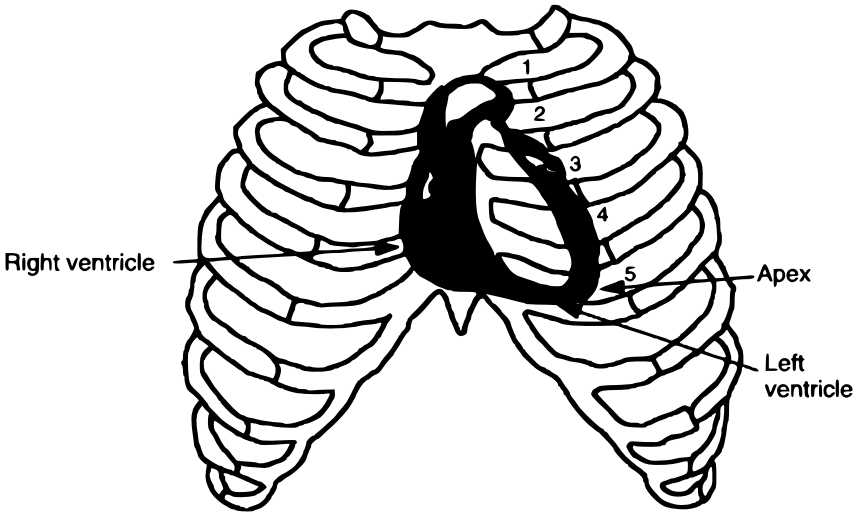


Figure 2.1: Heart location on rib cage [29].

2.1.2 Respiratory System

The respiratory system's primary function is the exchange of the gas allocated in the lungs. In other words, it is the exchange between the carbon dioxide and oxygen inhaled. In addition to this function, the respiratory system is also responsible for maintaining the balance of the acid-base levels [33].

Typically, the level of ventilation is measured with the respiratory rate and the tidal volume. The organ systems, like the nervous system, the cardiovascular system, the respiratory system, and the excretory system, can cause alterations in the respiratory system's function, meaning that the respiratory system is a mirror of the health state of the whole body. The respiratory rate, rhythm, regularity, depth, and volume can indicate the imbalance and the respiratory disorders. Some of these disorders, when continuously monitored, can be avoided. Below are presented a few examples of such disorders [30]:

- Respiratory distress syndrome;
- Pulmonary edema and embolism;
- Pneumonia;
- Chronic Obstructive Pulmonary Disease (COPD);
- Severe heart failure;
- Low blood volume.

Throughout the respiratory cycle, muscles contract to produce changes in the volume of the thorax. This causes pressure differences between the thorax and the external environment, moving the air inside or outside the lungs, going from a higher pressure level to a lower one. The thorax and the abdomen motion causes displacements at the skin surface, which are measurable by a radar [30]. Figure 2.2 illustrates the location of the muscles associated with breathing.

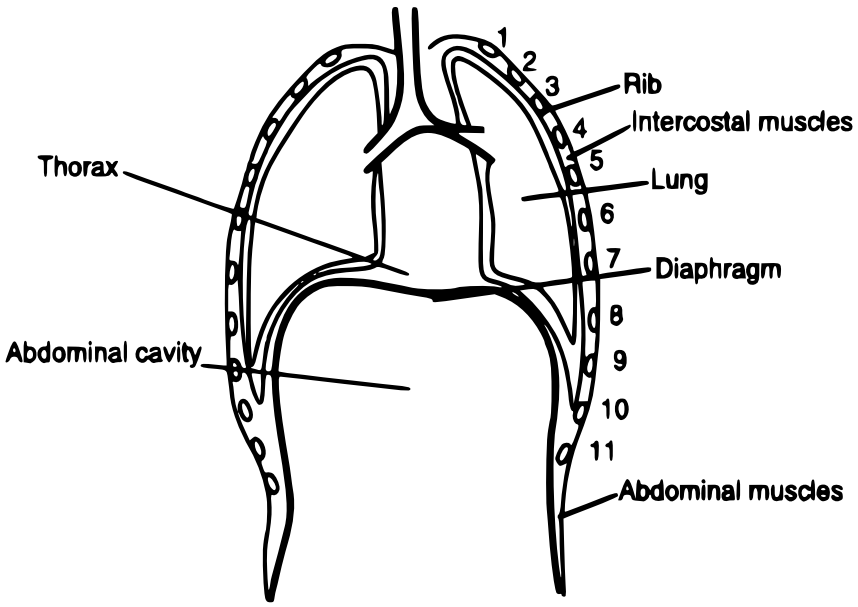


Figure 2.2: Location of respiration muscles [34].

During normal breathing, no muscles are contracted for expiration, except when the subject is speaking or singing, i.e., in quiet breathing, there is no implied effort for expiration. In inhalation, the diaphragm contracts, the thorax elongates, and the volume increases, pushing the abdominal area forward. Figure 2.3 shows the behavior of the thoracic cavity during the breathing process (inhalation on the left and expiration on the right). In light inhalations, the diaphragm extends 1-2 cm, while in deep inspiration it can reach 10 cm [35, 36].

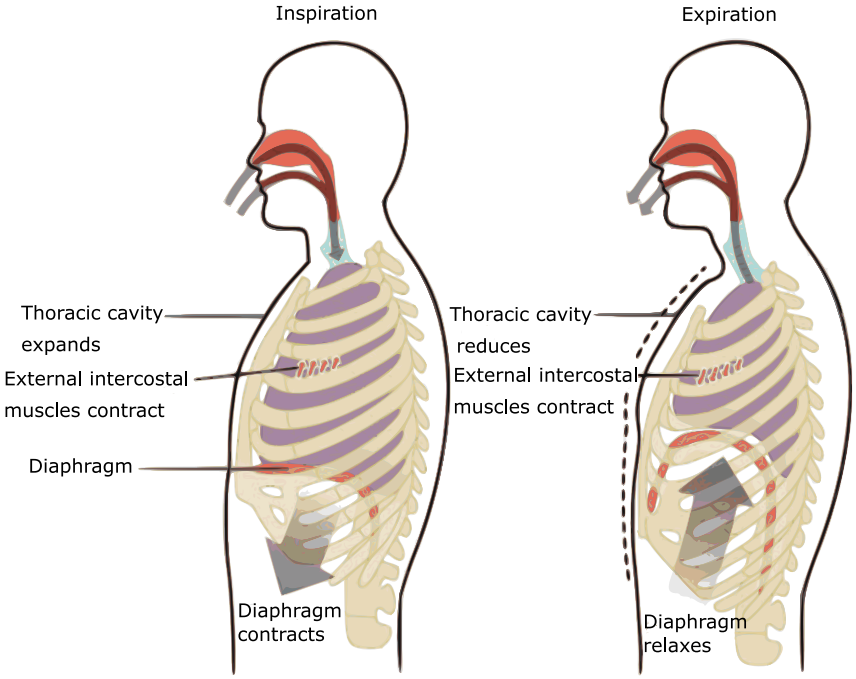


Figure 2.3: Thoracic cavity during the breathing process [37].

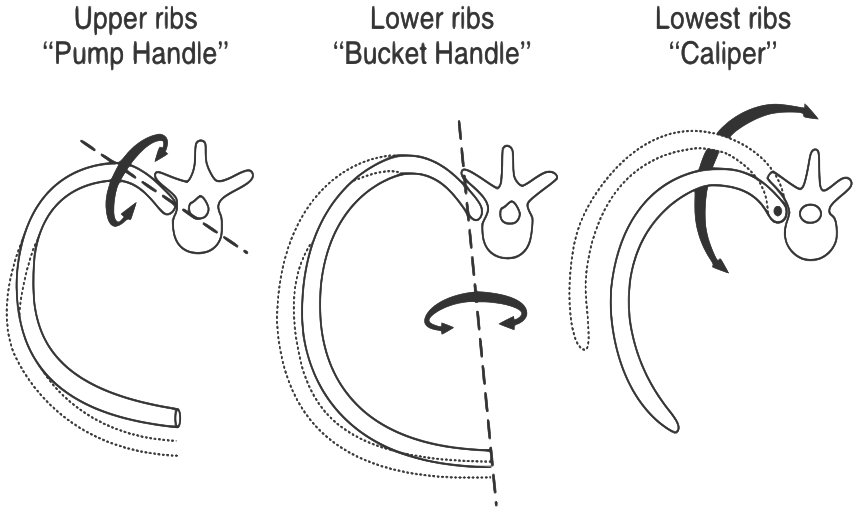


Figure 2.4: Rib cage motion during the breathing process [34].

There are three types of rib movement at different points in the rib cage: the "pump-handle" motion of the upper ribs, the "bucket-handle" motion of the lower ribs, and the "caliper" motion of the lowest ribs [34]. In Figure 2.4 it is represented the rib cage motion characterization:

- Pump Handle: Consists in an upward rotation of the upper ribs around their axis [36].
- Bucket Handle: Lower ribs can glide and rotate because they connect to the spine differently from the upper ribs [36].

- Caliper: Lowest ribs displacement occurs because they are not connected to the sternum [36].

2.1.3 Physiologic Gender Variability

The physical differences between men and women are evident. Besides these, the physiologic ones are also notable. The respiratory function is directly related to gender differences, specifically, with physiologic differences between both (men and women) [17]. In this subsection, it will be presented some differences between genders. They might impact respiratory function and, consequently, radar signals acquisition.

2.1.3.1 Lung Differences

Firstly, this study will focus on the impact of gender on the lungs. The difference between gender at the physical level during the respiratory cycle starts immediately with the difference in the size of the nasal cavity, where males have larger nasal cavities [17]. Their lungs are bigger, not only in terms of volume, but also in terms of their volume variations [18–21]. Furthermore, the respiratory rate is slightly higher in women than in men [18]. Men have significantly higher mean values for all lung variables, both lung volumes and maximal flow rates, except for resistance which is considerably lower in men, according to [38, 39].

Men have bigger lungs than women and that fact has been shown using different approaches: standard morphometric methods [40], chest radiographs [23] and three-dimensional geometric morphometric methods on computed tomography scans [22]. The lung forms differ between males and females, being more pyramidal in the male's cases and more prismatic in the female's cases [22]. Sex differences are also present in the volume and configuration of the rib cage. Women are characterized by a disproportionately smaller rib cage size than men. Specifically, the cross-sectional area, the anterior-posterior internal, and the lateral diameters are lower in different lung volumes [23, 41].

However, in literature, there are evidences that the differences in pulmonary function (namely lung volumes, maximal expiratory flow rates, diffusion surfaces, and maximal pulmonary ventilation) between females and males are primarily due to the smaller height and trunk size in women [41].

2.1.3.2 Chest Wall Differences

In terms of Total Lung Capacity (TLC), women present a more rounded rib cage than men. The different thoracic configurations in women are also shown by a different ratio between rib cage cross-sectional area and diaphragm dome height [23]. The higher rib cage dimensions in men are obtained in radiographs, among other techniques, as mentioned before (lungs dimension). With infrared imaging techniques, it has been proved that the male rib cage has a higher anteroposterior diameter and larger perimeters, cross-sectional area, and volume [24]. In addition to being larger, the ribs of males are also deeper than the ribs of females with the same height, and this is linked to a larger rib cage volume in males [25]. In women, the movement of the ribs is predominantly "pump-handle", and in men is "bucket-handle", as suggested by [22].

In [42], a study was developed about the thorax and abdomen movement distances in 100 healthy subjects (50 males and 50 females). The population dimension is relatively large and complete, so it is possible to draw generalized conclusions about the impact of gender on respiratory movement. Nonetheless, the state of the art of [24] study is controversial, i.e. while some authors [43, 44] reported a relatively greater rib cage motion in women, others as [45, 46] did not. The results refer that upper thoracic movement was significantly decreased with age, with females exhibiting slightly more rib cage movement during deep breathing. Also, in [24], it is possible to validate the previous statement, where female subjects were characterized by smaller dimensions of the rib cage compartment and during quiet breathing by lower tidal volume (amount of air that moves in or out of the lungs with each respiratory cycle).

2.1.3.3 Diaphragm Differences

Relatively to the diaphragm, in both Residual Volume (RV) and TLC or Functional Residual Capacity (FRC), the length of the diaphragm is approximately 9% shorter in females than in males [23]. The results of [42] also refer that there was less abdominal movement in females than in males. In deep breathing, males showed predominantly more diaphragmatic breathing when compared with females.

2.1.3.4 Body Mass Index Differences

Regarding BMI, it is common knowledge that weight increase leads to a BMI increase. Weight affects respiratory parameters as it causes short airway dysfunction, limitation of respiratory flow, respiratory muscle strength, decreased pulmonary gas exchange,

and decreased respiratory control [47]. In [47] the authors also refer decreasing of lung expansion if fat deposits over the diaphragm, abdomen, and intercostal muscles. Other studies mention that higher perceived body fat is associated with lower lung volumes [26, 48]. These characteristics can affect a received signal because it may have a lower amplitude when BMI has a higher value.

2.1.3.5 Aging Impact

Relatively to aging, [49] proves that the body stature changes to the thoracic cage cause chest wall compliance. The same study reveals that stiffening of the thoracic cage from calcification to the rib cage and age-related kyphosis [1] from osteoporosis reduces the ability of the thoracic cage to expand during inspiration and places the diaphragm at a mechanical disadvantage to generate effective contraction [49]. In sum, and as [51] refers, all the components of the respiratory system are affected by aging, such as the lung elastic recoil decreasing, the muscles losing strength, and the respiratory center being less sensitive. These characteristics can affect a received signal, as it can have less amplitude; this means that the received signal may be weaker and more irregular.

2.1.3.6 Discussion

In summary, men have most parts of trunk dimensions bigger than women [17, 22–24]; this means that men may have more space to accommodate lungs during inhaling. Moreover, men have higher tidal volume for the same weight than women, which implies that the movement of the men's rib cage may be greater than women with the same weight [24]. As a result, during the analysis of the acquired signal, a greater amplitude may be observed in the case of men compared with women.

The study done in [52] compares the power absorption of females and males, where it is preferred that the Whole-Body-Average Specific Absorption Rate (WBA-SAR) in the band 0.1 GHz to 4 GHz increase with the subcutaneous fat. Considering that women may have more subcutaneous fat than men in the thorax, this means that at these frequencies, females may present more absorption in this zone. Consequently, the reflected wave may have a lower amplitude in females than in males. However, the prototype used in this work operates at a higher frequency (5.8 GHz). Hence, the reflection of the incident wave in the chest wall will be greater [53]. This way, the difference in the reflected wave amplitude between females and males might decrease.

[1] Kyphosis is an exaggerated, forward rounding of the upper back [50]

Also related to subcutaneous fat is the BMI, relating also with the study, where it states that the respiratory amplitude is lower the higher the BMI value is [47]. So, it is expected that the respiratory signal has a lower amplitude in individuals with higher BMI.

With advancing age, all respiratory functions are affected, as is the expansion of the rib cage on inspiration, the lung elastic recoil decreasing, the muscles losing strength, and the respiratory center being less sensitive [49, 51]. Thus, breathing gets weaker and weaker with advancing age.

2.2 Remote Detection of Vital Signals Using the Bio-Radar

Given the recent events, the idea that we can be infected by airborne viruses changed the way we live. This fact catalyzed the investigation of the acquisition of vital signs, such as heart and respiration rates of a human subject, with non-contact techniques.

The radar is a device that allows object detection and location by transmitting an electromagnetic wave and detecting the echo reflected by the target, such as an aircraft or a body. Military forces secretly used the first radar system in the Second World War to track targets over long distances. Over time, radar technology has been applied in other areas of interest, such as healthcare. The Bio-Radar concept has not emerged recently. It was presented for the first time in 1975 as a radar solution for healthcare [2]. Thus, the Bio-Radar technology aims to combine the concept of radar and biomedical measurements to detect vital signs such as breathing and heartbeat without using electrodes or sensors [3].

2.2.1 Radar Operation Description

The Doppler Radar operates according to the micro Doppler effect. It detects moving targets and can compute their velocity, one of the most common applications being traffic control. The Doppler effect is the apparent difference between the received wave frequency and the transmitted wave frequency, which is observed whenever the target is moving to the radar [30].

Radars with different operation modes can be used to capture vital signs from the human chest surface. In this system it is possible to have three operation modes: Continuous Wave (CW), Frequency-Modulated Continuous Wave (FMCW) and Ultra Wideband (UWB) [54]. Figure 2.5 represents the three operation modes mentioned later in the document.

Since the radar system is stationary, the received frequency is the same if the subject is also static since the same number of Wave/s is received. However, assuming that the target moves towards or away from the radar, different frequencies are received: in case the movement is towards the radar, a higher frequency is received due to the higher number of received Wave/s. In contrast, if the target goes backward, a lower frequency is perceived due to the lower number of Wave/s received [30].

Knowing that each wavelength corresponds to a phase change equal to 2π radians, the frequency shift effect can also be measured through a phase change.

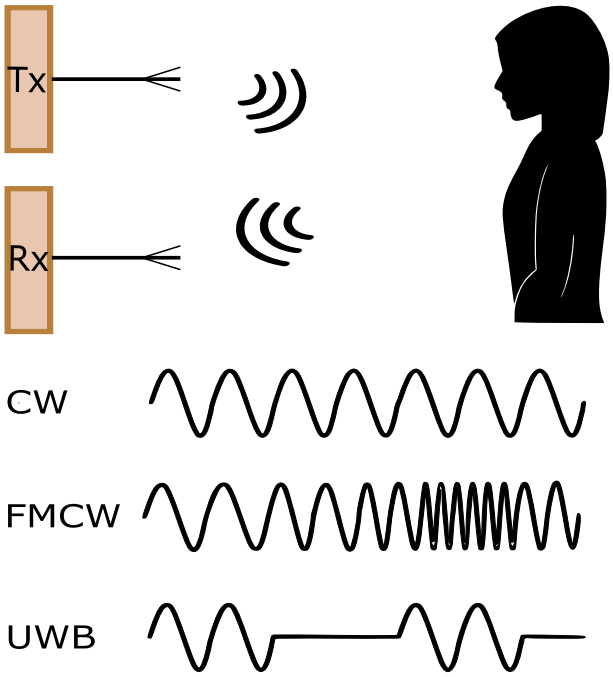


Figure 2.5: Bio-radar with different operation modes.

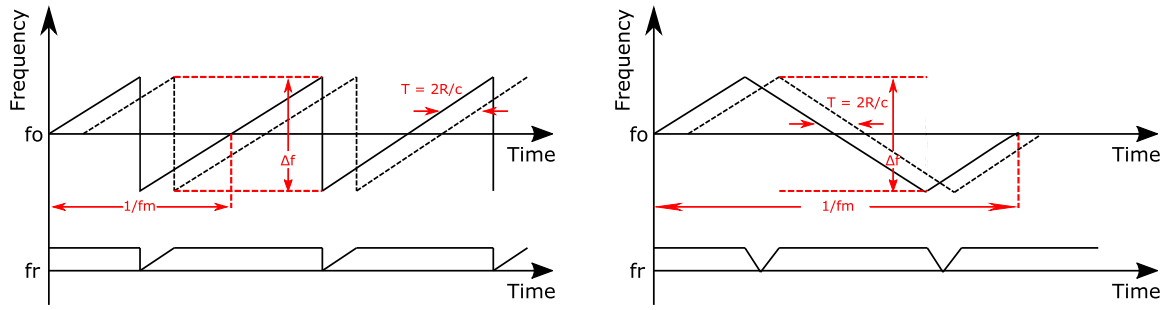
2.2.1.1 Continuous Wave

The CW radars transmit a continuous signal with the same frequency. The radars with CW mode are known for their simplicity, so they are the most common type. A typical radar transmits and receives a Radio Frequency (RF) continuous signal with a narrow bandwidth. A system based on this operation mode is composed of a signal generator, which generates the signal that will be transmitted and down-converts the received signal at reception. Since the system usually is handled with narrowband signals, it is possible to perceive the frequency shift due to the Doppler effect when the target is moving. In that way, the radar can measure the velocity of the target's motion, and this characteristic enables one to distinguish between a moving target and stationary

objects independently of its distance (considering the radar's range), or its velocity [55].

2.2.1.2 Frequency-Modulated Continuous Wave

Still within the context of CW radar, the FMCW transmits a continuous signal, which is frequency modulated. The FMCW allows the computation of the target's velocity and the distance between the target and the radar. This type of radar improves the range resolution, which is not the case with CW [56]. There are two best-known forms of frequency modulation: linear and triangular. In this way, in both cases, the frequency increase and decrease linearly over time. However, the frequency decrease in linear modulation is instantaneous, as Figure 2.6a illustrates, contrarily to the Figure 2.6b. In both cases, represented in Figure 2.6, the transmitted signal is represented by the solid triangular waveform, and the dashed curve represents the received echo signal from a stationary target. The second curve is delayed concerning the first $T = \frac{2R}{c}$ seconds (where R is the distance between the sensor and the target, and c is the wave propagation speed). Based on Figure 2.6, the bandwidth of the modulated signal is equal to Δ_f, and it determines the precision of the range measurement. Thus, f_m is the rate at which modulation is done, and it determines the maximum detectable range without ambiguity. As a result of the relation between the transmitted and the received echo signal is a frequency difference, represented by f_r, it is expected that it changes according to target movements [30].



(a) Linear frequency modulation.

(b) Triangular frequency modulation.

Figure 2.6: Frequency–time relation in an FMCW radar with linear (a) and triangular (b) frequency modulation [30].

2.2.1.3 Ultra Wideband Radar

The UWB Radar is the most common and used derivation of pulsed radar. In contrast to the CW and FMCW examples presented above, the UWB radar transmits a pulse, so it has to transmit and receive in different moments. Consequently, the Tx and Rx do not operate simultaneously; this implies that the time for the pulse to repeat must be longer than the travel time of the wave transmitted by the radar and the echo reflected

on the target [36].

Because reflections are not acquired during signal transmission, transmitter leakage and parasitic reflections are separated temporarily from long-range targets. This means that reflections that occur immediately from short-range objects are not detected since waiting for the full transmission is necessary. This radar can measure the target distance through the received pulse delay [36].

2.2.2 Comparison

In Sections 2.2.1.1, 2.2.1.2, and **??**, three operation modes were described. In this section, their advantages and disadvantages will be discussed to refute the choice of the radar type used in this dissertation.

Since the FMCW and the UWB radars use a aconsiderable amount of bandwidth, they can have range isolation between the target and the clutter; this means that CW radars are not able to estimate the target distance and discriminate the target in noisy environments. On the other hand, the range resolution of the CW radars is limited just only by the system noise, whereas in the FMCW and UWB radars is also constrained by the used frequency bandwidth [30].

As mentioned in Section 2.2.1.3, in UWB radars, the interval between the transmission and the reception must be longer than the travel time of the wave transmitted by the radar, and the echo reflected on the target. This implies that in case the target is positioned at different distances, the interval between the transmission and the reception needs to be adjusted [57].

Both radars, CW and FMCW estimate vital signs using phase measurements, which can be seen as an advantage for this type of radar since they are able to determine the target's and object's velocity, and distinguish between a moving target and stationary objects [55, 56]. Concerning the CW radar, this system is simple to implement because it has a single oscillator for the transmitter and receiver. Considering that this radar estimates the vital signs using phase measurements and that it could provide high precision results [30], it will be an advantage for this application. Additionally, the CW is the only one that is able to operate in Industrial, Scientific, and Medical (ISM) band, by the fact that FMCW, and UWB radars should enforce compliance with emission masks issued by regulatory bodies [57]. This way, the radar used for the experimental studies in this dissertation is the CW.

2.3 Vital Signals Acquisitions Using CW

This section summarizes some studies on acquiring vital signals using the CW radar over time. A brief description of the prototype used later in this work, specifically the hardware that constitutes it, is also provided.

2.3.1 CW Radar Constitution

As mentioned before, the Bio-Radar prototype used in this dissertation was previously developed in [27], and its operating principle is depicted in Figure 2.7. This system is constituted by a CW Doppler radar which continuously transmits a sinusoidal carrier that is generated digitally and receives the reflected signal by the target. As a result of the Doppler effect, there is a phase shift as the subject's chest wall moves towards or away from the radar, creating a phase modulation in the received signal [30]. The main purpose of this system is to estimate respiratory and heart signals. This information can be obtained through the phase extraction of the received signal.

The signals from the subject have low amplitude, with their bandwidth occupying short ranges close to DC. They are susceptible to several noise sources, such as clutter from the scenario reflections. The Digital Signal Processing (DSP) Module executes an algorithm developed for the rigorous extraction of information from bio-signals. Considering these characteristics, the signals system is defined as a scenario in which the only disturbance source is the parasitic reflections from nearby standing objects. Figure 2.8 represents the implemented DSP algorithm [58]. After applying a low-pass filter to the acquired signal, the resultant signal is decimated to reduce its original sampling rate to a lower one. DC offset removal is the next step in the DSP module. Considering that an arc represents the received signal, in an ideal scenario, its center should be on the referential origin of the polar plot, and the DC offset is the deviation of the arc center from the referential origin. Therefore, it should be compensated. Lastly, to represent the received vital sign in the time domain, it is mandatory to apply the arctangent demodulation [59]. Finally, the Recovered Signal Module displays in real-time the respiratory signal and computes its rate [27].

The carrier frequency selection is an important topic because it affects the fundamental operation of the system. It also influences physical characteristics in the transmitting environment, such as the antenna size, the range for subject detection, and resolution. The usage of high frequencies brings advantages, such as an increase of the Signal-to-Noise Ratio (SNR) and the possibility to use more compact and portable radar modules

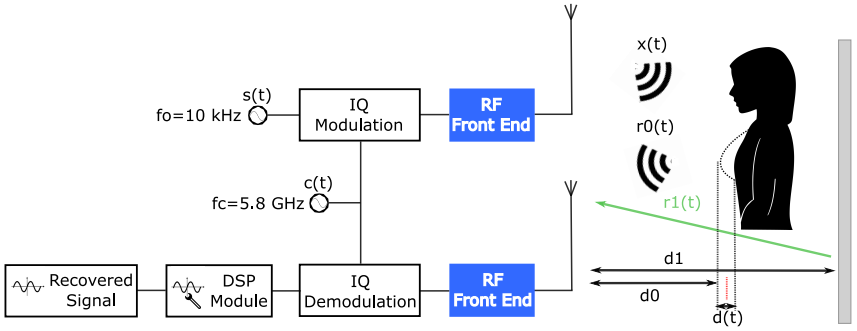


Figure 2.7: Block diagram of the bio-radar system, adapted from [27].

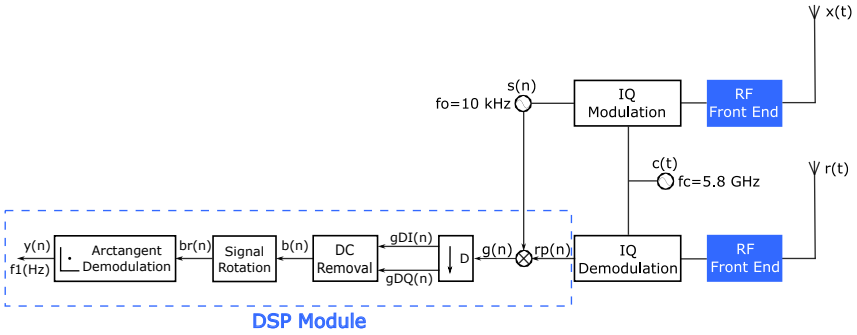


Figure 2.8: Block diagram of the bio-radar system including DSP module, adapted from [60].

(smaller size antennas can produce the same gain and directivity) [30]. However, carriers with higher frequencies make it challenging to measure breathing and heart rate simultaneously [27] because, with higher frequencies, the arc length increase and consequently can become a circumference. The chosen carrier frequency is 5.8 GHz and belongs to one unlicensed band. The system uses two 2×2 patch array antennas, one for transmission and the other for signal reception [27], as Figure 2.9 illustrates.

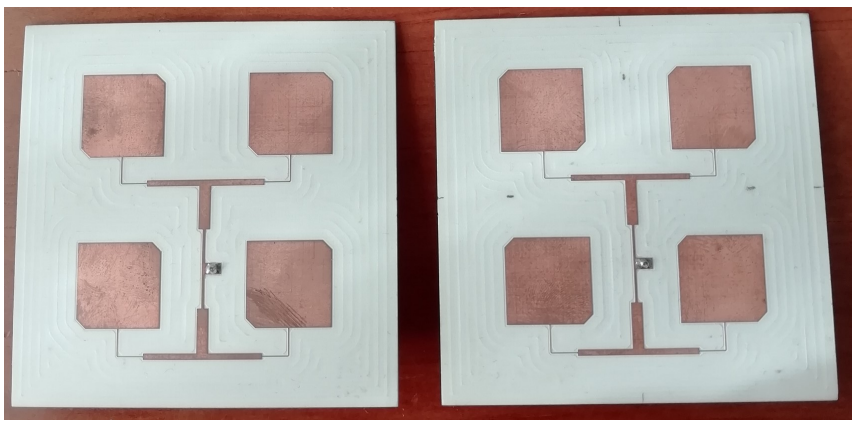


Figure 2.9: Patch antennas - Tx and Rx.

The main hardware module of this prototype is the USRP B200 platform, which is

represented in Figure 2.10, with a continuous frequency coverage from 70 MHz – 6 GHz [61]. Through this platform, it is possible to program the characteristics, such as the power of the transmitted wave or its sampling frequency, of the Bio-Radar.

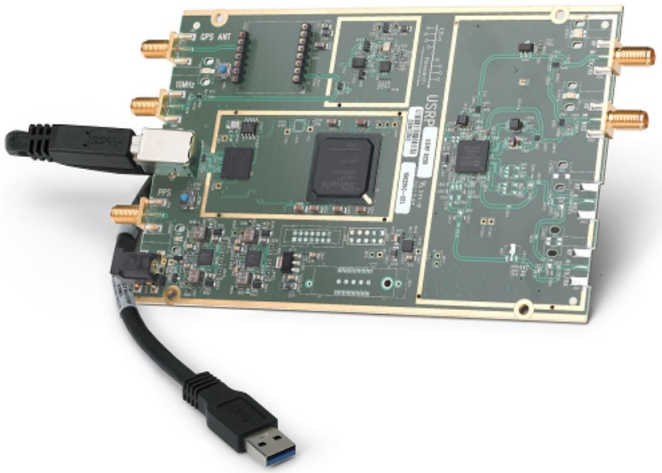


Figure 2.10: USRP B200 platform [61].

2.3.2 State of the Art Review

Numerous experimental studies on acquiring vital signs with Bio-Radar have been presented over time. In this section, a brief descriptive summary of some studies is developed, describing how the experiment is carried out, the frequency of the Bio-Radar, the population under study, and the results obtained. The Bio-Radar can be helpful in different applications, such as sleep monitoring, emotion recognition, child occupation detection in the vehicle interior, heart tracking, and general vital signs acquisition. Of all the CW radar studies read for developing this state of the art, twenty-five were selected, shown in Table 2.1. Of the selected studies, three do not refer to the number of subjects analyzed, four make no distinction between genders, and only one concludes about the impact of gender on the results.

Going back in time to 2002, in [62], a system was developed to detect vital signs. This system is based on the motion sensing principle of the Doppler effect. The experimental part was measured by just one person, using CW, operating at 10 GHz.

In 2009, in [63], a system was developed to measure the heartbeat and respiration rate. In single-channel mode, these measurements have a mean error of ± 3 beats/min. On the other hand, with arctangent demodulation and autocorrelation (the method found to make the system more accurate and reliable) was obtained ± 1 beat/min, in an In-phase and Quadrature (IQ) receiver configuration. This conclusion was made using an

experiment, with unknown population size, and without reference to the gender of the participants, where they would be 1.5m or less away from the CW Doppler radar at a frequency of 2.4 GHz [63]. Also, in 2009 an investigation was presented on the Heart Rate Variability (HRV), and the Respiratory Sinus Arrhythmia (RSA) indices using data obtained from Doppler radar cardiopulmonary remote sensing [64]. This method was validated using data obtained from 12 subjects in seated and supine positions. Results from radar correlated well with the ones obtained through respiratory piezoresistor chest belts. In this study, the experimental measurements were made in 12 subjects using 2.4 GHz [64].

Moving forward a few years, in 2012, a radar signal processing method was introduced to extract the respiration and heart rates from a single person [65]. To prove the method's validity, vital signs were acquired from stationary males by a CW Doppler radar unit. The sensor operating at 24 GHz was located 0.5 meters away from the subject (in this study, the vital signs of 2 males were measured) [65].

In 2014, a sensor was developed to reconstruct the chest wall movement caused by cardiopulmonary activities accurately [66]. An algorithm was developed to estimate the respiratory and heartbeat rate and some indicators of HRV. The sensor system included a CW radar operating at 5.8 GHz. It was tested on 10 volunteers, 5 males, and 5 females. However, the results did not refer to the differences between vital signs acquisitions on both genders [66].

One year later, in [67], a contactless sleep sensing system was presented that continuously and unobtrusively tracks sleep quality. During the experiment with 8 participants (once again, unknown gender), tests were performed with five sensor orientations: in front of the bottom of the feet, in front of the chest, in front of the side of the torso (which has a lower Mean Absolute Error (MAE)), in front of the top of the head and front of the back. The sensor (CW radar) transmits a 24 GHz signal over a distance range of 0.5 m to 2 m. In the same year, a Doppler radar sensor system was developed for vital signs detection, and activity monitoring inside a RF shielded room [68]. Throughout the experience with a CW Doppler radar operating at 34 GHz, the ten male volunteers were in four different positions, performing different actions: walking, standing, and lying [68].

In 2016, a time-window-variation technique was developed for fast acquisition of Heart Rate (HR) [69]. This technique appeared to solve the issue of insufficient spectrum resolution and respiration harmonics. The experiments were performed on four subjects under controlled laboratory conditions, by a 5.8 GHz CW Doppler radar [69].

One year later, a wavelet-transform-based data-length-variation technique was proposed to realize the fast detection of HR [70]. This study uses the CW radar and the same frequency as in the previous study. In both experimental studies, the number of subjects and their gender are unknown [70].

Moving to 2018, heart sound detection by radar systems was introduced, which enables a touch-free, and continuous monitoring of heart sounds [71]. A CW radar operating at a frequency of 24.17 GHz was used for the experimental tests, with seven male volunteers and four females. This study concludes that using the detected heart sounds considerably improves radar-based heartbeat monitoring, and the achieved performance is competitive compared to phonocardiography. In these results, nothing is concluded about gender variability [71]. Also that year, in [72], a method for estimating the Beat-to-Beat Interval (BBI) based on Doppler sensors, using the precise peak detection method. Through the experiments (using a CW bio radar at a frequency of 24 GHz) on ten subjects, in the cases where a subject was sitting still, typing, and speaking, it was confirmed that the proposed method improved the previous state-of-the-art [72].

In 2019, a 915 MHz CW Doppler Radar sensor for detecting vital signs were proposed. Six participants, one female and five males, were involved in [73] study. The authors concluded that the proposed radar sensor is insensitive to surface movements caused by respiration and heartbeat due to the operating frequency being 915 MHz. However, path-loss calculation at 915 MHz and 2.45 GHz indicates that the received power by the radar sensor, which includes movement inside the human body, is higher operating at 915 MHz than at a higher frequency [73]. Also, that year, an accurate method was proposed for detecting vital signs obtained from a CW radar sensor, which operates at 2.45 GHz. Experimental results with six subjects (2 males and four females) show that the proposed method can be used to obtain the heart rate with high accuracy. However, we cannot obtain the information for an HRV analysis due to the intrinsic characteristics of the radar sensor [74].

In 2020, a dataset of clinically recorded radar vital signs with synchronized referenced sensor signals was presented, and it consists of twenty-four hours of synchronized data from radar and reference devices [75]. The implemented CW radar operates at 24 GHz and is focused on the chest while the subjects are lying on a tilt table. Thirty healthy subjects were measured, 14 males and 16 females [75]. In the same year, a radar-recorded heart sounds and vital signs dataset, including synchronized reference sensor signals, was also presented [76]. Similar to the studies described above, this study is supported by experiments in eleven subjects (seven males and four females), measured in different defined scenarios and at several measurement positions such as

at the carotid, the back, and several frontal positions on the thorax [76]. The authors stated that the exact anatomy of every person is a little different (i.e., the exact location of the heart inside the chest), and the different fat distributions impact the measurements. They also refer that, since the radar illuminates a relatively large area, the exact alignment of the antenna does not play an essential role.

Also, that year, a review of vital human signals detection methods and potential was done using radars [77]. The experimental study was carried out using CW radar operating at a frequency of 10 GHz and with two volunteers, one female, and one male. The presented results are given for the male and female subjects, respectively. A breath rate of 18 Breaths Per Minute (BrPM) and a heart rate of 58 Beats Per Minute (BPM) were obtained for the male subject. Moreover, for the female subject these results were 20 BrPM and 78 BPM [77].

A new simple, effective, and high-accuracy settlement to maintain cardiac activity sensing comprehensively and systematically for the first time, for the estimation of BBI and cardiac timings, was presented in the middle of this year. In this study, a CW radar operating at 24 GHz was used to measure a population of three males and three females [78].

Table 2.1: Several studies in the area of Bio-Radar.

Reference	Population	Gender distinction in result	Respiratory rate	Heart rate	Frequency [GHz]	Error	RR Error	HR Error
[62]	1 S	-	X	X	10	-		
[63]	U	-	X	X	2.4	ME	-	$\pm 3/\pm 1$ [beats/min]
[64]	12 S	-	X	X	2.4	-		
[65]	2 M	-	X	X	24	-		
[66]	5 M; 5 F	No		X	5.8	RE	-	2.53 - 4.83 [%]
[68]	10 M	-	X	X	34	-		
[69]	4 S	-	X	X	5.8	E	-	< 5 [%]
[70]	U	-		X	5.8	ME	-	3.5 [%]
[71]	7 M; 4 F	No		X	24.17	E	-	44.2 [ms]
[72]	8 M; 2 F	No		X	24	-		
[73]	5 M; 1 F	No	X	X	0.915	MER	0.27 [%]	2.39 [%]
[74]	2 M; 4 F	No		X	2.45	E	-	3.22 [%]
[75]	14 M; 16 F	No	X	X	24	RMSE	-	25 [ms]
[76]	7 M; 4 F	No	X	X	24	RMSE	-	47.85 [ms]
[77]	1 M; 1 F	Yes	X	X	10	-		
[78]	3 M; 3 F	No		X	24	MRE	-	0.51 - 1.06 [%]
[79]	50 NB	-	X	X	24	-		
[80]	14 M; 21 F	No	X	X	24	-		
[81]	1 M; 4 F	No	X		2.4	-		
[82]	6 M; 4 F	No	X	X	2.4	MAE	3.5 [beats/min]	0.8 [beats/min]
[10]	1 M	-	X	X	2.4	-		
[83]	U	-	X	X	2.4	-		
[67]	8 S	-	X	X	24	MAE	1.18 [breaths/min]	3.29 [beats/min]

M - Male, **F** - Female, **S** - Subject, **U**[1] - Undefined, **NB** - Newborn, **ME** - Mean Error
RE - Relative Error, **MER** - Mean Error Rate, **RMSE** - Root Mean Square Error
MRE - Mean Relative Error, **MAE** - Mean Absolute Error
[1] The study performs tests on subjects, however the number or gender is not specified

The studies mentioned above are summarized in Table 2.1. Mostly the presented studies summarized in Table 2.1 do not give information about volunteers' gender, and the others that refer to how many subjects are male or female do not distinguish differences

between gender. The work [77] is the only study that relates the data obtained to each gender, in which women present a greater number of cycles per minute, both at a cardiac and respiratory level. However, this conclusion cannot be considered since only one subject per gender was considered, and with this population size, it is impossible to draw general conclusions. Regarding the analysis of the impact of the volunteers' stature on the acquired signal, none of the studies does it. This information is only provided as an indication and for the readers' knowledge, but there is no relation with the obtained results.

In addition to the theory-practical studies, there are three more datasets in the context of CW radar. The first one acquired a digital stethoscope, an Electrocardiogram (ECG), and a respiration sensor, which measured 11 test subjects [76]. The second, similar to ours, measured the heartbeat and respiration; however, only were measured 30 healthy subjects [75], is a relatively low number, considering what we intend to do in this work. Exist another dataset acquired by the FMCW system, where were realized tests at 50 children. The purpose of this study is to develop a child safety sensor for intelligent vehicles [84].

3

Bio-Radar Implementation and Validation

3.1 Bio-Radar Model

Before starting with the study of physiological and body stature variability in the acquisition of vital signs using the Bio-Radar, we must familiarize ourselves with the overall system and the signals that it acquires.

The system under study, as mentioned in Chapter 2, consists of two antennas, Tx and Rx, where the Tx antenna transmits a signal focused on chest wall of the subject. This signal undergoes a phase modulation due to the chest movement, reflecting the wave to the Rx antenna. After receiving the signal, it is down-converted to the baseband. The baseband signal can be expressed by Equation (3.1):

$$g(n) = Ae^{j\varphi(n)} \tag{3.1}$$

Where $\varphi(n)$ is the phase modulation function, and considering an ideal scenario, it is given by Equation (3.2).

$$\varphi(n) = \left(\frac{4\pi d_0}{\lambda} + \Theta\right) + \frac{4\pi R(n)}{\lambda} \tag{3.2}$$

The phase modulation function is composed of two parts: the first one depends on

the d_0 component (distance between Rx antenna and the subject), and the second and last depends on the $R(n)$ component (chest wall movement). Both parts depend on wavelength (λ). The Θ angle is constant and it is related to the distance between the mixer and the antenna and the phase shift in rib cage surface [85].

The relation between wavelength (λ) and operation frequency (f_c) is represented by Equation (3.3). It is considered that the propagation speed (c) is equal to the light speed in a vacuum.

$$\lambda = \frac{c}{f_c} \tag{3.3}$$

Finally, $R(n)$, in Equation (3.2), describes the subject chest movement, and this function is represented by the Equation (3.4).

$$R(n) = \Delta_r \cos(2\pi f_1 n) \tag{3.4}$$

Where Δ_r is the amplitude of respiratory movement of the chest wall and f_1 represents the RR.

3.2 Simulation

A script was developed using MATLAB to simulate the impact of some parameters of the Bio-Radar model previously presented. Thus it is possible to analyse the influence of their variations on the received signal. Initially, it started by fixing values, as seen in Table 3.1.

Table 3.1: Simulation initial values.

A	d_0 [m]	λ [m]	Θ [rad]	Δ_r [m]	$f1$ [Hz]
0.1	1	0.0517	$\pi/8$	0.00325	0.3

Considering that the respiratory signal will be analyzed in the complex domain, the received signal is represented in an arc form.

In Figure 3.1 it is represented the $g(n)$ function on a polar graph. The distance between the complex origin and the arc, is represented by A (Equation (3.1)). The arc length is described by $\varphi(n)$ (Equation (3.2)), which represents the chest wall movement. This motion depends on respiratory signals amplitude (Δ_r). Ideally, the arc is centered with

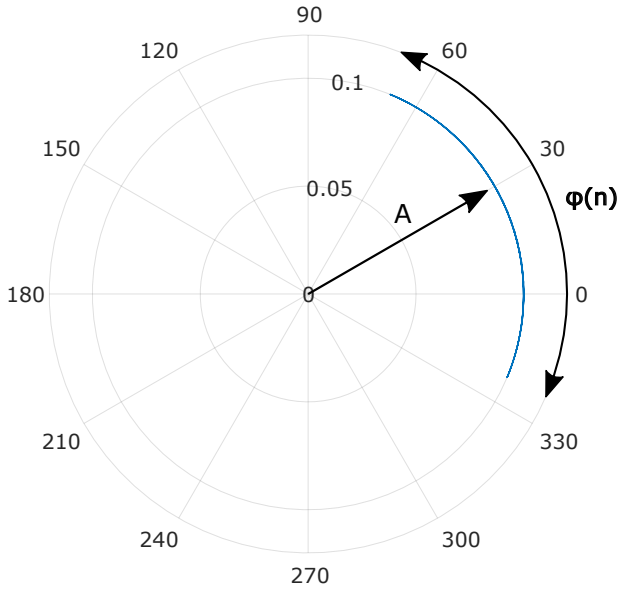
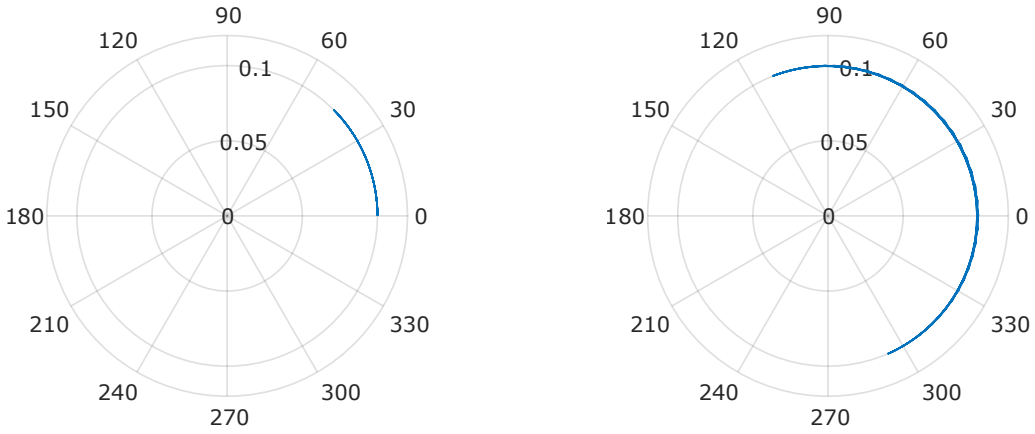


Figure 3.1: Example of a received baseband signal.

the graph origin, and the radius is A. In order to simplify system analysis, it can be considered that the A amplitude is constant, because $\Delta_r << d_0$.

The increase of Δ_r value is directly proportional to the increase of the breathing intensity, and consequently, the arc length increases too. So, as it is possible to verify in Figure 3.2, if it is less intense, the arc length decreases (Figure 3.2a), otherwise increases (Figure 3.2b).



(a) Δ_r = 1.6 mm (b) Δ_r = 6.4 mm

Figure 3.2: Influence of respiratory intensity.

As mentioned previously, A represents the distance between the graph origin and the arc, or in other words, its radius. The increasing of the signal power, A, implies a

deviation from the origin of the graph, as it is possible to verify in Figure 3.3.

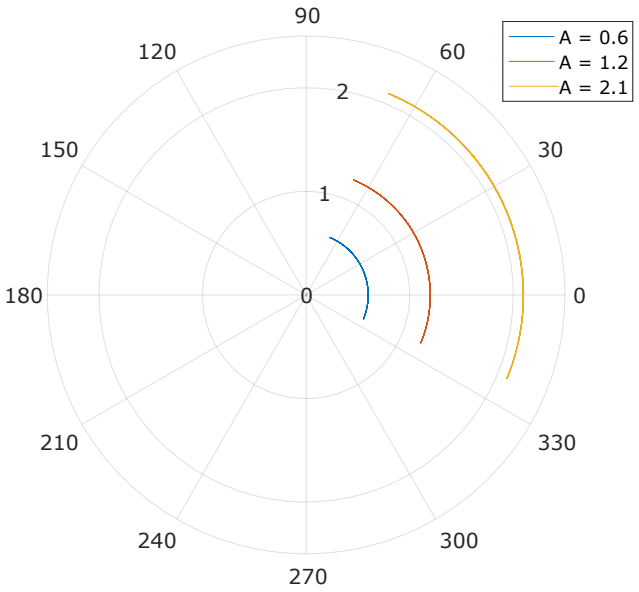
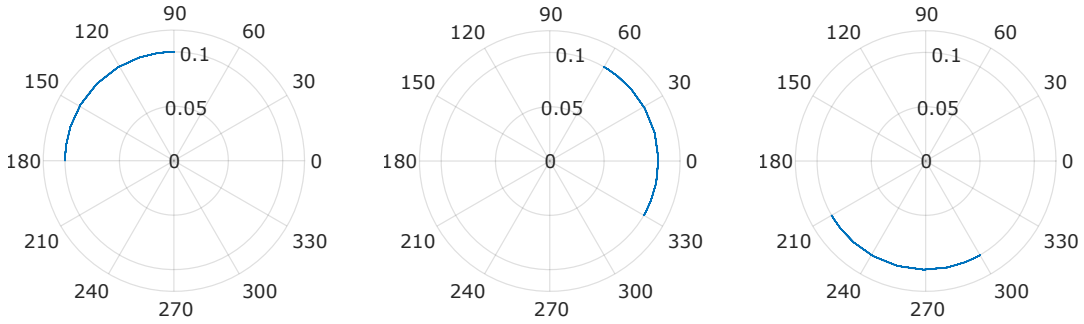


Figure 3.3: Representation of the received signal in function of the power A.

Changing the distance between the subject and the antennas, d_0, affects only the arc position in the polar plane. The distance to the graph origin does not change. As shown in Figure 3.4, a clockwise movement of the arc is observed when d_0 increases. Thus, it is possible to state that there is an arc delay due to the signal's increased traveled distance.

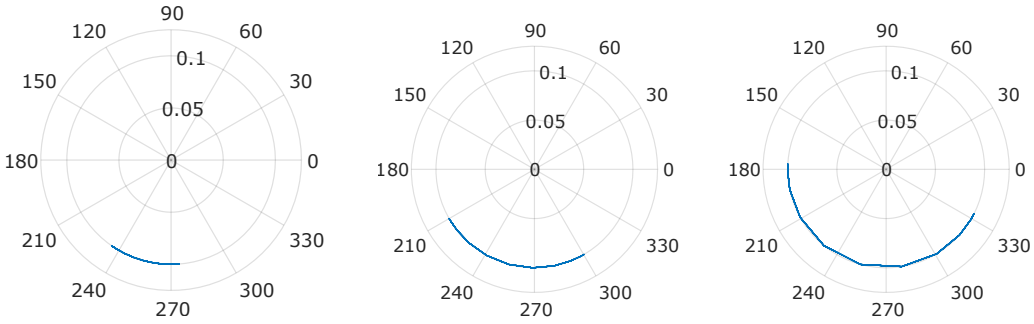


(a) d_0 = 0.5 m. (b) d_0 = 1.5 m. (c) d_0 = 2.5 m.

Figure 3.4: Representation of the received signal in function of the distance d_0.

According to the Equation (3.3), the carrier frequency is inversely proportional to the wavelength. If the wavelength decreases, the arc length increases, as shown in Figure 3.5.

As mentioned at the beginning of this section, for all these simulations it was considered an ideal scenario. However, there are always noise, and other reflections sources



(a) f_c = 2.5 GHz. (b) f_c = 5.8 GHz. (c) f_c = 10 GHz.

Figure 3.5: Influence of the carrier frequency in the received signal, when $\Delta r = 0.00325$.

in a real scenario, such as clutter from the scenario reflections.

3.2.1 Impact of the DC Offset

Assuming that all sources of clutter are stationary, these reflections do not vary over time and therefore are perceived as DC offsets. The DC offset is the principal source of the signal disturbance in CW radars. Taking this information into account, the received signal considering the DC offset is given by the Equation (3.5):

$$g(n) = Ae^{j\varphi(n)} + A_{DC}e^{j(\theta_{DC})} \tag{3.5}$$

Where A_{DC} is the amplitude of the DC offset and θ_{DC} is its phase.

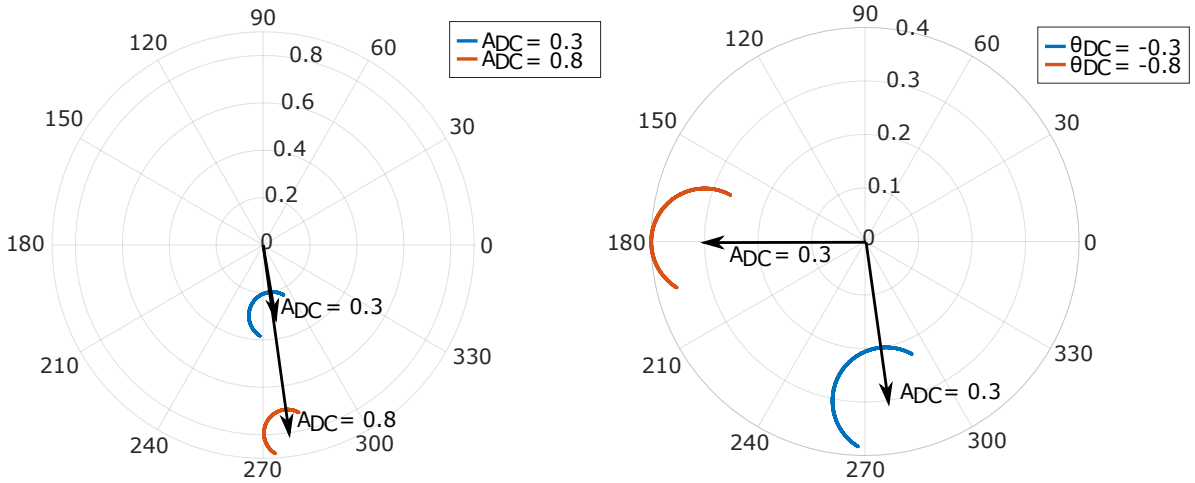
Before going on with the removal of the DC offset, it is necessary to study and understand how it affects the signal. This component influences the arc center and depends on two factors, which are: its amplitude, A_{DC}, and its phase, θ_{DC}.

Focusing first on its amplitude, if it increases, the arc center will be deviating from the graph origin, according to the vector of length A_{DC}, as it is represented in Figure 3.6a.

Relatively to its phase, the arc will move clockwise if it increases, as Figure 3.6b illustrates.

3.3 Signal Processing

Now that the signal behavior is understood, signal processing is carried out, and it is summarized in Figure 3.7. This processing will allow us to obtain the signal that will later be used to calculate the RR, and other important features.



(a) DC offset amplitude A_{DC}. (b) DC offset phase θ_{DC}.

Figure 3.6: Received signal representation with DC offset.

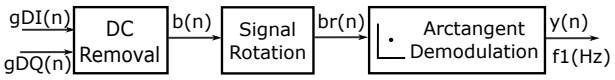


Figure 3.7: Diagram of the signal processing.

Knowing that there are sources of signal disturbance in a real scenario, the signal that arrives at the system will have DC offset. As the first step, it is mandatory to remove the DC offset so that when analyzing the signal, only the signal reflected by the chest wall is analyzed, without the reflections coming from the obstacles in the scenario.

3.3.1 DC Offset Removal

The transmitted signal is reflected not only by the subject's chest wall, but also by other objects located in the vicinity. All these reflections are equally received by the radar front-end. Assuming that these objects are stationary, these reflections are perceived as DC offsets added to the arc center. Figure 3.8 depicts an example of a respiratory signal acquired in a real context scenario.

The method used in this work to remove the DC offset, the Optimized Cost (OC), was developed in [60]. Similarly to other methods presented in literature this method is based on a cost-function minimization. However, this method is robust to weak signals, considering that weak signals are signals that have reduced power and amplitude. Most of the classic methods, based on circle fitting, when the signal is weak, give a solution to the DC offset (center of the arc) inside the radar samples, assuming that the radar samples form a circle and not an arc of a circumference. In contrast, this

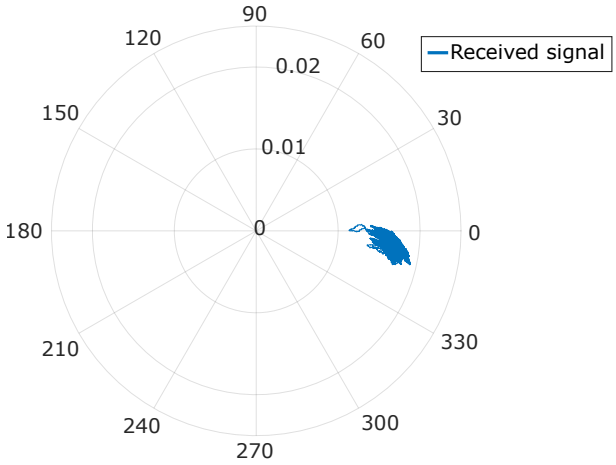
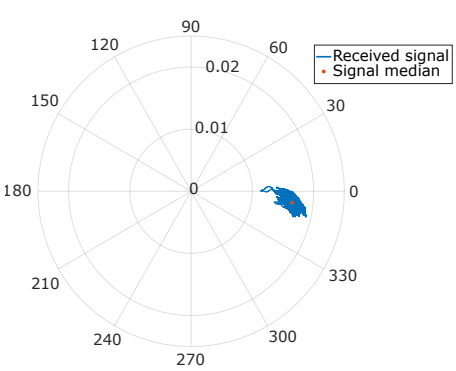
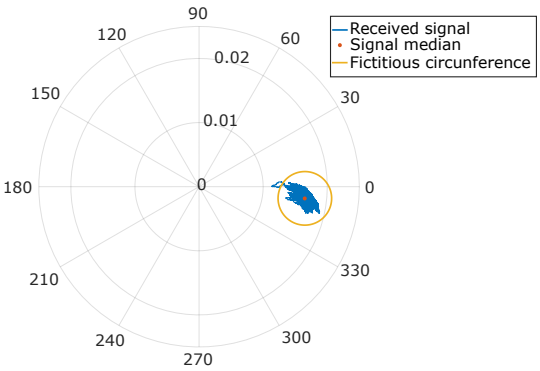


Figure 3.8: Example of a real received respiratory signal.

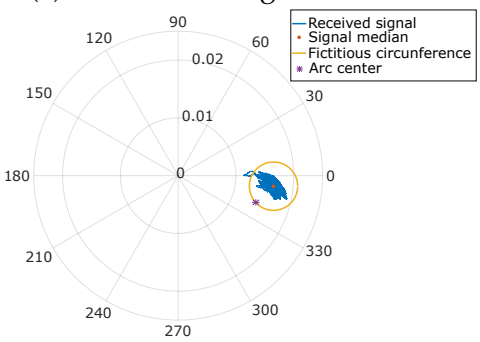
method assumes that the radar samples represent an arc and not a circumference, providing a DC offset closer to the true one. As a disadvantage, the OC method considered that a respiratory signal is always an arc, ignoring the possibility of the respiratory signal being a circle. This study guarantee that the signal is always an arc through the prototype, namely the frequency. The Figure 3.9 shows the process of removing the DC offset from a real signal.



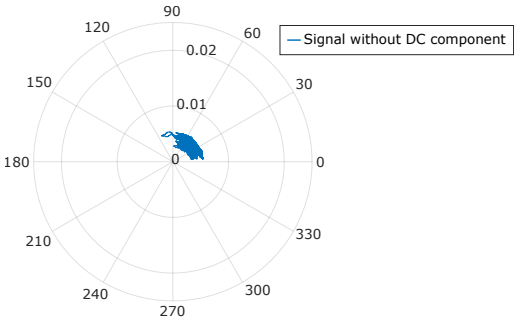
(a) Discover the signal median.



(b) Fictitious circumference.



(c) Lowest cost point.



(d) Signal without DC offset.

Figure 3.9: DC offset removal process.

Additionaly, this method uses a fictitious circumference to limit the searching points of for possible DC offset solutions, and thus the identified solution is guaranteed to be outside the radar samples (see Figure 3.9c). Its center is given by the calculation of the signal median. This technique applies a median to avoid outliers coming from impulsive movements of the subject. Using the cost function presented in [60], it is achieved the point, which presents the lowest cost in relation to the respiratory signal, and it is presented in Figure 3.9c. This point represents the DC offset present in the signal. Removing the point to the radar samples will result in the signal of Figure 3.9d.

3.3.2 Signal Rotation

The process that follows the removal of the DC offset is to rotate the signal in phase while maintaining the distance to the graph's center. This rotation consists of the signal varying around a phase equal to zero by subtracting the signal angle. The reasons why this process is developed may not be obvious. However, hypothesizing that the signal will vary around 180º and knowing that this location corresponds to not only 180º but also -180º, the signal will have wraps. The reason to resort to this method is to avoid these wraps, making the signal under analysis a signal with no average phase component, since it will vary around 0º phase. The representation of the resulting signal is shown in Figure 3.10.

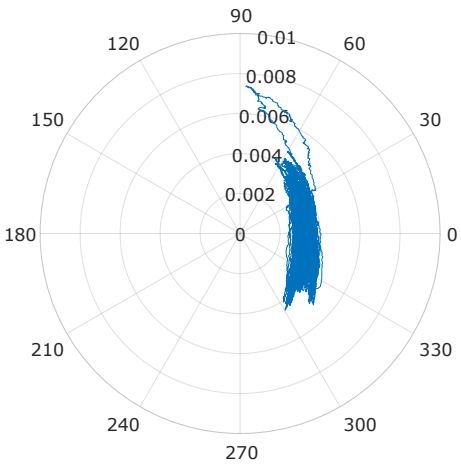


Figure 3.10: Example of a rotated signal.

3.4 Signal Parameters Computation

In Chapter 2, the physiological and body stature variabilities studied are mostly related to respiratory amplitude and RR. Thus, in this section two parameters of the respiratory signal are obtained, its amplitude and its rate.

The first parameter obtained is the RR. This parameter informs how many BrPM are taken. The second, and last, parameter obtained is the respiratory signal amplitude. The amplitude of the arc reveals information regarding the expansion of the subject's rib cage, namely its forward movement.

3.4.1 Respiratory Rate Calculation

One of the respiratory signal characteristics under analysis is the RR as a parameter for further analysis. Before starting with the validation process, it is necessary to briefly explain the three methods used for this validation.

3.4.1.1 Spectral Analysis

For the spectral analysis, the Fast Fourier Transform (FFT) algorithm is used to convert the original signal domain into a frequency domain. Thus, using the FFT, we can retrieve the signal's fundamental frequency. To achieve the signal into a frequency domain, it was used the Welch method, where the inputs are the respiratory signal, at which was removed its average, and the sampling frequency. In the function outcome, there is a peak, where the magnitude is maximum, and the frequency at which this peak occurs is the RR.

In Figure 3.11 represents the FFT result of a respiratory signal as an example. In this case, the RR corresponds to a maximum of the spectrum in the desired bandwidth (0.1 Hz a 0.5 Hz) [86], being 0.32 Hz.

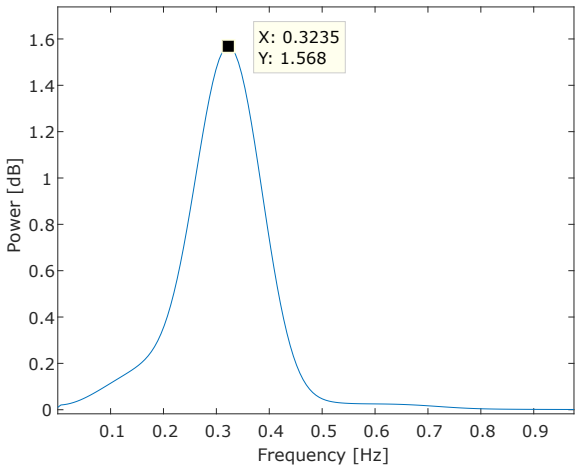


Figure 3.11: Example of the spectral representation.

3.4.1.2 Autocorrelation Analysis

Since autocorrelation is the correlation of a signal with a delayed copy of itself as a delay function, this method will reveal information about the periodicity of the signal.

Figure 3.12 presents an example of the autocorrelation result. Before using the *xcorr* MATLAB function, it is necessary to remove the signal average from it so that it can be an input of the function. The absolute value of the *xcorr* result was applied to get the aspect of the Figure 3.12. The peak situated in the first sample corresponds to a null delay. The second one corresponds to the correlation of the signal with its copy in phase opposition. Third, both signals have the same phase, and it is the peak that gives us information about the signal frequency.

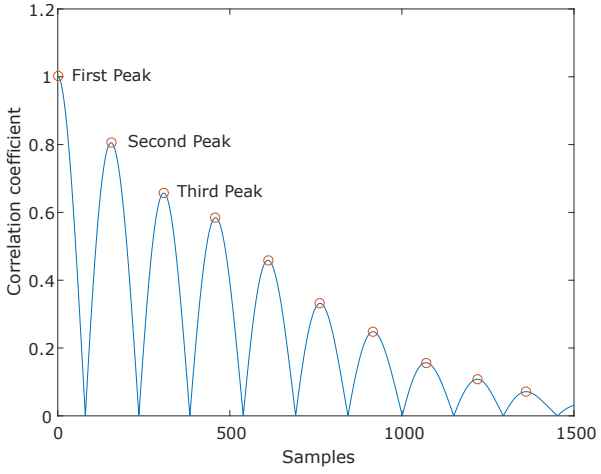


Figure 3.12: Example of auto correlation method.

3.4.1.3 Find Peaks

The *find peaks* method consists of finding the maximums of the respiratory signal, which corresponds to the transition between inspiration and expiration. When two maximums are found, the time difference between them is stored. This process is performed consecutively until the last peak is found. The interval between peaks is referred to as Δ_p. To determine the RR, the average Δ_p of the signal segment is calculated.

Figure 3.13 presents an example of an acquired respiratory signal, with one minute and the respective maximums indicated.

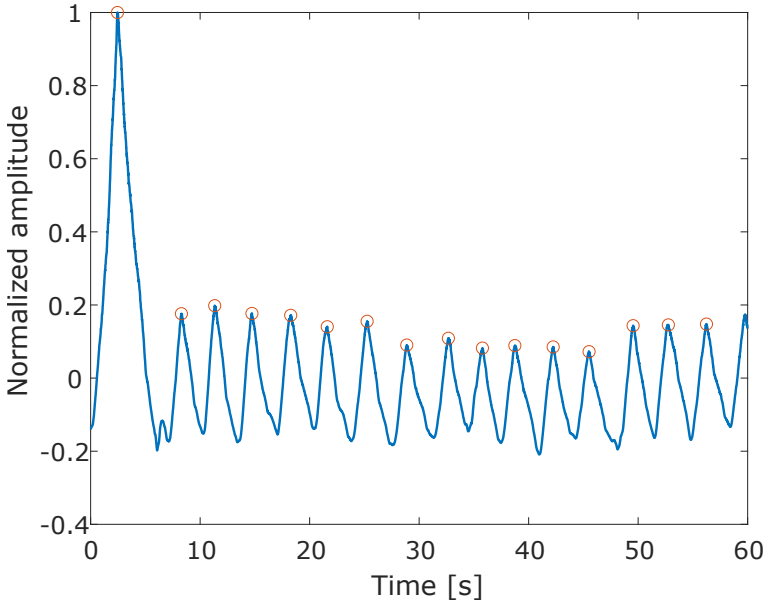


Figure 3.13: Example of *find peaks* method.

3.4.1.4 System Validation

In this way, it is necessary to validate the Bio-Radar system's accuracy in acquiring vital signs, in order to verify if it can be used in this study. For this reason, the validation is separated into two stages, as presented in Figure 3.14. The first involves a short dataset and a certified measuring instrument, the BIOPAC system. In order to overcome the influence of the BIOPAC outliers, this first validation stage was implemented, aiming to select the best method to use as reference. In the second, the BIOPAC system is used as reference and the Bio-Radar system error is calculated in relation to the first system.

Starting with the first validation stage of the bio-radar system, the vital signs of 20 subjects (10 men and 10 women) were acquired, who have the following characteristics,

presented in Table 3.2.

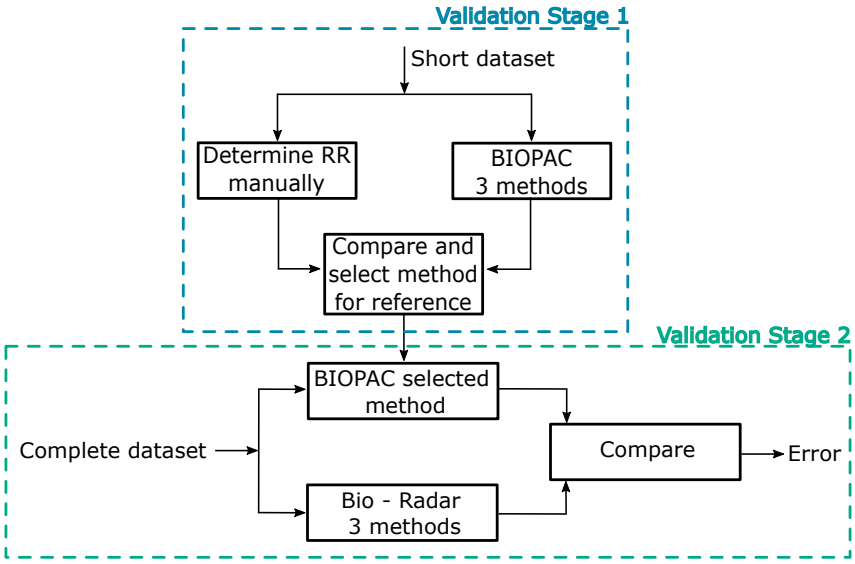


Figure 3.14: RR determination diagram.

Table 3.2: Description of the physical statures of the subjects considered for the bio-radar signal validation.

ID	Gender	Height [m]	CWP [cm]	ID	Gender	Height [m]	CWP [cm]
01	F	1.62	80	11	M	1.86	84
02	M	1.78	80	12	F	1.64	80
03	M	1.88	87	13	F	1.75	80.5
04	F	1.61	73	14	F	1.71	80
05	F	1.68	77	15	M	1.74	81
06	M	1.75	87	16	M	1.62	91
07	M	1.81	98.5	17	M	1.77	94
08	F	1.54	71	18	M	1.78	108
09	F	1.61	72	19	F	1.57	80
10	M	1.88	94	20	F	1.60	73

F - Female, M - Male, CWP - Chest Wall Perimeter

In this short dataset the respiratory signal was measured using two systems simultaneously: the Bio-Radar and the BIOPAC (MP160 Data Acquisition System), which uses the AcqKnowledge 5 software. The respiratory signal is acquired through the RSP100C module with a transducer chest band attached to the subject's chest wall. Thus, a synchronization method was used, which is a respiratory pattern composed of three quick and deep breaths and an apnea period of 10 seconds. After this apnea, the subject exhales, taking out all the air he has in his lungs, as it is represented in Figure 3.15. The trigger was made at this very moment.

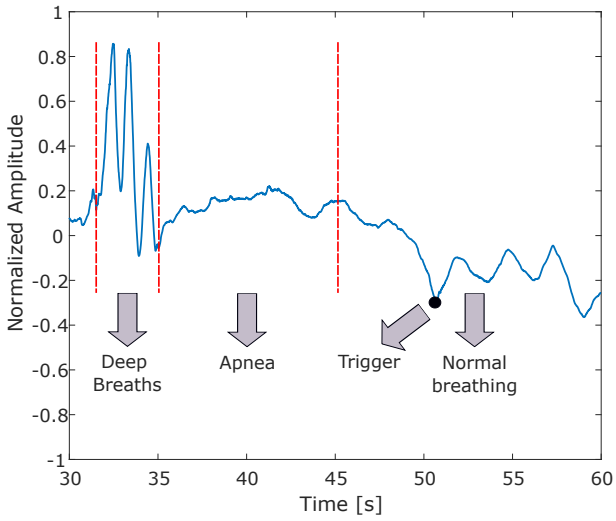
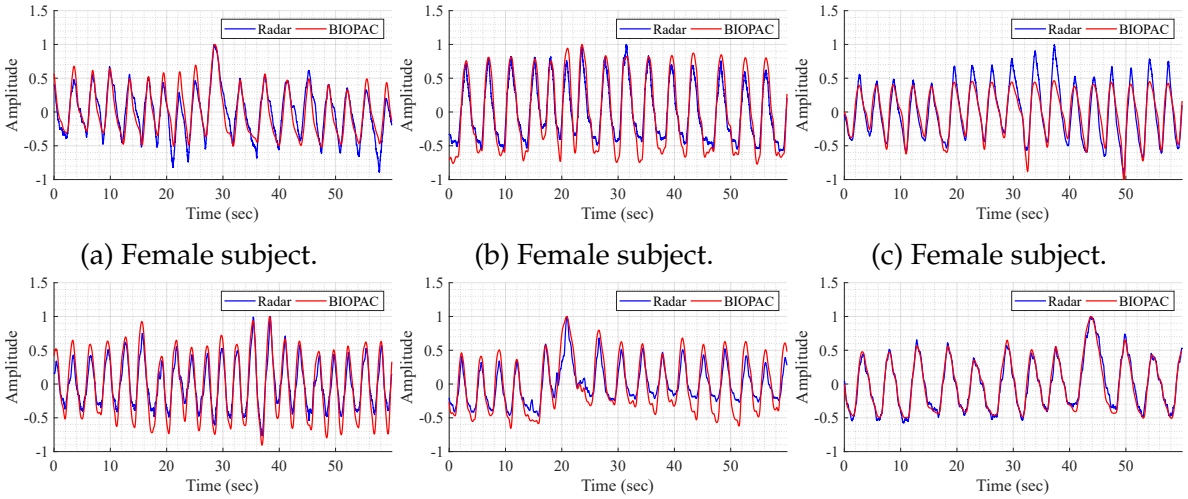


Figure 3.15: Example of a subject respiratory pattern.

In Figure 3.16 it is presented a one-minute window of some individuals' signals acquired by the two systems. By analyzing this figure, it is possible to state that the signals acquired by both systems are synchronized and similar.



(a) Female subject. (b) Female subject. (c) Female subject.

(d) Male subject. (e) Male subject. (f) Male subject.

Figure 3.16: Examples of signals acquired by the Bio-Radar and the BIOPAC simultaneously.

The next step in this process involves the RR determination of the short BIOPAC dataset, which consists of one minute out of a total acquisition of five minutes for each participant. A preliminary test is performed, which is to manually count the number of peaks in the signal from this system. After this, the three methods of RR calculation mentioned before were used in the BIOPAC signals and then compared to the results of the previous step. The method that comes closest to the manual determination will be the method used as a reference for the BIOPAC system.

Since the reference method was already selected, the second stage of validation is reached, considering the complete dataset. The RR of the signals from the BIOPAC system is calculated with the reference method selected in the first validation stage. In order to calculate the measurement error of the Bio-Radar compared to the BIOPAC, the three RR determination methods are used again. The method that will have the best performance, i.e., the lowest error, is the method that will be used later in this work.

Once the process for determining the method for calculating RR has been described, it will proceed to the choice of the method at each stage. To evaluate the performance of each method, three different errors were considered: the Absolute Error (AE), as in Equation (3.6), and then calculate the MAE, as in Equation (3.7) and the Root Mean Squared Error (RMSE), as in Equation (3.8).

$$AE = |RV - OV| \tag{3.6}$$

$$MAE = mean(AE) \tag{3.7}$$

$$RMSE = \sqrt{mean(AE^2)} \tag{3.8}$$

Where:
RV is the true value, which is from BIOPAC;
OV is the observed value.

In the first validation stage, it is selected the method to determine the RR of the BIOPAC signal. In Table 3.3 is presented the MAE of this system relatively to the RR determined manually. Given the values presented, the choice of the reference method is clear, since the method with the smallest error is chosen, i.e. the *find peaks* method.

Table 3.3: Summary of errors for the selection of the reference method.

Method	MAE
Spectral Analysis	0.6679
Autocorrelation Analysis	0.6679
Find peaks	0.2794

In the second validation stage, it is selcted the method which is applied in the Bio-Radar signals. Table 3.4 shows the values of the three methods errors relative to the *find peaks* method used in the signals acquired by the BIOPAC system. By analyzing it,

it is possible to verify that the *find peaks* is the method that presents the lowest MAE. To corroborate this statement, we developed a CDF graph, presented in Figure 3.17. This graph reveals that the *find peaks* method presents the lowest error for a large percentage of the dataset. In other words, for 80% of the dataset, the *find peaks* presents an error of 0.5 BrPM, whereas the autocorrelation presents an error of around 1 BrPM, and the FFT presents an error of upper to 1 BrPM.

Table 3.4: Summary of errors for the selection of the method to apply in Bio-Radar signals.

Method	MAE	RMSE
Spectral Analysis	0.9177	1.4255
Autocorrelation Analysis	0.7570	1.3583
Find peaks	0.6552	1.8322

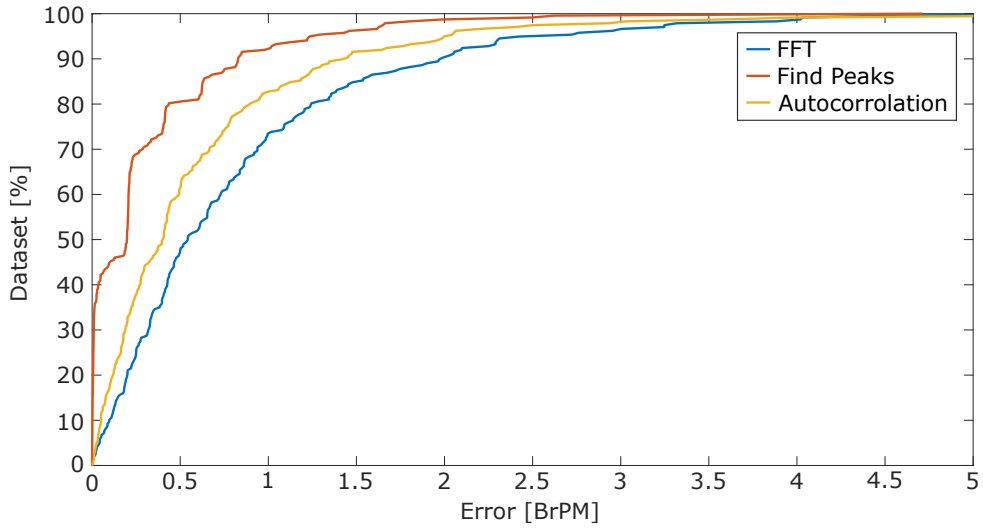


Figure 3.17: CDF graph of the three methods.

Finally, the computation time of each method is determined, resorting to the *tic* and *toc* MATLAB functions, for the 5-minute signal. The results are presented in Table 3.5, and it can be seen that the autocorrelation approach stands out positively, with the shortest time. The second best time is the *find peak* approach, and lastly the FFT.

Table 3.5: Computing time of each method.

Method	Computing Time [s]
FFT	0.0895
Auto correlation	0.0202
Find peaks	0.0491

Based on the data revealed, it can be seen that even though the *find peaks* method does not have the shortest computation time, it is the method that stands out positively concerning MAE and CDF relative to the other methods. Therefore, it will be the method used to calculate the RR for the new dataset.

3.4.2 Arc Length Calculation

In addition to RR, this study will look into respiratory amplitude. The amplitude of the arc reveals information regarding the expansion of the subject's rib cage, namely its forward movement. The respiratory amplitude is given by the arc signal length.

Firstly, to determine the arc length, it is necessary to determine the midpoint of the arc, which could be obtained through the signal mean. However, in this case, it was used the signal median to discard the impulsive or involuntary movements of the subject, as it is visible in Figure 3.18. In this way, the signal midpoint C is given by the Equation (3.9):

$$C = median(P), \tag{3.9}$$

where P are all the points of the arc formed by the respiratory signal. After this, it is calculated the median of the distance between the signal midpoint and all the P points of the signal (Equation (3.10)).

$$d = median(|P - C|) \tag{3.10}$$

This median distance, d, corresponds approximately to the size of a quarter of the signal. Thus the arc length of the respiratory signal will be approximately four times this distance (Equation (3.11)).

$$L = 4 \times d \tag{3.11}$$

To validate the Equation (3.11), is drawn a circumference centered in C and with diameter L. As it is possible to verify in Figure 3.18, the arc length of the respiratory signal corresponds to L.

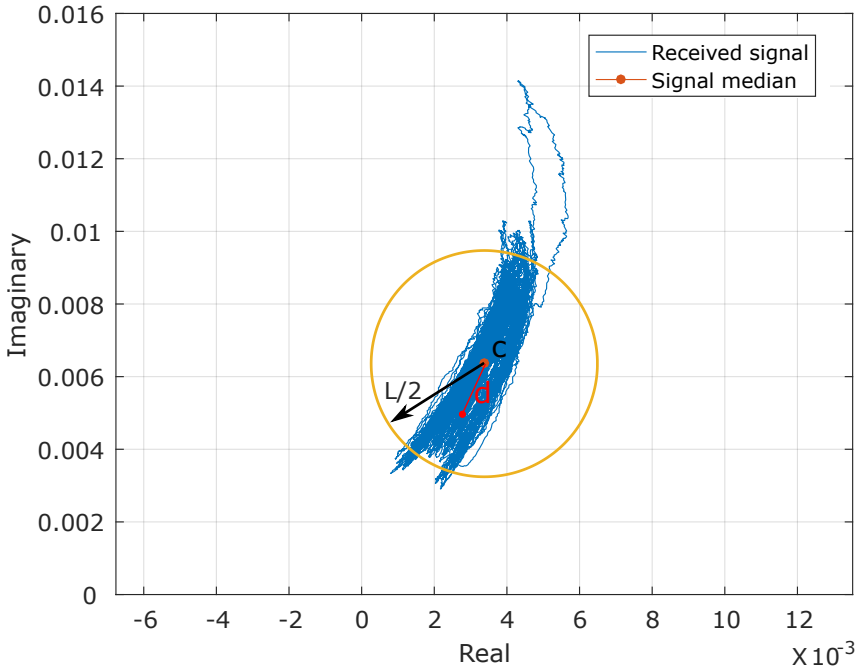


Figure 3.18: Respiratory arc length calculation.

4

Bio-Radar Signals Statistical Analysis

4.1 Vital Signs Acquisition

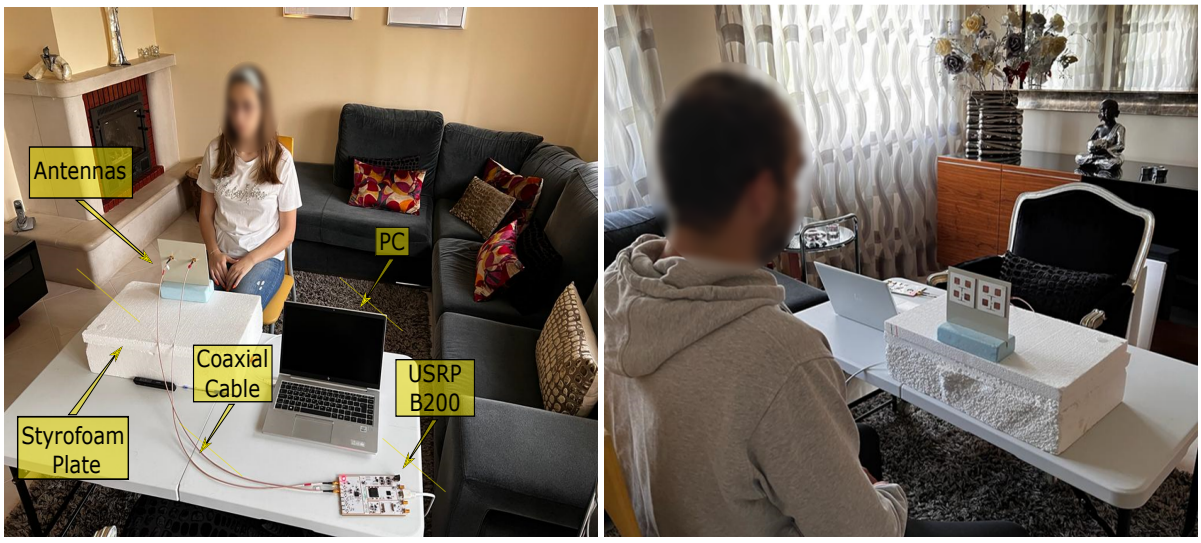
One of the principal and most crucial components of this work is, in fact, the experimental component. This component involves the vital signs non-contact acquisition to build a dataset including 92 individuals and its subsequent analysis. To the best of our knowledge, this dataset is the first one with such a large number of people. This dataset can benefit researchers who do not have access to radar hardware, do not have time to acquire such a large one, or want to increase theirs own dataset.

4.1.1 Setup

First, it is essential to describe the environment in which the experiment took place. The experience occurred in only one physically unchanged environment of 18 m^2. This environment is a typical living room (with a sofa, furniture, and television).

The setup consists of the Bio-Radar antennas as RF front-end. Both are placed, according to Figure 4.1a, placed on a plastic table. Between the antennas and the table, there are always styrofoam plates to adjust the antennas position in relation to the subject's chest wall.

Figure 4.1a and Figure 4.1b show the experimental setup with a woman and a man participant for data collection, including a CW radar, which its prototype, as mentioned before, was developed in [27].



(a) Woman participant. (b) Man participant.

Figure 4.1: Overview of the measurement setup and the system configuration.

4.1.2 Dataset

For data acquisition, 92 subjects (46 males and 46 females) were recruited, aged between 18 and 50. The study was approved by the Ethics and Deontology Committee of the University of Aveiro, Portugal (No.29-CED/2021). The implemented procedure was in line with the Declaration of Helsinki, and informed consent was obtained from all the subjects. Our main objective was to have a balanced dataset in relation to the gender, age, CWP, and BMI distributions.

Additionally to the Bio-Radar signals, some information related to the participants was collected namely age, CWP, height, and weight, as it shown in Table A.1.

The experiment carried out consists of 5-minute of acquisition, where the volunteer breaths normally.

4.2 Physiological and Body Stature Analysis

At the beginning of this document, it was mentioned and discussed the possibility that the physiological and body stature variability may influence the received signals acquired by the Bio-Radar system. Furthermore, the influence might be more noticeable in features related to the breathing amplitude, i.e., to the rib cage expansion of and to the RR of the subject. For this reason, two features of the respiratory signal were extracted, the RR and the respiratory arc length (respiratory signal amplitude) L. Table

4.1 shows that men have, on average, a lower value of RR than women. On the other hand, men have a higher value of respiratory amplitude than women.

Table 4.1: Statistics values of RR and L.

Gender	RR mean $\pm$ STD [BrPM]	RR median [BrPm]	L mean $\pm$ STD	L median
Male	14.6252 $\pm$ 4.0037	14.7602	0.0064 $\pm$ 0.0034	0.0056
Female	16.5741 $\pm$ 3.8281	16.6840	0.0040 $\pm$ 0.0023	0.0034

Besides the gender influence, graphs were drawn to represent the variations of the L and the RR with the BMI and the CWP, and they are shown in Figure 4.2. Individuals who have the highest values of BMI not only have lower values of L, as is shown in Figure 4.2a, but also have the lowest values of RR, as presented in Figure 4.2b. These facts can be justified by the studies mentioned in Section 2.1.3, that state that people who present a higher fat percentage in the diaphragm zone have a lower capacity to move the chest wall [47].

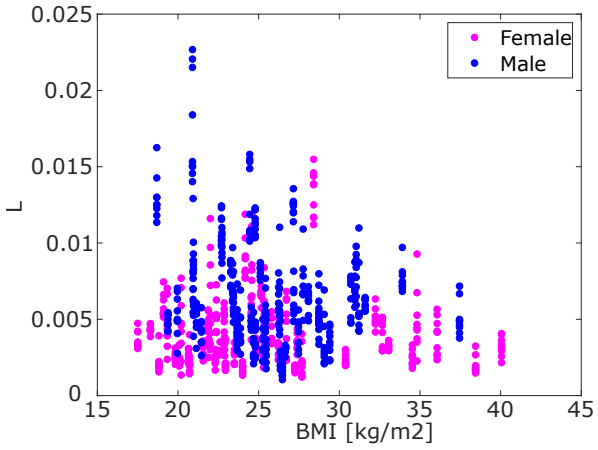
Relatively to the CWP, analysing Figure 4.2c, the highest values of L belong to people with a medium CWP. People who have a lower CWP, have a higher RR, as presented in Figure 4.2d. Taking into account that the people who have a smaller CWP may present a lower lung volume, they may have to breathe faster to obtain the oxygen they need.

Associating the L at RR, it is possible to conclude that higher values of L correspond to lower values of RR, as it is possible to confirm through Figure 4.2e. This analysis corroborates the statements made by looking at Table 4.1, where men have a higher L and a lower RR, and women have the opposite.

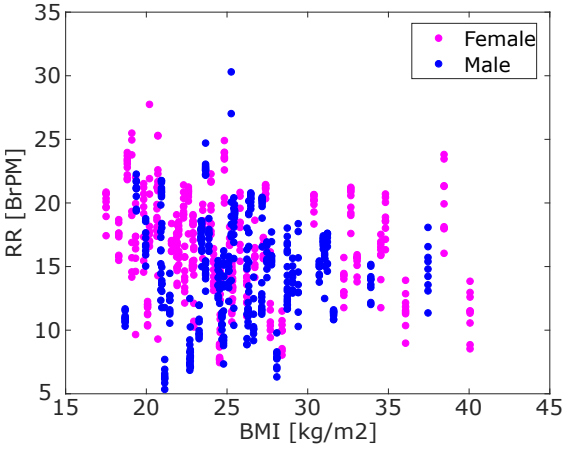
4.3 Selecting the Signal Features

The Bio-Radar signal features were selected taking into account the ones tipically used in the literature [9, 12, 80, 87–90], resulting in the following list:

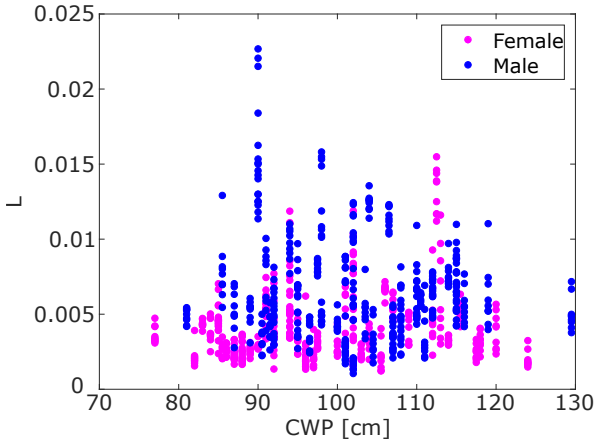
1. **L:** Arc length of the respiratory signal, in other words, amplitude of the respiratory signal;
2. **med ibi:** Interbreath Interval (IBI) median;
3. **mean d1 signal resp:** Mean of the absolute value of the respiratory signal first derivative;



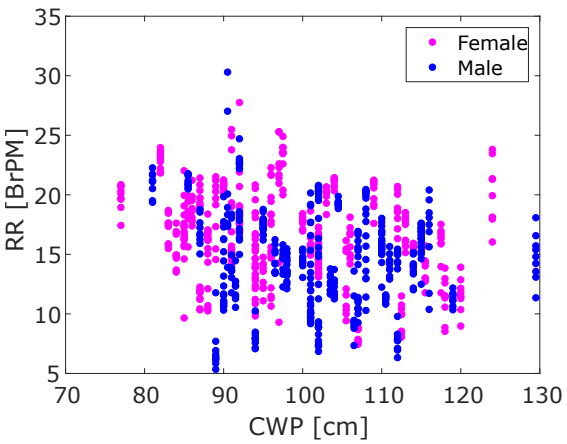
(a) Variation of L with BMI.



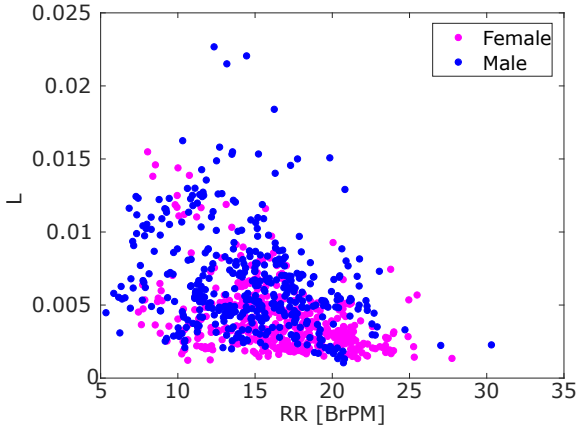
(b) Variation of RR with BMI.



(c) Variation of L with CWP.



(d) Variation of RR with CWP.



(e) Variation of L with RR.

Figure 4.2: Overview of the statistical results.

4. **mean Nd1 signal resp:** Mean of the absolute of the normalized respiratory signal first derivative;

5. **mean d2 signal resp:** Mean of the absolute of the respiratory signal second derivative;

6. **mean Nd2 signal resp:** Mean of the absolute of the normalized respiratory signal second derivative;
7. **skewness T inhale:** Skewness[1] inhalation time;
8. **med T inhale:** Median of inhalation time;
9. **iqr T inhale:** Inter-Quartile Range (IQR) of inhalation time;
10. **mean T inhale:** Mean of inhalation time;
11. **skewness T exhale:** Skewness of exhalation time;
12. **med T exhale:** Median of exhalation time;
13. **iqr T exhale:** IQR of exhalation time;
14. **mean T inhale:** Mean of exhalation time;
15. **RR:** Respiratory Rate;
16. **p01:** Average Power Spectral Density (PSD) energy in the band 0 - 0.1 Hz;
17. **p02:** Average PSD energy in the band 0.1 - 0.2 Hz;
18. **p03:** Average PSD energy in the band 0.2 - 0.3 Hz;
19. **p04:** Average PSD energy in the band 0.3 - 0.4 Hz;
20. **p05:** Average PSD energy in the band 0.4 - 0.9 Hz;
21. **p06:** Average PSD energy in the band 0.9 - 1.5 Hz;
22. **med sign:** Signal median;
23. **mean sign:** Signal mean;
24. **iqr sign:** Signal IQR;
25. **iqr ibi:** IBI IQR;
26. **mean ibi:** IBI mean;
27. **skewness ibi:** Skewness of the IBI;
28. **AppEn max:** Maximums approximate entropy;
29. **kurtosis:** Respiratory signal kurtosis;
30. **SDNN Resp:** Respiratory Standard Deviation of IBI (SDNN);
31. **SAM:** Number of samples above the average of the signal;
32. **RMSSD Resp:** Root Mean Square of Successive Differences (RMSSD) between respirations;

[1]Skewness refers to a distortion or asymmetry that deviates from the symmetrical bell curve, or normal distribution, in a set of data [91]

33. **skewness sign:** Skewness of the signal;
34. **RRV:** Respiratory Rate Variability (RRV);
35. **RMS:** Quadratic mean value of the signal;
36. **freq resp:** Respiratory frequency;
37. **mean peak valley width:** Average of the peaks and valleys width of the respiratory signal in the time domain;
38. **Ratio PSD LF HF:** PSD of the ratio between Low Frequency (LF) (0.1 - 0.4 Hz) and High Frequency (HF) (0.4 - 1.5 Hz);
39. **var sign:** Variance of the respiratory signal;
40. **AppEn resp:** Aproximate entropy of the respiratory signal.

Features 1 and 15 are determined in Section 3.4.2 and 3.4.1, respectively. Comparing feature 15 with feature 36, the only difference is that the first is expressed in BrPM, and the second is expressed in Hz. Relatively to the PSD energy of the signal, represented by the features 16-21, and 38, they are calculated resorting to the *pwelch* MATLAB function. This function is applied to the signal after centering its average at 0.

Features 2, 25-27, 30, and 34 are based on cardiac features, such as Interbeat Interval and Heart Rate Variability. As for the features related to inspiration and expiration, 7-14 are statistical measures extracted from the maximums and minimums of the respiratory signal. Like the last features, features 3-6, 22-24, 28-33, 37, 39, and 40 are also statistical measures concerning the entire respiratory signal.

Feature 34 is determined in [12], using the equation (4.1):

$$RRV = (100 - \frac{H1}{DC})\% \tag{4.1}$$

where $H1$ and DC are the amplitude powers of the first and the zero harmonic peaks, respectively [12].

The feature 35 is given by the equation (4.2):

$$RMS = \sqrt{\sum_{i=0}^{N_s} \frac{x(i)^2}{N_s}} \tag{4.2}$$

where $x(i)$ represents the signal samples, and N_s represents the number of the respiratory signal samples.

4.4 Statistical Analysis

In this section, we will explain all the statistical processes developed in MATLAB. The process that will be described is based on the one conducted in [92], and it is summarized in Figure 4.3.

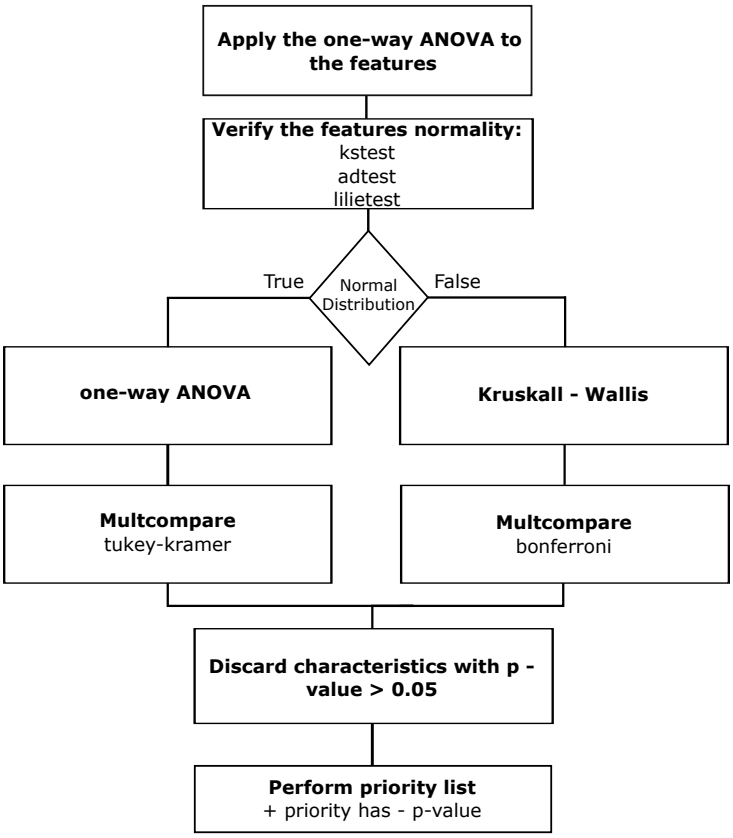


Figure 4.3: Statistical analysis diagram.

This statistical study was conducted for four different tests, where each was dedicated to verify the possibility to distinguish subjects with the specific characteristics. The tests under investigation are gender, age, BMI, and CWP. These tests are defined to obtain a binary classification according to the following:

- **Gender:**
  - Male: 46 people;
  - Female: 46 people.
- **Age:**
  - 18 - 29 years old: 46 people;

  - 30 - 50 years old: 46 people.

- **BMI:**
  - 18 - 25 kg/m^2: 47 people;
  - $\geqslant$ 25 kg/m^2: 45 people.

- **CWP:**
  - $\leqslant$ 100 cm: 46 people;
  - > 100 cm: 46 people.

The first step is to apply the one-way ANOVA test to determine the residuals of this test and check whether their residuals present a normal distribution. The One-sample Kolmogorov-Smirnov, Anderson-Darling, and Lilliefors tests were used for this validation and all have return a decision test for the null hypothesis that residuals are from a normal distribution [93]. In other words, if two or more tests returned the value 0, the null hypothesis is confirmed. Otherwise, it does not come from such a normal distribution. The tests result informed which test can be applied in the next step.

After this, and taking into account the previous step, two tests are applied to obtain the *p*-values [2] of the feature. These tests are the one-way ANOVA, in case the (obtained in the previous step) have a normal distribution, and the Kruskall-Wallis if they have not [95]. After the one-way ANOVA and the Kruskall-Wallis tests, the Multcompare MATLAB function is apllied to verify if no other features can be discarded. The *p*-values are used, in this case, to discard features with values greater than 0.05 and and then build a priority features list. The features that better characterize each test are different between them. In this way the priority list is build each test, sorting the *p*-values in ascending order.

After defining the features priority list for each class, the correlation of the features was analyzed. For this purpose, a correlation matrix was drawn, which again will further restrict the features [96]; this means that there may be features that reveal the same signal information, i.e., are highly correlated. First, a feature was considered to be highly correlated with other when the absolute value of the correlation was superior to 0.5, and then superior to 0.7. When the matrix shows a value equal to or greater than the established value, the feature that presents a higher position in the priority list is the one considered.

[2] *p*-value is the probability of obtaining results at least as extreme as the observed results of a statistical hypothesis test [94].

4.4.1 Discussion of Features Selection

As expected, the priority lists are not equal between the tests since, for each test, some features categorize better in their classes. The Table 4.2, B.1, B.2, and B.3 present the priority lists defined for each test, namely for the gender, age, BMI, and CWP, respectively. The first column of each table shows the initial list of features, the second column shows the ones resulted from the analysis of the correlation matrix considering a correlation threshold of 0.7 and the last column shows the selected features for 0.5 correlation. In addition to the priority lists and according with the statistical process being conducted, the correlation matrices of the tests are also presented in Figure 4.4, B.1, B.2, and B.3.

The features resulting from this process are the ones that will be used for machine learning to classify in the defined tests. The resulting features are different for each test and they are now being discussed.

4.4.1.1 Gender Test

Table 4.2 and Figure 4.4 represent the priority lists and correlation matrices, respectively, of the gender test.

Relatively to the features that better categorize the gender test, it is possible to verify that the RR and the arc of the respiratory amplitude (L) are two of these features. Since women have a higher average RR, while men have a higher average respiratory amplitude, as already mentioned in Section 4.2. Knowing that the features presented in the first priority list can reveal the same signal information, a correlation of 0.7 is applied, which, compared to 0.5, is less restrictive. In addition to the features already mentioned above, RR and L, the signal PSD in different frequency bands were the features that show relevance. Using the 0.5 value is reduced in eight features, compared to 0.7. The features remaining are: L, RR, med T inhale, med sign, iqr sign, AppEn max, and SDNN Resp.

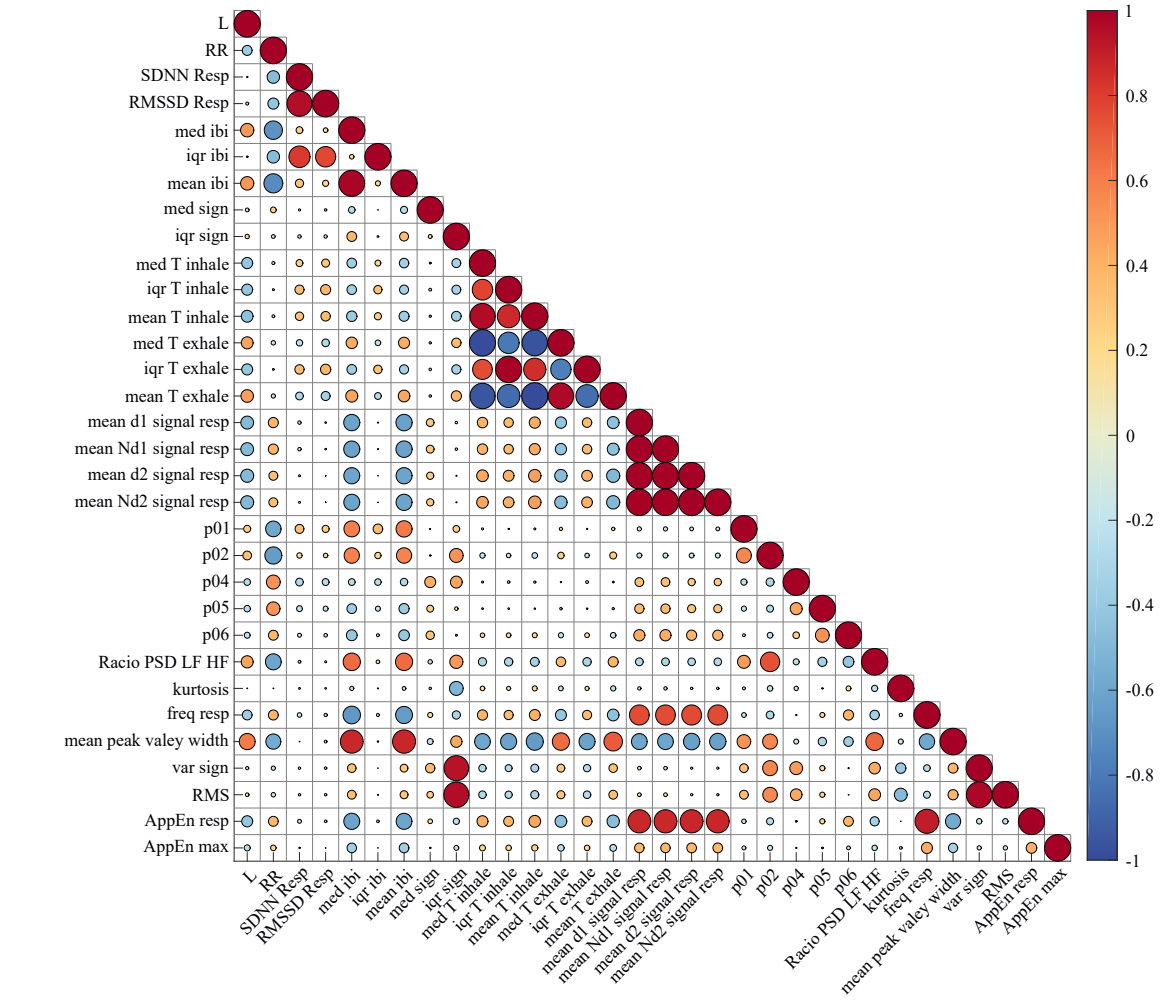
Table 4.2: Priority lists of gender test.

Initial		With 0.7 of correlation		With 0.5 of correlation	
Features	p-value	Features	p-value	Features	p-value
L	1.09×10^{-36}	L	1.09×10^{-36}	L	1.09×10^{-36}
mean peak valley width	5.29×10^{-24}	med ibi	2.66×10^{-20}	RR	1.80×10^{-12}
AppEn resp	2.48×10^{-23}	mean d1 signal resp	3.78×10^{-13}	med T inhale	4.59×10^{-12}
mean ibi	9.06×10^{-21}	RR	1.80×10^{-12}	med sign	7.65×10^{-04}
freq resp	2.66×10^{-20}	med T inhale	4.59×10^{-12}	iqr sign	1.94×10^{-03}
med ibi	2.66×10^{-20}	p02	7.01×10^{-11}	AppEn max	7.93×10^{-03}
mean d2 signal resp	3.00×10^{-16}	p05	7.58×10^{-08}	SDNN Resp	1.57×10^{-02}
mean Nd2 signal resp	3.00×10^{-16}	p04	2.27×10^{-05}		
iqr T inhale	8.63×10^{-15}	p06	2.59×10^{-04}		
mean T exhale	9.00×10^{-15}	p01	2.72×10^{-04}		
med T exhale	1.40×10^{-14}	med sign	7.65×10^{-04}		
iqr T exhale	6.82×10^{-14}	iqr sign	1.94×10^{-03}		
Racio PSD LF/HF	9.39×10^{-14}	AppEn max	7.93×10^{-03}		
mean d1 signal resp	3.78×10^{-13}	kurtosis	1.54×10^{-02}		
mean Nd1 signal resp	3.78×10^{-13}	SDNN Resp	1.57×10^{-02}		
mean T inhale	5.38×10^{-13}				
RR	1.80×10^{-12}				
med T inhale	4.59×10^{-12}				
p02	7.01×10^{-11}				
p05	7.58×10^{-08}				
p04	2.27×10^{-05}				
p06	2.59×10^{-04}				
p01	2.72×10^{-04}				
RMS	4.93×10^{-04}				
var sign	4.95×10^{-04}				
med sign	7.65×10^{-04}				
RMSSD Resp	8.72×10^{-04}				
iqr sign	1.94×10^{-03}				
iqr ibi	4.71×10^{-03}				
AppEn max	7.93×10^{-03}				
kurtosis	1.54×10^{-02}				
SDNN Resp	1.57×10^{-02}				

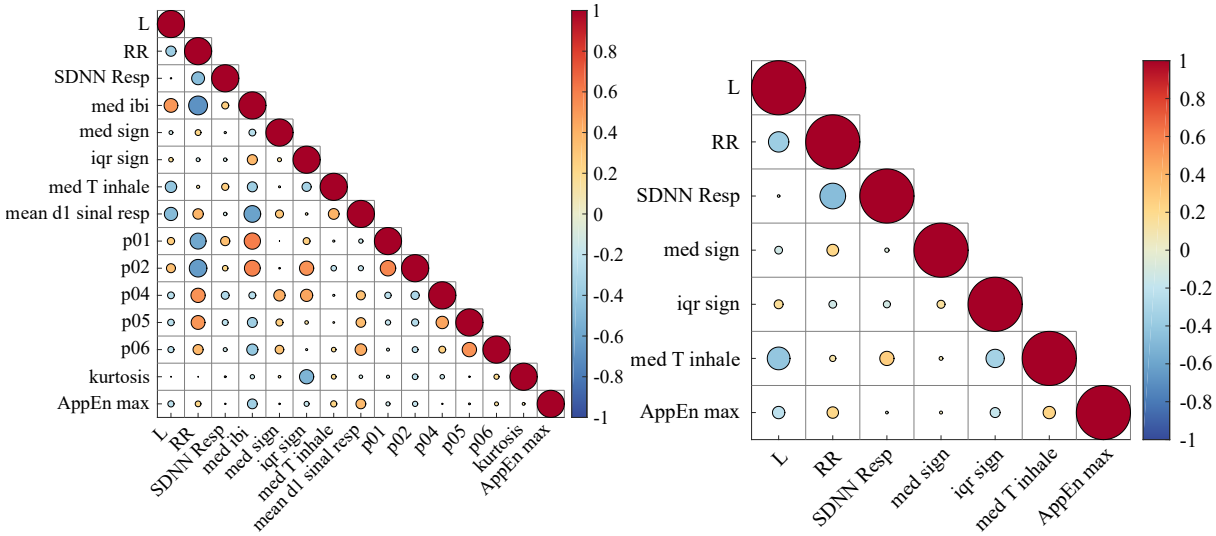
4.4.1.2 Age Test

The representation of the priority lists, and the correlation matrices, of the age test, are in Table B.1 and Figure B.1, respectively.

In this case, the number of features decreased regarding the gender test; this may be because the features used in this study reveal more information about the gender test. Employing a correlation value of 0.7, there is a reduction of 50% in the features. Using a correlation value of 0.5, eight features remain under study. The age test is not categorized neither by the RR value or by the periodicity measures. Aging causes a decrease in the capacity of the thorax to expand [49].



(a) Initial correlation matrix.



(b) Correlation matrix using 0.7 of correlation. (c) Correlation matrix using 0.5 of correlation.

Figure 4.4: Correlation matrices associated at gender.

4.4.1.3 Body Mass Index Test

In Table B.2 is shown the priority lists of the BMI test and in Figure B.2 is shown the correlation matirces also of the BMI test.

The categorization of the BMI is made through fewer features than in the previous tests. In [26, 48], was reported that a higher percentage of abdominal fat is associated with a lower lung volume leading to a higher RR value, which was indeed selected as one of the most important features for this study. The RMSSD Resp may be an indicator of these facts too. Another consequence of fat present in a diaphragm region is that it can hinder the expansion of the chest wall, leading also to a decrease in the respiratory signal amplitude. However, in contrast to what was expected, this feature did not revealed being representative, since it was not selected for this case. As a result of hindering chest wall expansion, the inhalation time can increase, as indicated by the selection of the feature mean T inhale.

4.4.1.4 Chest Wall Perimeter Test

Table B.3 represents the priority lists of the CWP test and Figure B.3 represents the correlation matirces also of the CWP test.

Looking at the features, the CWP test is categorized by the respiratory signal irregularity, RRV, as well as by the RR, and by the signal arc length. These results may be correlated with the people who have a higher CWP to be capable of accommodating a higher amount of air on inspiration and, thus, not having to perform so many breaths. When a lower correlation value of 0.5 is used, the number of resulting features is the most reduced comparing with the remaining tests.

5

Bio-Radar Signals for Subject Classification

5.1 Machine Learning

After obtaining the most relevant features for each test, it is possible to train the ML models. This training aims to verify if it is possible to distinguish the classes inside each test. In this study is denominated test at gender, age, BMI, and CWP, and classes were created for each test in order to define a binary problem. The classes were defined, as defined in Section 4.4, taking into account the number of people inside each one to obtain a quantitative balance between them.

In consideration of this information, it was decided to apply three ML classifiers: SVM, KNN, and Random Forest, which will be explained in the Sections 5.1.1, 5.1.2, and 5.1.3, respectively.

5.1.1 Support Vector Machine

The SVM is a non-parametric classification technique based on statistical learning theory [97]. The main goal of SVM is to find the hyperplane separating the data so that the distance from this hyperplane to each data class is as large as possible. In the example of Figure 5.1, there are two classes, green and blue. The hyperplane is determined using the "support vectors", as it is possible to verify in the example, each "bounding

hyperplane" is composed of two samples of each class. These vectors are the closest support points to the hyperplane in the data sample used for its construction [98].

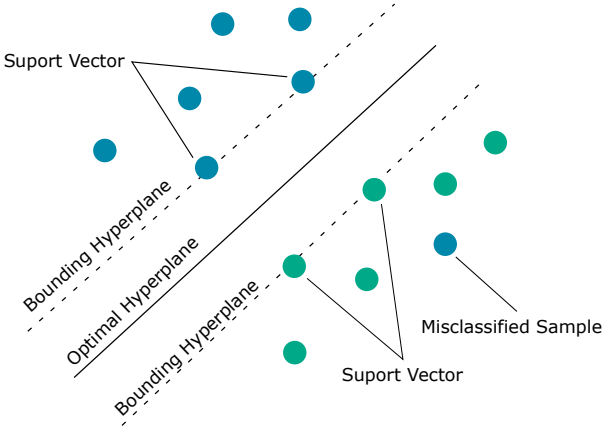


Figure 5.1: The SVM classifier, modified from [99].

5.1.2 K-Nearest Neighbor

The KNN classifies a new observation by measuring the distance of this observation from its nearest points and then categorizing it in the class to which most of the K points belong. A higher number of the K neighbors will imply a smoother classification boundary. On the other hand, a small number of K produces a more flexible boundary with a high variance but low bias. In Figure 5.2 it is presented an example of a KNN classification [100], in which there are two classes, blue and green. When inserting a new sample and considering K=3, it will be classified as green. If it is considering K=7, the new sample will be classified as blue.

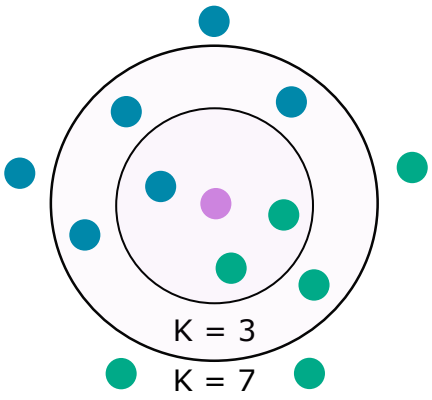


Figure 5.2: The KNN classifier, modified from [101].

5.1.3 Random Forest

This classifier is used in regression and classification problems. The Random Forest consists of a combination of tree predictors. Each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The generalization error of a forest of tree classifiers depends on the strength of the individual trees in the forest and the correlation between them [102]. The final result of this model is the result returned from the trees with the highest vote, or the average of these. The Figure 5.3 represents graphically what was explained above.

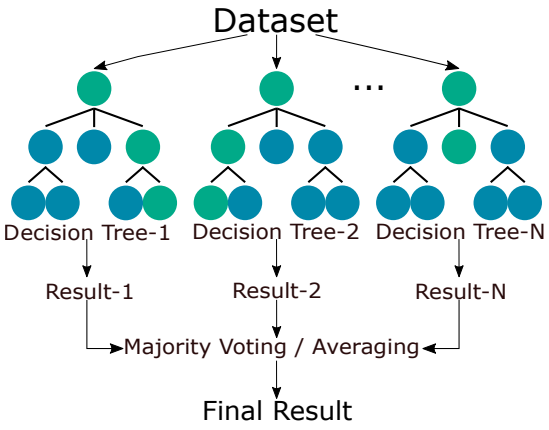


Figure 5.3: Random forest classifier, modified from [103].

5.2 Classifiers Evaluation

The data resulting from Section 4.4 are split into training, test, and new subjects. This dataset is composed by 92 individuals, that have nine 1-minute signal windows. With these 92 individuals, two strategies were conducted for the data splitting for training and testing: the first test uses two subjects as new subjects, and the remaining are divided into 70% for training and 30% for testing; the second test uses five subjects as new subjects, and the remaining are divided into 70% for training and 30% for testing. The Figure 5.4 presents the dataset division tests. Twenty iterations were performed with the same follow-up. Figure 5.5 represents the implemented process. In this study, four tests were carried out to find the condition that would lead to the best results:

- a) Correlation value of 0.7 and 2 people as new subjects;

- b) Correlation value of 0.7 and 5 people as new subjects;
- c) Correlation value of 0.5 and 2 people as new subjects;
- d) Correlation value of 0.5 and 5 people as new subjects.

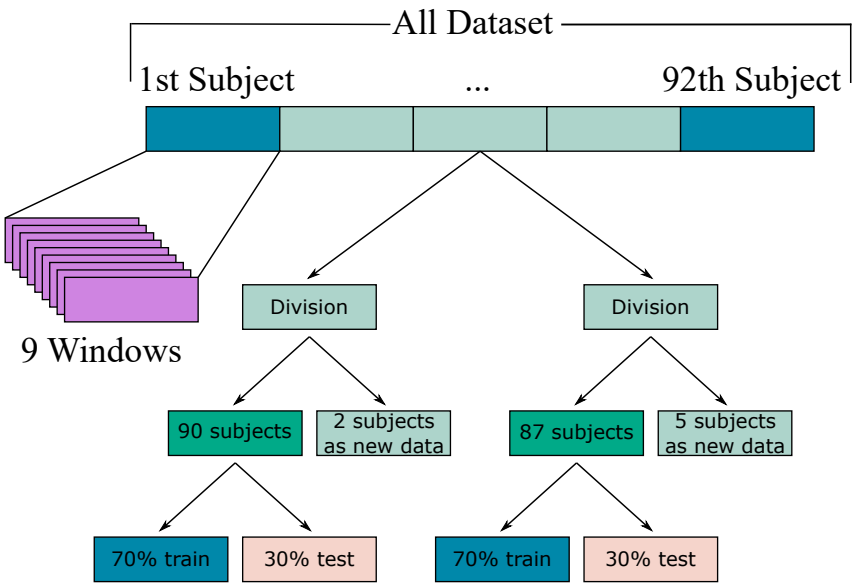


Figure 5.4: Data splitting.

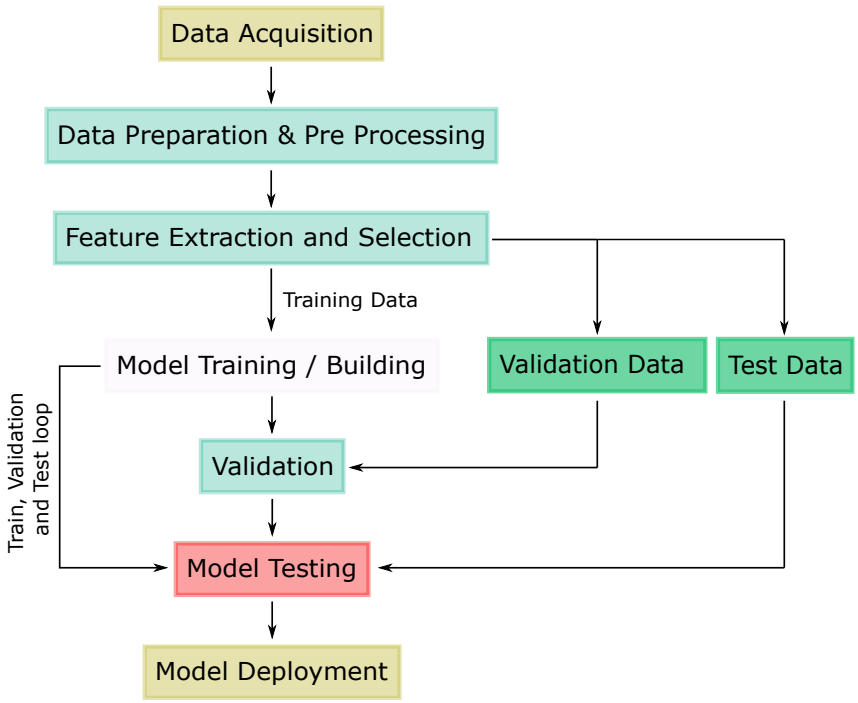


Figure 5.5: Classification workflow, modified from [104].

The hyperparameters selection was made for each ML model in different phases. Firstly, the hyperparameters were obtained with a default range of the MATLAB tool, resorting to a loop with twenty iterations. The results of the first step were annotated, and we decreased the range. This process was repeated until the results were unanimous, and it was defined the models' hyperparameters with the values that had the highest that were more often selected considering all iterarions. Regarding to these defined hyperparameters for the three ML models, they are presented in Table 5.1.

Table 5.1: Selected hyperparameters to the performed tests.

ML Algorithm	Hyperparameters	Values			
		Gender	Age	BMI	CWP
SVM	KernelFunction	gaussian	gaussian	gaussian	gaussian
	Standardize	1	1	1	1
	BoxConstraint	12	7	5	19
	KernelScale	1.7	1.3	2.7	1.3
KNN	NumNeighbors	5	4	10	4
	Standardize	1	1	1	1
	Distance	minkowski	minkowski	minkowski	minkowski
	Exponent	-	0.5	0.6165	-
	DistanceWeight	squaredinverse	squaredinverse	squaredinverse	squaredinverse
Random Forest	Method	regression	regression	regression	regression
	OOBPredictorImportance	On	On	On	On
	MinLeafSize	5	5	5	5

Relatively to the results of the trained models, these are presented by test, for the gender in Section 5.3.1, for the age in Section 5.3.2, for the BMI in Section 5.3.3, and for the CWP in Section 5.3.4. In the subsequent subsections are presented one common table, which reveals the mean accuracy with the corresponding Standard Deviation (STD). Besides these values, obtained through *predict* MATLAB function, it is also presented the accuracy values for the Cross-Validation (CV). In the CV was used the *KFold*, with K assuming a value of 5, which consists of randomly dividing the observations into subsamples, each of which has approximately the same number of observations, as it is shown in Figure 5.6.

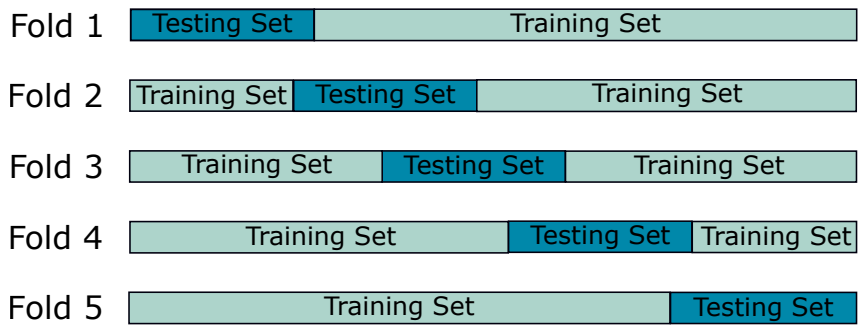


Figure 5.6: Example of *KFold* validation, with $K = 5$, modified from [105].

5.3 Results Discussion

This section is focused on the presentation of the classifiers implementation results and their discussion. Results are presented according to the following order: first the gender accuracies obtained, followed by the age, BMI, and CWP.

5.3.1 Gender

Using Table 5.2 values as reference, it can be seen that the results obtained in the Random Forest algorithm are considerably higher relative to SVM and KNN. Therefore, in the gender classification, an accuracy of 82.96% is obtained in the CV method. In test data, 76.66% is obtained, and an accuracy of 63.33% is achieved in new subjects.

Table 5.2: ML results for the gender test.

Algorithm	Parameters	a)	b)	c)	d)
		Acc $\pm$ STD [%]	Acc $\pm$ STD [%]	Acc $\pm$ STD [%]	Acc $\pm$ STD [%]
SVM	CV	76.02 $\pm$ 1.75	76.00 $\pm$ 1.66	68.73 $\pm$ 2.05	69.15 $\pm$ 2.16
	Test	77.90 $\pm$ 2.47	77.91 $\pm$ 3.29	68.97 $\pm$ 2.91	69.62 $\pm$ 3.79
	New subjects	62.78 $\pm$ 20.16	63.11 $\pm$ 12.48	55.56 $\pm$ 21.78	54.89 $\pm$ 11.56
KNN	CV	74.22 $\pm$ 1.29	74.91 $\pm$ 2.11	68.95 $\pm$ 1.90	68.18 $\pm$ 2.09
	Test	75.95 $\pm$ 2.87	74.85 $\pm$ 2.62	68.87 $\pm$ 2.15	68.95 $\pm$ 2.39
	New subjects	64.17 $\pm$ 26.21	60.78 $\pm$ 15.42	61.67 $\pm$ 16.41	59.67 $\pm$ 14.06
Random Forest	CV	82.96 $\pm$ 0.39	83.22 $\pm$ 0.68	81.49 $\pm$ 0.52	81.60 $\pm$ 0.74
	Test	76.66 $\pm$ 3.02	76.43 $\pm$ 2.65	73.64 $\pm$ 2.74	72.76 $\pm$ 2.44
	New subjects	63.33 $\pm$ 30.77	61.22 $\pm$ 18.43	62.50 $\pm$ 16.99	66.44 $\pm$ 14.98

Acc - Accuracy; CV - Cross Validation; STD - Standard Deviation;
a) - Correlation = 0.7 and New subjects = 2; b) - Correlation = 0.7, New subjects = 5;
c) - Correlation = 0.5, New subjects = 2; d) - Correlation = 0.5, New subjects = 5.

5.3.2 Age

In the age test, the results obtained with the Random Forest algorithm were also better as it is shown in Table 5.3. In this case, the test that presents better accuracy results is the b). In the b) test is obtained 80.19% of accuracy in the CV, 71.13% of accuracy in the test data, and 55.78% of accuracy in the new subjects.

5.3.3 Body Mass Index

Random Forest is the one that, once again, shows better results in this test. In this algorithm, one test stands out from the others, namely the b), since it has the best

Table 5.3: ML results for the age test.

Algorithm	Parameters	a)	b)	c)	d)
		Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]
SVM	CV	67.25 ± 1.55	67.70 ± 1.74	64.17 ± 1.63	64.11 ± 1.88
	Test	68.46 ± 1.95	68.10 ± 2.93	64.92 ± 3.18	64.47 ± 2.60
	New subjects	54.72 ± 19.93	52.11 ± 10.75	56.11 ± 13.11	50.33 ± 11.96
KNN	CV	72.68 ± 2.21	73.78 ± 1.74	66.88 ± 2.51	67.15 ± 2.07
	Test	73.52 ± 2.00	73.44 ± 2.57	67.08 ± 2.80	67.61 ± 2.73
	New subjects	47.22 ± 21.21	48.11 ± 12.33	47.22 ± 17.15	49.00 ± 11.20
Random Forest	CV	80.15 ± 0.50	80.19 ± 0.71	78.96 ± 0.66	79.17 ± 0.58
	Test	70.60 ± 3.15	71.13 ± 3.04	67.41 ± 2.49	68.33 ± 1.99
	New subjects	59.44 ± 23.78	55.78 ± 12.73	65.83 ± 19.01	55.56 ± 13.91

Acc - Accuracy; CV - Cross Validation; STD - Standard Deviation;
a) - Correlation = 0.7 and New subjects = 2; b) - Correlation = 0.7, New subjects = 5;
c) - Correlation = 0.5, New subjects = 2; d) - Correlation = 0.5, New subjects = 5.

result on the test data relative to the other tests. Thus, for the categorization of the BMI test we get accuracy values such as: 80.58% on CV, 72.52% on test data and 56.67% on new subjects.

Table 5.4: ML results for the BMI test.

Algorithm	Parameters	a)	b)	c)	d)
		Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]
SVM	CV	71.31 ± 2.21	72.65 ± 2.24	64.46 ± 2.50	64.17 ± 2.21
	Test	71.42 ± 2.90	70.96 ± 2.44	64.28 ± 3.44	64.62 ± 1.84
	New subjects	56.67 ± 23.68	55.22 ± 14.46	46.39 ± 21.58	53.33 ± 13.31
KNN	CV	71.34 ± 1.48	71.32 ± 1.52	63.75 ± 1.87	64.24 ± 1.84
	Test	72.33 ± 2.84	72.52 ± 2.15	64.07 ± 2.66	63.91 ± 3.20
	New subjects	55.00 ± 21.70	51.67 ± 12.51	51.94 ± 16.74	50.89 ± 7.69
Random Forest	CV	80.76 ± 0.75	80.58 ± 0.49	78.48 ± 0.61	78.53 ± 0.65
	Test	70.60 ± 3.11	72.52 ± 1.39	63.72 ± 2.76	65.88 ± 2.24
	New subjects	59.17 ± 22.96	56.67 ± 11.34	53.89 ± 14.65	51.22 ± 13.44

Acc - Accuracy; CV - Cross Validation; STD - Standard Deviation;
a) - Correlation = 0.7 and New subjects = 2; b) - Correlation = 0.7, New subjects = 5;
c) - Correlation = 0.5, New subjects = 2; d) - Correlation = 0.5, New subjects = 5.

5.3.4 Chest Wall Perimeter

As in all the previous tests, the Random Forest algorithm performs better, as it is possible to confirm in Table 5.4. As in all the previous tests, the Random Forest algorithm performs better. As in the gender test, test a) achieving higher better accuracy values. In this test, the following accuracy values are obtained: 82.08% on the CV, 74.61% on the test data, and 66.11% on the new subjects.

Table 5.5: ML results for the CWP test.

Algorithm	Parameters	a)	b)	c)	d)
		Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]	Acc ± STD [%]
SVM	CV	76.40 ± 2.03	75.28 ± 2.21	70.90 ± 1.77	70.51 ± 2.20
	Test	76.44 ± 2.31	76.82 ± 2.23	70.62 ± 1.77	71.50 ± 2.72
	New subjects	59.17 ± 18.93	57.11 ± 16.32	47.78 ± 19.36	53.89 ± 14.18
KNN	CV	75.29 ± 1.49	75.43 ± 1.76	69.74 ± 1.65	70.95 ± 2.12
	Test	75.78 ± 3.12	75.56 ± 2.82	70.21 ± 2.25	70.43 ± 3.37
	New subjects	51.67 ± 17.85	54.00 ± 11.13	50.28 ± 17.98	47.00 ± 11.88
Random Forest	CV	82.08 ± 0.51	82.22 ± 0.50	80.38 ± 0.54	80.54 ± 0.74
	Test	74.61 ± 2.21	73.29 ± 2.65	70.43 ± 2.37	70.17 ± 2.38
	New subjects	66.11 ± 25.42	61.56 ± 15.83	60.00 ± 15.78	58.56 ± 17.08

Acc - Accuracy; CV - Cross Validation; STD - Standard Deviation;
a) - Correlation = 0.7 and New subjects = 2; b) - Correlation = 0.7, New subjects = 5;
c) - Correlation = 0.5, New subjects = 2; d) - Correlation = 0.5, New subjects = 5.

5.3.5 Discussion

Considering the results presented in the previous subsections, all tests were obtained accuraccies over 50%, and in two cases, gender and CWP, were obtained accuraccies over 60%. The gender test is easiest to classify, since the most accurate classification rate that was achieved with test subjects, 76.66%. Males have in average higher respiratory amplitudes, and lower values of RR. In contrast, females have higher values of RR, and lower values of respiratory amplitudes. For these reason, it is the simplest to classify. Contrarily, the age test is the hardest test to classify, due to the lowest accuracy of classification with new subjects, 71.13%. The characteristics mentioned in the Section 2.1.3.5 may be more meaningful at older ages, and for this reason, the age test has the lowest accuracy. These results achieved in the four tests reveal that the features extracted from the respiratory signal can distinguish test classes with reasonably reliable accuracy.

Analyzing the accuracy results of the CV, and the Test, it is possible to verify that the differences are, in most cases, similar. Revealing the robustness of the model and the inexistence of overfit.

Depending on the correlation values, the results achieved in the tests were significantly different. With a correlation value of 0.7, more features were selected than with a correlation value of 0.5. Analysing the results accuracies, these are higher with a correlation value of 0.7. This means that the features that are included in the correlation value of 0.7 and that are not included in the correlation value of 0.5 are significant for the tests under study. The number of new subjects was also changed, and the highest results correspond to a larger number of people. The reason for this might be the fact that

there were more people for training and testing. However the values achieved with a higher number of people have no significant difference from the values obtained with a smaller number of new subjects, always remaining above 50% in any test. This may mean that the dataset of 90 people is representative of the population.

6

Conclusion and Future Work

6.1 Conclusion

Completed the dissertation, all objectives presented were successfully achieved. With this work, a complete study on the effect of physiological and body stature variability on the respiratory signal acquired by the Bio-Radar was achieved. In addition to this study, ML models were developed to distinguish subjects between genders and age, BMI and CWP.

The first conclusion presented in this work is related to the best approach for determining RR. Section 3.4.1 presents the process to achieve the approach. For this objective, three methods are compared: the autocorrelation analysis, the spectral analysis, and the find peaks. The find peaks method was chosen since it has the lowest MAE, 0.6552 BrPM.

Regarding the influence of gender, BMI and CWP on the respiratory signal. According to the averages of the respiratory signal amplitude and the RR, men have a greater amplitude of the respiratory signal than women. Conversely, women have a higher respiratory rate than men. Regarding the influence of BMI on the respiratory signal, it affects both components. Individuals with a higher BMI have a smaller respiratory amplitude and a higher RR. As far as the influence of CWP is concerned, the people who have a smaller CWP may present a lower lung volume, they may have to breathe faster to obtain the oxygen they need. In the same context, and relating the RR with the

respiratory amplitude, these concepts are inversely correlated, i.e., when the individual RR is higher, the smaller respiratory amplitude is.

In order to achieve good results regarding the main objective, it was essential to have a dataset with several people. Thus, knowing the lack of public datasets, performing vital signs acquisitions through the system under study was necessary. Vital signs were acquired from 92 people (46 males and 46 females) aged 18-50. This dissertation holds the data set with the most significant physiological and body stature variability to our knowledge. Based on the dataset, with high variability presented, it is possible to prove that the relationship between them, stated in studies in other scientific areas, is perceptible with the acquisition system used.

Features were extracted from the respiratory signals, in order to categorize the established tests. Through the statistical process, we achieved the best features for this categorization. Regarding gender, it was possible to conclude that the periodicity signals measures do not help to categorize this test, contrary to RR and L, which are the features that stand out in this test. Regarding age, values such as RR and periodicity measures do not categorize this test. However, in agreement with the earlier study, which states that aging decreases the thoracic expansion capacity, measurements such as the respiratory signal arc length (L) help categorize the test. Concerning the influence of BMI on respiratory signals, it affects both components. Individuals with a higher BMI have a smaller respiratory amplitude and a higher RR. Finally, the CWP test is categorized by respiratory arc length, RR, and RRV.

After the statistical study had been applied to the processed signals, ML models were trained to try to distinguish gender and age, BMI and CWP range. In order to get the best possible results, not only three algorithms were used, but also four tests were applied. The tests were focus on the correlation value used in the statistical process, taking the values of 0.5 or 0.7, and on the number of individuals considered as new data, taking the values of 2 or 5 individuals. Analyzing the results, two tests stand out. Considering both tests using the 0.7 correlation, where just the number of subjects as new data changes, demonstrated that restricting even more features does not help to categorize the tests. Besides, features that are being used in test categorization at a correlation of 0.7 and not at a correlation of 0.5 are essential for test categorization since the results of tests using a correlation of 0.7 are better than those using a correlation of 0.5. The test a) scores better on gender and CWP characterization, and test b) best characterizes age and BMI.

Still, within the scope of the application of machine learning, it was possible to obtain the following values of accuracy in test data for the respective tests:

- **Gender:** 76.66%;
- **Age:** 71.13%;
- **BMI:** 72.52%;
- **CWP:** 74.61%.

Finally, the easiest test to classify is gender, followed by CWP, then BMI, and finally, age. The new subjects results are not as good as the test data, however, and taking into account that these results are above the 50%, it is possible distinguish subjects. In order to improve the new subjects results it is necessary increase the dataset. Therefore, the population will be better represented and the model can be trained with even more respiratory signals.

6.2 Future Work

Analyzing all the work developed and already looking toward the near future, it is possible to highlight four works:

- Include the cardiac signal in the study;
- Apply genetic algorithms or neural networks for the gender, age, BMI, and CWP distinguish;
- Biometric identification through Bio-Radar signals;
- Increase the dataset, and thus can also increase the age range.

Considering that the Bio-Radar system acquires both vital signs, the cardiac and the respiratory signal, it is possible to apply an algorithm for the extraction of the cardiac signal. In this way, one of the works derived from this dissertation applies to the study developed for the cardiac signal.

To try to improve the classification between gender, age, BMI, and CWP, other technologies could be used for this purpose, such as neural networks and genetic algorithms. Taking into account that in this dissertation is proved the existence of physiological and body stature variability, this is a step towards the study of biometric identification since this has already been performed in another system.

Since the individuals who constitute the dataset of this dissertation are aged between 18 - 50, it would be interesting to enrich the dataset by widening the age range. This new dataset is intended to verify whether the conclusions drawn from this data hold.

References

[1] Qiang An *et al.*, "A novel approach to simulate chest wall micro-motion for bio-radar life detection purpose", in *Target and Background Signatures II*, SPIE, vol. 9997, 2016, pages 263–268.

[2] James C. L., "Noninvasive microwave measurement of respiration", *Solid State Electronics*, vol. 63, pages 1530–1530, Feb. 1975.

[3] T. Li, T. Qiu, and H. Tang, "Optimum heart sound signal selection based on the cyclostationary property", *Computers in biology and medicine*, vol. 43, no. 6, pages 607–612, 2013.

[4] Carolina Gouveia *et al.*, "Study on the usage feasibility of continuous-wave radar for emotion recognition", *Biomedical Signal Processing and Control*, vol. 58, page 101 835, 2020.

[5] Xiaochao Dang, Zetong Chen, and Zhanjun Hao, "Emotion recognition method using millimetre wave radar based on deep learning", *IET Radar, Sonar & Navigation*, 2022.

[6] Li Zhang *et al.*, "Non-contact dual-modality emotion recognition system by CW radar and RGB camera", *IEEE Sensors Journal*, vol. 21, no. 20, pages 23 198–23 212, 2021.

[7] Hafeez Ur Rehman Siddiqui *et al.*, "Respiration based non-invasive approach for emotion recognition using impulse radio ultra wide band radar and machine learning", *Sensors*, vol. 21, no. 24, page 8336, 2021.

[8] Masayuki Kagawa, Kazuki Suzumura, and Takemi Matsui, "Sleep stage classification by non-contact vital signs indices using doppler radar sensors", in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2016, pages 4913–4916.

[9] F. Lin *et al.*, "Sleepsense: A noncontact and cost-effective sleep monitoring system", *IEEE transactions on biomedical circuits and systems*, vol. 11, no. 1, pages 189–202, Feb. 2016.

[10] L. Zhang *et al.*, "Sleep stages classification by CW doppler radar using bagged trees algorithm", in *2017 IEEE Radar Conference (RadarConf)*, IEEE, Seattle, WA, USA, May 2017, pages 788–791.

[11] Jie Liu *et al.*, "Non-contact human fatigue assessment system based on millimeter wave radar", in *2021 IEEE 4th International Conference on Electronics Technology (ICET)*, IEEE, 2021, pages 173–177.

[12] Georges Matar, Jean-Marc Lina, Julie Carrier, and Georges Kaddoum, "Unobtrusive sleep monitoring using cardiac, breathing and movements activities: An exhaustive review", *IEEE Access*, vol. 6, pages 45 129–45 152, 2018.

[13] Jaehoon Jung *et al.*, "Cnn-based driver monitoring using millimeter-wave radar sensor", *IEEE Sensors Letters*, vol. 5, no. 3, pages 1–4, 2021.

[14] Yu Rong, Kumar Vijay Mishra, and Daniel W Bliss, "Sparse processing for driver respiration monitoring using in-vehicle mmwave radar", in *2022 IEEE/MTT-S International Microwave Symposium-IMS 2022*, IEEE, 2022, pages 440–443.

[15] Ivan D Castro *et al.*, "Physiological driver monitoring using capacitively coupled and radar sensors", *Applied Sciences*, vol. 9, no. 19, page 3994, 2019.

[16] Gila Benchetrit, "Breathing pattern in humans: Diversity and individuality", *Respiration physiology*, vol. 122, no. 2-3, pages 123–129, 2000.

[17] A. Lomauro and A. Aliverti, "Sex differences in respiratory function", *Breathe*, vol. 14, pages 131–140, Jun. 2018.

[18] P. B. Dominelli *et al.*, "Dysanapsis and the resistive work of breathing during exercise in healthy men and women", *Journal of Applied Physiology*, vol. 119, no. 10, pages 1105–1113, Nov. 2015.

[19] P. H. Quanjer *et al.*, "Implications of adopting the global lungs initiative 2012 all-age reference equations for spirometry", *European Respiratory Journal*, vol. 42, no. 4, pages 1046–1054, Oct. 2013.

[20] J. Schwartz *et al.*, "Sex and race differences in the development of lung function", *American Review of Respiratory Disease*, vol. 138, no. 6, pages 1415–1421, 1988.

[21] P. H. Quanjer, G. L. Hall, S. Stanojevic, T. J. Cole, and J.others Stocks, "Age- and height-based prediction bias in spirometry reference equations", *European Respiratory Journal*, vol. 40, no. 1, pages 190–197, Jul. 2012.

[22] N. Torres-Tamayo *et al.*, "3D analysis of sexual dimorphism in size, shape and breathing kinematics of human lungs", *Journal of anatomy*, vol. 232, no. 2, pages 227–237, 2018.

[23] F. Bellemare, A. Jeanneret, and J. Couture, "Sex differences in thoracic dimensions and configuration", *American Journal of Respiratory and Critical Care Medicine*, vol. 168, no. 3, pages 305–312, May 2003.

[24] M. Romei *et al.*, "Effects of gender and posture on thoraco-abdominal kinematics during quiet breathing in healthy adults", *Respiratory physiology & neurobiology*, vol. 172, no. 3, pages 184–191, May 2010.

[25] X. Shi *et al.*, "A statistical human rib cage geometry model accounting for variations by age, sex, stature and body mass index", *Journal of biomechanics*, vol. 47, no. 10, pages 2277–2285, Apr. 2014.

[26] Tim JT Sutherland *et al.*, "The relationship between body fat and respiratory function in young adults", *European Respiratory Journal*, vol. 48, no. 3, pages 734–747, 2016.

[27] Gouveia C., "Bio-radar", M.S. thesis, Universidade de Aveiro, 2017.

[28] Medical News Today. "What to know about the cardiovascular system". (Jan. 2022), [Online]. Available: https://www.medicalnewstoday.com/articles/cardiovascular-system.

[29] A. Flint, *A practical treatise on the diagnosis, pathology, and treatment of diseases of the heart*. Philadelphia: Blanchard and Lea, 1859.

[30] Lubecke V. M. *et al.*, *Doppler Radar Physiological Sensing*. John Wiley & Sons, 2015.

[31] MedBullets Team Step1. "Muscles of respiration - respiratory - medbullets step 1". (Nov. 2021), [Online]. Available: https://step1.medbullets.com/respiratory/117007/muscles-of-respiration.

[32] E.H. Awtry *et al.*, "Evaluation of the patient with cardiovascular disease", *Cecil Essentials of Medicine, 5th edition, WB Saunders, Philadelphia*, pages 30–42, 2001.

[33] Gary P Carlson, "Fluid, electrolyte, and acid-base balance", in *Clinical biochemistry of domestic animals*, vol. 5, Elsevier, 1997, pages 485–516.

[34] DG Osmond, "Functional anatomy of the chest wall", *Lung biology in health and disease*, vol. 29, pages 199–233, 1985.

[35] JR Rodarte and FR Shardonofsky, "Respiratory system mechanics", *Textbook of Respiratory Medicine (3rd ed.) Philadelphia: Saunders*, pages 91–117, 2000.

[36] A. D. Droitcour, "Non-contact measurement of heart and respiration rates with a single-chip microwave doppler radar", Ph.D. dissertation, Stanford University, Jun. 2006.

[37] D. Reichl. "The benefits of deep diaphragmatic breathing". (Nov. 2021), [Online]. Available: https://www.mindcoolness.com/blog/deep-diaphragmatic-breathing-benefits.

[38] L. J. Brooks and K. P. Strohl, "Size and mechanical properties of the pharynx in healthy men and women", *American Review of Respiratory Disease*, vol. 146, no. 6, pages 1394–1397, 1992.

[39] L. J. Brooks *et al.*, "Relationship between lung volume and tracheal area as assessed by acoustic reflection", *Journal of Applied Physiology*, vol. 64, no. 3, pages 1050–1054, Mar. 1988.

[40] W. M. Thurlbeck, "Postnatal human lung growth", *Thorax*, vol. 37, no. 8, pages 564–571, Aug. 1982.

[41] J. F. Bellemare *et al.*, "Thoracic dimensions at maximum lung inflation in normal subjects and in patients with obstructive and restrictive lung diseases", *Chest*, vol. 119, no. 2, pages 376–386, Sep. 2000.

[42] H. Kaneko and J. Horie, "Breathing movements of the chest and abdominal wall in healthy subjects", *Respiratory care*, vol. 57, no. 9, pages 1442–1451, Sep. 2012.

[43] A. R. Fugl-Meyer, "Relative respiratory contribution of the rib cage and the abdomen in males and females with special regard to posture", *Respiration*, vol. 31, no. 3, pages 240–251, 1974.

[44] R. Gilbert, J. H. Auchincloss Jr., and D. Peppi, "Relationship of rib cage and abdomen motion to diaphragm function during quiet breathing", *Chest*, vol. 80, no. 5, pages 607–612, 1981.

[45] J. T. Sharp *et al.*, "Relative contributions of rib cage and abdomen to breathing in normal subjects", *Journal of Applied Physiology*, vol. 39, no. 4, pages 608–618, Oct. 1975.

[46] J. A. Verschakelen and M. G. Demedts, "Normal thoracoabdominal motions. influence of sex, age, posture, and breath size", *American journal of respiratory and critical care medicine*, vol. 151, no. 2, pages 399–405, 1995.

[47] Urooj Bhatti, Zulfiqar Ali Laghari, and Binafsha Manzoor Syed, "Effect of body mass index on respiratory parameters: A cross-sectional analytical study", *Pakistan Journal of Medical Sciences*, vol. 35, no. 6, page 1724, 2019.

[48] Carolyn S Ray *et al.*, "Effects of obesity on respiratory function", *American Review of Respiratory Disease*, vol. 128, no. 3, pages 501–506, 1983.

[49] Gulshan Sharma and James Goodwin, "Effect of aging on respiratory system physiology and immunology", *Clinical interventions in aging*, vol. 1, no. 3, page 253, 2006.

[50] Mayo Clinic. "Kyphosis". (2022), [Online]. Available: https://www.mayoclinic.org/diseases-conditions/kyphosis/symptoms-causes/syc-20374205\%7D\%7D.

[51] A Rossi *et al.*, "Aging and the respiratory system", *Aging Clinical and Experimental Research*, vol. 8, no. 3, pages 143–161, 1996.

[52] L Sandrini, Alessandro Vaccari, C Malacarne, L Cristoforetti, and Rolando Pontalti, "RF dosimetry: A comparison between power absorption of female and male numerical models from 0.1 to 4 ghz", *Physics in Medicine & Biology*, vol. 49, no. 22, page 5185, Oct. 2004.

[53] K Sasaki, K Wake, and S Watanabe, "Measurement of the dielectric properties of the epidermis and dermis at frequencies from 0.5 GHz to 110 GHz", *Physics in Medicine and Biology*, vol. 59, no. 16, pages 4739–4747, Aug. 2014.

[54] C. Gouveia, J. Vieira, and P. Pinho, "A review on methods for random motion detection and compensation in bio-radar systems", *Sensors*, vol. 19, no. 3, page 604, Jan. 2019.

[55] M. I. Skolnik, "Introduction to radar", *Radar handbook*, vol. 2, page 21, 1962.

[56] Øyvind A., "Radar monitoring of heartbeats and respiration", Ph.D. dissertation, Faculty of Mathematics and Natural Sciences, University of Oslo, 2013.

[57] S. Pisa, E. Pittella, and E. Piuzzi, "A survey of radar systems for medical applications", *IEEE Aerospace and Electronic Systems Magazine*, vol. 31, no. 11, pages 64–81, Dec. 2016.

[58] C. Gouveia *et al.*, "Different antenna designs for non-contact vital signs measurement: A review", *Electronics (Switzerland)*, vol. 8, no. 11, page 1294, Nov. 2019.

[59] Byung-Kwon Park, Olga Boric-Lubecke, and Victor M Lubecke, "Arctangent demodulation with dc offset compensation in quadrature doppler radar receiver systems", *IEEE transactions on Microwave theory and techniques*, vol. 55, no. 5, pages 1073–1079, 2007.

[60] Carolina Gouveia *et al.*, "Dynamic digital signal processing algorithm for vital signs extraction in continuous-wave radars", *Remote Sensing*, vol. 13, no. 20, page 4079, 2021.

[61] Ettus Research. "Usrp b200 (board only)". (Dec. 2021), [Online]. Available: https://www.ettus.com/all-products/ub200-kit.

[62] O. B. Lubecke, P. W. Ong, and V. M. Lubecke, "10 GHz doppler radar sensing of respiration and heart movement", in *Proceedings of the IEEE 28th Annual Northeast Bio engineering Conference*, IEEE, Philadelphia, PA, USA, Apr. 2002, pages 55–56.

[63] R. Ichapurapu *et al.*, "A 2.4 GHz non-contact biosensor system for continuous vital-signs monitoring", in *2009 IEEE 10th Annual Wireless and Microwave Technology Conference*, IEEE, Clearwater, FL, USA, Apr. 2009, pages 1–3.

[64] W. Massagram *et al.*, "Assessment of heart rate variability and respiratory sinus arrhythmia via doppler radar", *IEEE Transactions on Microwave Theory and Techniques*, vol. 57, no. 10, pages 2542–2549, Sep. 2009.

[65] H. Tan, D. Qiao, and Y. Li, "Non-contact heart rate tracking using doppler radar", in *2012 International Conference on Systems and Informatics*, IEEE, Yantai, China, May 2012, pages 1711–1714.

[66] W. Hu *et al.*, "Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor", *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pages 725–735, Nov. 2013.

[67] T. Rahman *et al.*, "Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar", in *The 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, vol. 10, Osaka, Japan, Sep. 2015, page 1145.

[68] J. Kuutti *et al.*, "Evaluation of a doppler radar sensor system for vital signs detection and activity monitoring in a radio-frequency shielded room", *Measurement*, vol. 68, pages 135–142, May 2015.

[69] J. Tu and J. Lin, "Fast acquisition of heart rate in noncontact vital sign radar measurement using time-window-variation technique", *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 1, pages 112–122, Jan. 2016.

[70] M. Li and J. Lin, "Wavelet-transform-based data-length-variation technique for fast heart rate detection using 5.8 GHz CW doppler radar", *IEEE Transactions on Microwave Theory and Techniques*, vol. 66, no. 1, pages 568–576, Jan. 2018.

[71] C. Will *et al.*, "Radar-based heart sound detection", *Scientific reports*, vol. 8, no. 1, pages 1–14, Jul. 2018.

[72] K. Yamamoto, K. Toyoda, and T. Ohtsuki, "Spectrogram-based non-contact rri estimation by accurate peak detection algorithm", *IEEE Access*, vol. 6, pages 60 369–60 379, Oct. 2018.

[73] J. H. Park *et al.*, "915 MHz continuous wave doppler radar sensor for detection of vital signs", *Electronics*, vol. 8, no. 5, page 561, Jul. 2019.

[74] J. Y. Kim *et al.*, "Peak detection algorithm for vital sign detection using doppler radar sensors", *Sensors*, vol. 19, no. 7, page 1575, Apr. 2019.

[75] S. Schellenberger *et al.*, "A dataset of clinically recorded radar vital signs with synchronised reference sensor signals", *Scientific data*, vol. 7, no. 1, pages 1–11, Sep. 2020.

[76] K. Shi *et al.*, "A dataset of radar-recorded heart sounds and vital signs including synchronised reference sensor signals", *Scientific data*, vol. 7, no. 1, pages 1–12, Feb. 2020.

[77] M. Kebe *et al.*, "Human vital signs detection methods and potential using radars: A review", *Sensors*, vol. 20, no. 5, page 1454, Mar. 2020.

[78] W. Xia, Y. Li, and S. Dong, "Radar-based high-accuracy cardiac activity sensing", *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pages 1–13, Jan. 2021.

[79] A. R. Diewald *et al.*, "RF based child occupation detection in the vehicle interior", in *2016 17th International Radar Symposium (IRS)*, IEEE, Krakow, Poland, 2016, pages 1–4.

[80] L. Anishchenko, "Challenges and potential solutions of psychophysiological state monitoring with bio-radar technology", *Diagnostics*, vol. 8, no. 4, page 73, Oct. 2018.

[81] Q. Gao *et al.*, "Non-contact emotion recognition via CW doppler radar", in *2018 Asia-Pacific Microwave Conference (APMC)*, IEEE, Kyoto, Japan, Nov. 2018, pages 1468–1470.

[82] Q. Liang *et al.*, "Research on non-contact monitoring system for human physiological signal and body movement", *Biosensors*, vol. 9, no. 2, page 58, Apr. 2019.

[83] O. Boric-Lubecke *et al.*, "Doppler radar sensing of multiple subjects in single and multiple antenna systems", in *TELSIKS 2005-2005 uth International Conference on Telecommunication in ModernSatellite, Cable and Broadcasting Services*, IEEE, vol. 1, Nis, Serbia, Sep. 2005, pages 7–11.

[84] Sungwon Yoo *et al.*, "Radar recorded child vital sign public dataset and deep learning-based age group classification framework for vehicular application", *Sensors*, vol. 21, no. 7, page 2412, 2021.

[85] Amy D Droitcour *et al.*, "Range correlation and i/q performance benefits in single-chip silicon doppler radars for noncontact cardiopulmonary monitoring", *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 3, pages 838–848, 2004.

[86] Delaram Jarchi *et al.*, "Validation of instantaneous respiratory rate using reflectance ppg from different body positions", *Sensors*, vol. 18, no. 11, page 3705, 2018.

[87] Lesya Anishchenko *et al.*, "Remote limb movement analysis during sleep by means of bioradar", in *2020 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, IEEE, 2020, pages 1–3.

[88] LN Anishchenko and EM Rutskova, "Estimation of rat's sleep-wake cycle using a bio-radar", in *2017 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, IEEE, 2017, pages 468–471.

[89] Tauhidur Rahman *et al.*, "Dopplesleep: A contactless unobtrusive sleep sensing system using short-range doppler radar", in *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pages 39–50.

[90] Lesya Anishchenko and Alina Turetzkaya, "Improved non-contact mental stress detection via bioradar", in *2020 International Conference on Biomedical Innovations and Applications (BIA)*, IEEE, 2020, pages 21–24.

[91] JAMES CHEN. "Skewness: Positively and negatively skewed defined with formula". (Jun. 2022), [Online]. Available: https://www.investopedia.com/terms/s/skewness.asp.

[92] Inês J., "Automatic audio signal analysis for the detection of anomalies in calls", M.S. thesis, Universidade de Aveiro, 2017.

[93] Taylor Arnold, Michael Kane, and Bryan W Lewis, *A computational approach to statistical learning*. Chapman and Hall/CRC, 2019.

[94] Brian Beers. "P-value". (May 2022), [Online]. Available: https://www.investopedia.com/terms/p/p-value.asp.

[95] Debbie L Hahs-Vaughn and Richard G Lomax, *Statistical concepts: A second course*. Routledge, 2020.

[96] Frost J., "A computational approach to statistical learning", *Journal of Chemical Information and Modeling*, vol. 9, 1689—1699, 2019.

[97] Vladimir N Vapnik, "Conclusion: What is important in learning theory?", in *The Nature of Statistical Learning Theory*, Springer, 1995, pages 167–175.

[98] Nello Cristianini *et al.*, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.

[99] Le Yu *et al.*, "Towards automatic lithological classification from remote sensing data using support vector machines", *Computers & Geosciences*, vol. 45, pages 229–239, 2012.

[100] Steven L Brunton and J Nathan Kutz, *Data-driven science and engineering: Machine learning, dynamical systems, and control*. Cambridge University Press, 2022.

[101] J. C. Chouinard. "K-nearest neighbors (knn) in python". (Sep. 2022), [Online]. Available: https://www.jcchouinard.com/k-nearest-neighbors.

[102] Leo Breiman, "Random forests", *Machine learning*, vol. 45, no. 1, pages 5–32, 2001.

[103] R. Abilash. "Applying random (classification) - machine learning algorithm from scratch with real datasets". (Sep. 2022), [Online]. Available: https://medium.com/@ar.ingenious/applying-random-forest-classification-machine-learning-algorithm-from-scratch-with-real-24ff198a1c57.

[104] Gisela Pinto *et al.*, "Multimodal emotion evaluation: A physiological model for cost-effective emotion classification", *Sensors*, vol. 20, no. 12, page 3510, 2020.

[105] MathWorks. "Cross-validation". (Sep. 2022), [Online]. Available: https://www.mathworks.com/discovery/cross-validation.html.

Summary of the subjects' physical statures description

Table A.1: Subjects' physical statures description.

ID	CWP [cm]	Height [cm]	Weight [kg]	BMI [kg/m^2]	Age	Gender	ID	CWP [cm]	Height [cm]	Weight [kg]	BMI [kg/m^2]	Age	Gender
ID01	112.0	180	91.0	28.09	27	M	ID47	129.5	179	120.0	37.45	47	M
ID02	88.0	164	58.0	21.56	27	F	ID48	105.5	170	80.0	27.68	23	F
ID03	84.0	157	54.0	21.91	22	F	ID49	120.0	169	103.0	36.06	42	F
ID04	115.0	177	97.0	30.96	27	M	ID50	112.0	172	103.0	34.82	22	F
ID05	90.0	164	60.8	22.61	22	F	ID51	108.0	183	91.0	27.17	38	M
ID06	94.0	155	62.0	25.81	25	F	ID52	95.0	155	63.0	26.22	45	F
ID07	98.0	168	69.0	24.45	25	M	ID53	83.0	162	48.0	18.29	23	F
ID08	87.0	166	55.0	19.96	26	M	ID54	94.0	161	58.0	22.38	24	F
ID09	96.0	162	63.0	24.01	25	F	ID55	94.0	164	66.1	24.58	37	F
ID10	104.0	179	87.0	27.15	26	M	ID56	115.0	186	108.0	31.22	41	M
ID11	85.5	165	57.0	20.94	21	M	ID57	101.0	163	71.0	26.72	43	F
ID12	101.0	190	84.0	23.27	20	M	ID58	91.0	162	60.0	22.86	46	F
ID13	114.0	170	98.0	33.91	27	M	ID59	112.0	174	93.0	30.72	48	M
ID14	87.0	164	60.0	22.31	28	F	ID60	106.0	165	68.5	25.16	47	F
ID15	114.0	165	90.0	33.06	23	F	ID61	87.0	164	54.0	20.08	45	F
ID16	110.5	164.0	85.0	31.60	31	M	ID62	110.0	172	85.0	28.73	45	M
ID17	102.0	183	76.0	22.69	26	M	ID63	92.0	175	72.0	23.51	49	M
ID18	118.0	158	100.0	40.06	23	F	ID64	92.0	155	55.0	22.89	45	F
ID19	102.0	180	85.0	26.23	23	M	ID65	81.0	170	56.0	19.38	20	M
ID20	102.0	164	65.0	24.17	18	F	ID66	106.5	183	83.0	24.78	22	M
ID21	115.5	159	81.5	32.24	19	F	ID67	109.0	175	100.0	32.65	43	F
ID22	85.0	160	56.0	21.88	18	F	ID68	100.0	178	75.0	23.67	46	M
ID23	97.5	184	85.0	25.11	29	M	ID69	95.0	173	70.0	23.39	28	M
ID24	91.5	170	62.0	21.45	26	M	ID70	82.0	163	50.0	18.82	18	F
ID25	97.5	158	62.0	24.84	22	F	ID71	101.0.	163	70.0	26.35	46	M
ID26	94.0	178	72.0	22.72	22	M	ID72	96.0	154	60.0	25.30	46	F
ID27	104.5	188	90.0	25.46	22	M	ID73	115.0	174	94.0	31.05	43	M
ID28	124.0	153	90.0	38.45	27	F	ID74	104.0	161	71.0	27.39	47	F
ID29	96.5	171	72.5	24.79	27	M	ID75	86.0	169	59.0	20.66	39	F
ID30	90.5	170	73.0	25.26	21	M	ID76	92.0	160	60.0	23.44	44	M
ID31	97.0	173	62.0	20.72	32	F	ID77	103.0	155	73.0	30.39	41	F
ID32	108.0	171	86.0	29.41	30	M	ID78	107.0	183	92.0	27.47	40	M
ID33	89.0	171	58.0	19.84	27	F	ID79	90.0	170	54.0	18.69	19	M
ID34	85.0	164	52.0	19.33	35	F	ID80	85.5	153	55.0	23.50	23	F
ID35	92.0	162	53.0	20.20	19	F	ID81	116.0	185	87.0	25.42	44	M
ID36	89.0	174	64.0	21.14	19	M	ID82	119.0	186	91.0	26.30	39	M
ID37	103.5	173	86.0	28.73	29	M	ID83	101.0	170	77.0	26.64	41	M
ID38	102.0	165	69.0	25.34	30	F	ID84	107.0	164	66.0	24.54	49	F
ID39	113.0	162	84.0	32.01	22	F	ID85	88.0	159	58.0	22.94	45	F
ID40	91.0	165	52.0	19.10	29	F	ID86	102.0	167	76.0	27.25	48	F
ID41	91.0	172	62.0	20.96	22	M	ID87	107.0	192	91.0	24.69	50	M
ID42	90.0	175	64.0	20.90	20	M	ID88	98.0	185	84.0	24.54	18	M
ID43	110.0	185	95.0	27.76	50	M	ID89	111.0	174	88.0	29.07	47	M
ID44	112.5	175	87.0	28.41	46	F	ID90	95.0	170	69.0	23.88	50	M
ID45	77.0	158	43.7	17.51	18	F	ID91	102.0	176	82.0	26.47	46	M
ID46	117.5	167	96.3	34.53	47	F	ID92	100.0	169	64.0	22.41	38	F

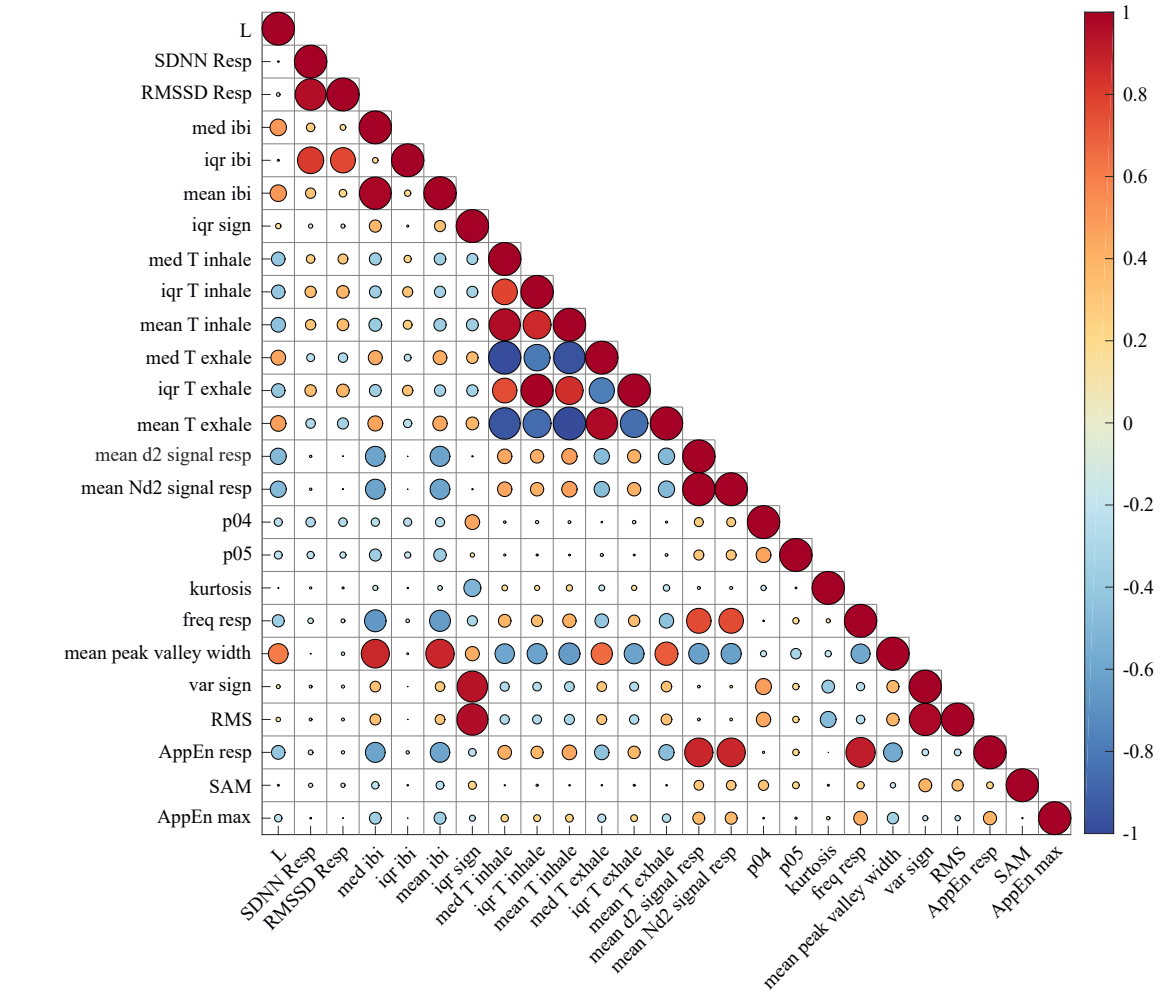
F - Female, M - Male, CWP - Chest Wall Perimeter, BMI - Body Mass Index

B

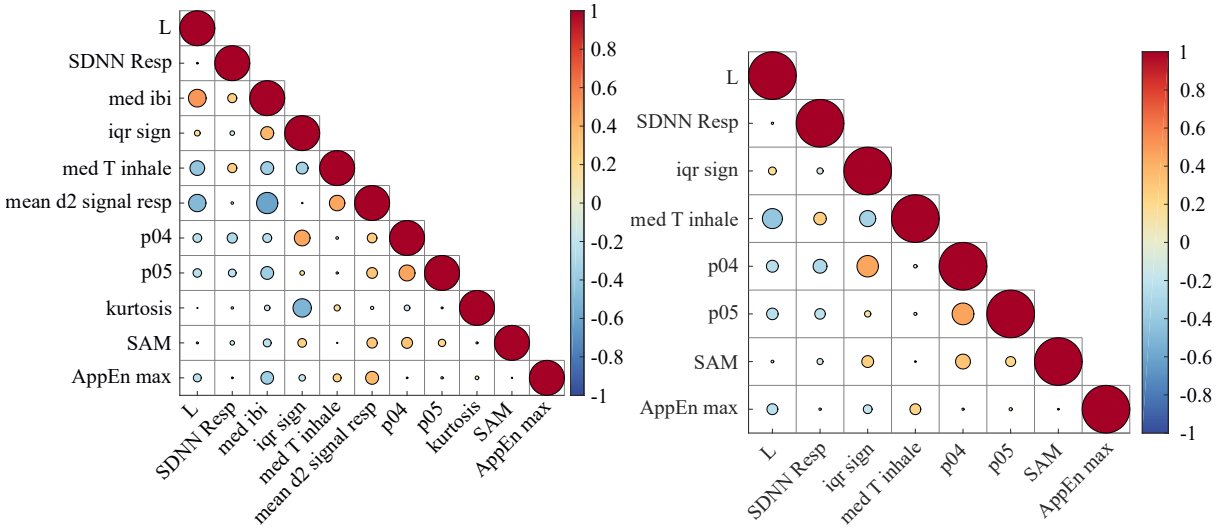
Figures and Tables of the Statistical Analysis Subsection

Table B.1: Priority lists of age test.

Initial		With 0.7 of correlation		With 0.5 of correlation	
Features	p-value	Features	p-value	Features	p-value
mean T exhale	5.28×10^{-08}	iqr sign	1.52×10^{-07}	iqr sign	1.52×10^{-07}
mean T inhale	5.28×10^{-08}	kurtosis	5.99×10^{-07}	med T inhale	8.68×10^{-07}
iqr sign	1.52×10^{-07}	med T inhale	8.68×10^{-07}	L	2.23×10^{-04}
iqr T exhale	1.77×10^{-07}	L	2.23×10^{-04}	SDNN Resp	4.61×10^{-03}
med T exhale	2.62×10^{-07}	SDNN Resp	4.61×10^{-03}	AppEn max	7.23×10^{-03}
iqr T inhale	2.95×10^{-07}	AppEn max	7.23×10^{-03}	p04	3.17×10^{-02}
kurtosis	5.99×10^{-07}	mean d2 signal resp	5.92×10^{-03}	SAM	3.43×10^{-02}
mean peak valley width	6.60×10^{-07}	med ibi	1.19×10^{-02}	p05	3.68×10^{-02}
med T inhale	8.68×10^{-07}	p04	3.17×10^{-02}		
var sign	9.02×10^{-07}	SAM	3.43×10^{-02}		
RMS	9.05×10^{-07}	p05	3.68×10^{-02}		
L	2.23×10^{-04}				
AppEn resp	1.41×10^{-03}				
iqr ibi	3.28×10^{-03}				
SDNN Resp	4.61×10^{-03}				
mean d2 signal resp	5.92×10^{-03}				
mean Nd2 signal resp	5.92×10^{-03}				
mean ibi	6.92×10^{-03}				
RMSSD Resp	7.15×10^{-03}				
AppEn max	7.23×10^{-03}				
freq resp	1.19×10^{-02}				
med ibi	1.19×10^{-02}				
p04	3.17×10^{-02}				
SAM	3.43×10^{-02}				
p05	3.68×10^{-02}				

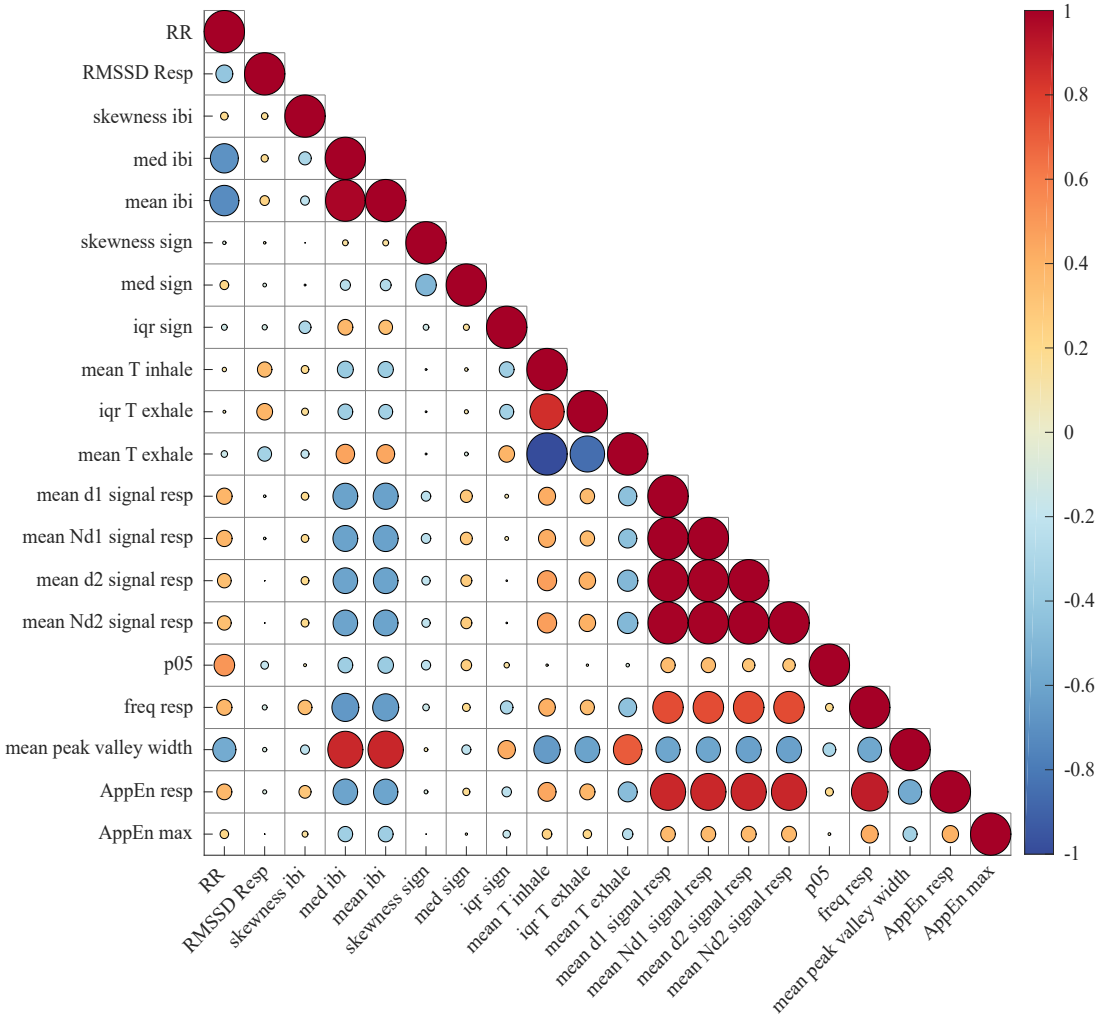


(a) Initial correlation matrix.

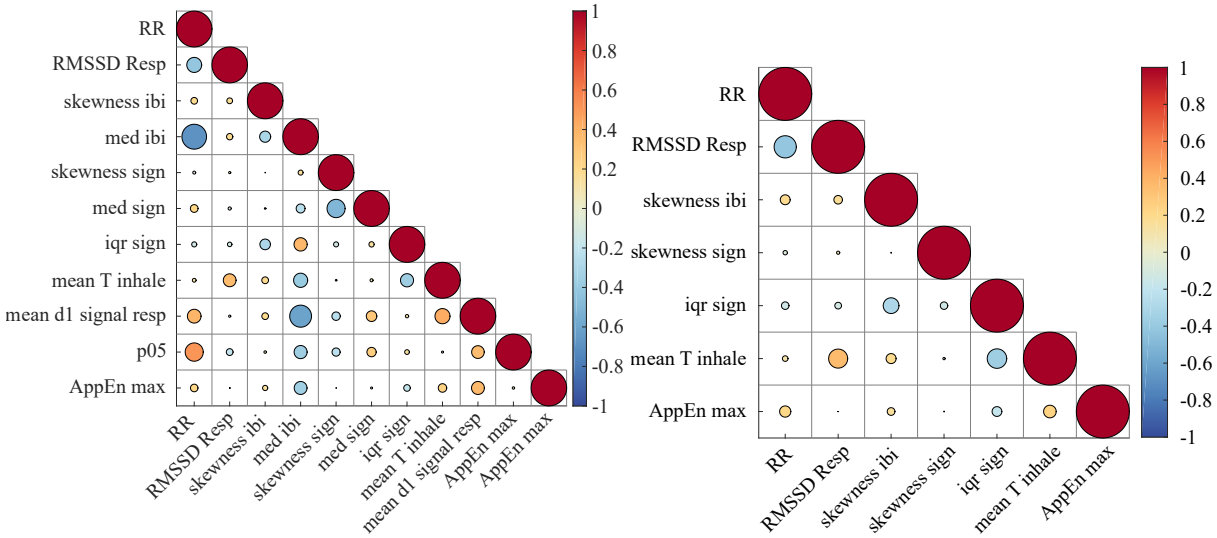


(b) Correlation matrix using 0.7 of correlation. (c) Correlation matrix using 0.5 of correlation.

Figure B.1: Correlation matrices associated at age.



(a) Initial correlation matrix.



(b) Correlation matrix using 0.7 of correlation.

(c) Correlation matrix using 0.5 of correlation.

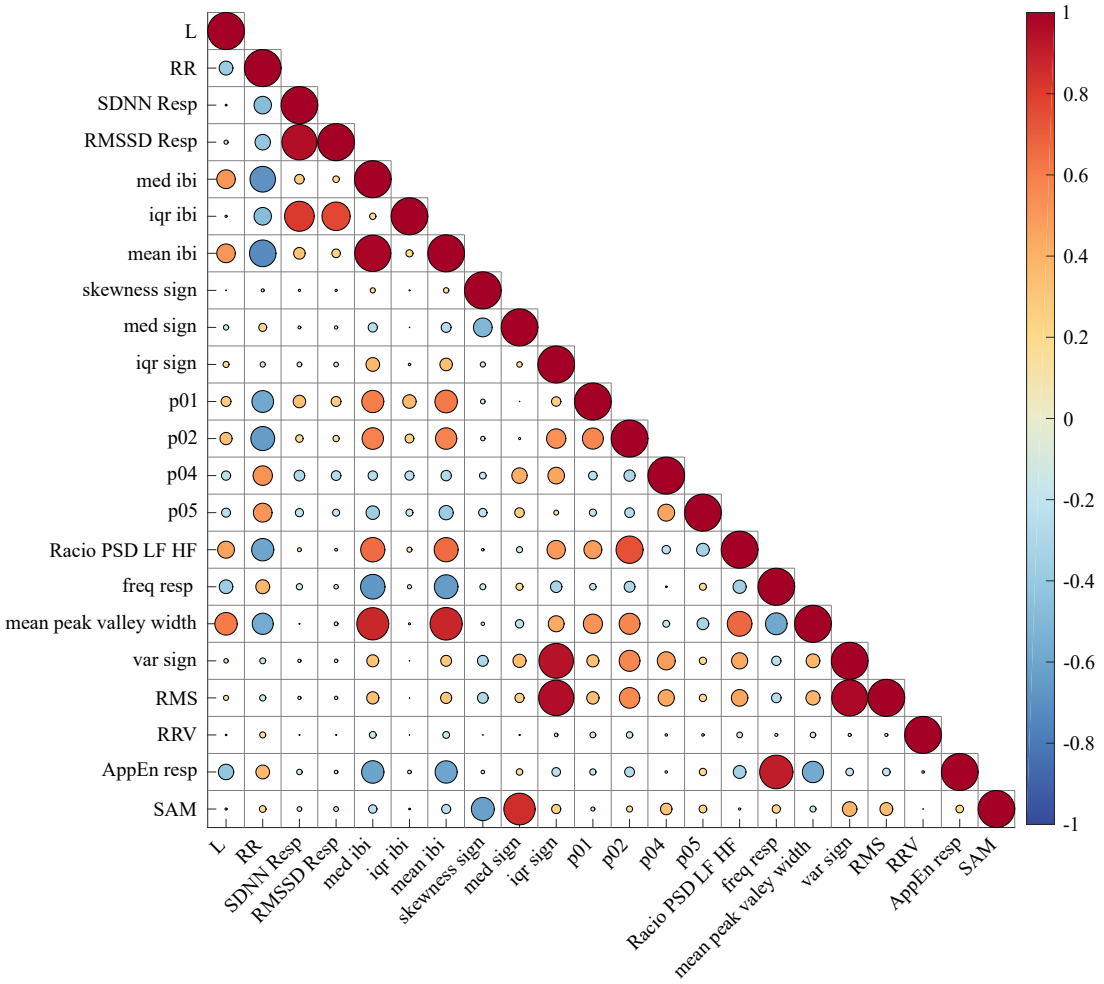
Figure B.2: Correlation matrices associated at BMI.

Table B.2: Priority lists of BMI test.

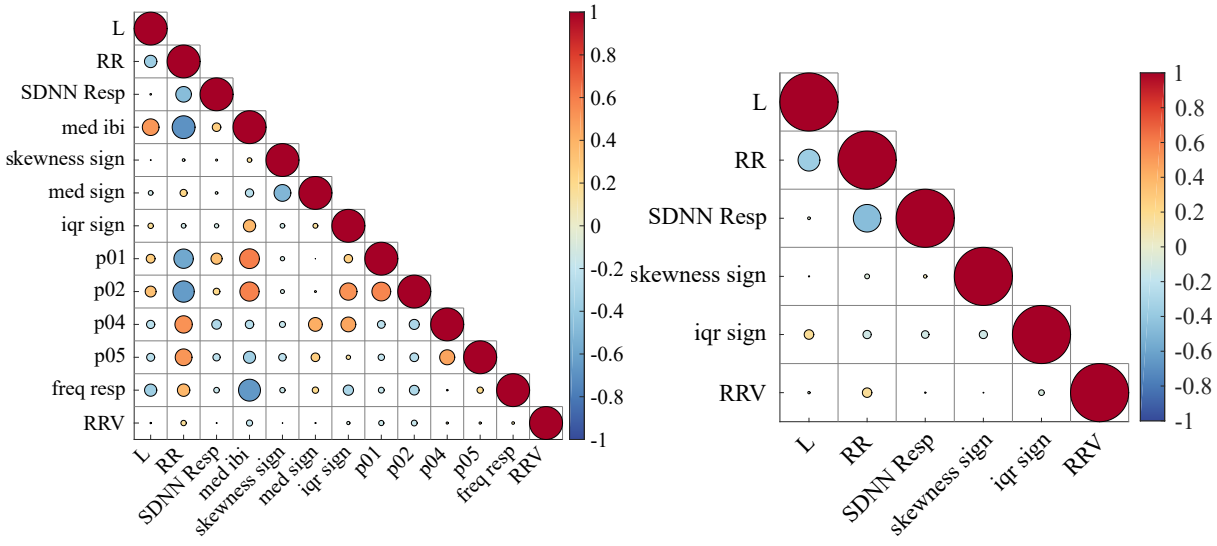
Initial		With 0.7 of correlation		With 0.5 of correlation	
Features	p-value	Features	p-value	Features	p-value
p05	1.04×10^{-05}	p05	1.04×10^{-05}	RR	9.79×10^{-03}
mean peak valley width	1.06×10^{-03}	RR	9.79×10^{-03}	AppEn resp	1.35×10^{-03}
AppEn resp	1.35×10^{-03}	AppEn resp	1.35×10^{-03}	skewness ibi	1.15×10^{-02}
RR	9.79×10^{-03}	mean d1 signal resp	1.09×10^{-02}	RMSSD Resp	2.98×10^{-02}
mean d1 signal resp	1.09×10^{-02}	skewness ibi	1.15×10^{-02}	mean T inhale	3.00×10^{-02}
mean Nd1 signal resp	1.09×10^{-02}	med ibi	1.70×10^{-02}	skewness sign	4.41×10^{-02}
skewness ibi	1.15×10^{-02}	RMSSD Resp	2.98×10^{-02}	iqr sign	2.35×10^{-02}
AppEn max	1.36×10^{-02}	med sign	2.99×10^{-02}		
freq resp	1.70×10^{-02}	mean T inhale	3.00×10^{-02}		
med ibi	1.70×10^{-02}	skewness sign	4.41×10^{-02}		
mean T exhale	2.16×10^{-02}	iqr sign	2.35×10^{-02}		
mean ibi	2.31×10^{-02}				
iqr sign	2.35×10^{-02}				
RMSSD Resp	2.98×10^{-02}				
med sign	2.99×10^{-02}				
mean T inhale	3.00×10^{-02}				
iqr T exhale	4.20×10^{-02}				
skewness sign	4.41×10^{-02}				
mean d2 sinal resp	4.94×10^{-02}				
mean Nd2 sinal resp	4.94×10^{-02}				
SAM					

Table B.3: Priority lists of CWP test.

Initial		With 0.7 of correlation		With 0.5 of correlation	
Features	p-value	Features	p-value	Features	p-value
RR	5.11×10^{-13}	RR	5.11×10^{-13}	RR	5.11×10^{-13}
p02	9.44×10^{-12}	p02	9.44×10^{-12}	skewness sign	2.84×10^{-05}
p05	3.44×10^{-11}	p05	3.44×10^{-11}	L	8.55×10^{-04}
p01	7.09×10^{-10}	p01	7.09×10^{-10}	iqr sign	2.11×10^{-03}
Racio PSD LF/HF	1.33×10^{-06}	skewness sign	2.84×10^{-05}	RRV	4.42×10^{-02}
SAM	8.58×10^{-06}	p04	3.79×10^{-05}	SDNN Resp	4.44×10^{-02}
skewness sign	2.84×10^{-05}	med sign	1.61×10^{-04}		
p04	3.79×10^{-05}	L	8.55×10^{-04}		
var sign	7.91×10^{-05}	freq resp	1.55×10^{-03}		
RMS	7.92×10^{-05}	med ibi	1.55×10^{-03}		
med sign	1.61×10^{-04}	iqr sign	2.11×10^{-03}		
L	8.55×10^{-04}	RRV	4.42×10^{-02}		
freq resp	1.55×10^{-03}	SDNN Resp	4.44×10^{-02}		
med ibi	1.55×10^{-03}				
iqr sign	2.11×10^{-03}				
mean ibi	2.67×10^{-03}				
iqr ibi	1.18×10^{-02}				
RMSSD Resp	1.86×10^{-02}				
mean peak valley width	1.91×10^{-02}				
AppEn resp	2.52×10^{-02}				
RRV	4.42×10^{-02}				
SDNN Resp	4.44×10^{-02}				



(a) Initial correlation matrix.



(b) Correlation matrix using 0.7 of correlation. (c) Correlation matrix using 0.5 of correlation.

Figure B.3: Correlation matrices associated at CWP.