# MECHANISM FOR THE CHARACTERIZATION AND REMOVAL OF LOW-INTENSITY MOVEMENT ARTIFACTS ON PPG SIGNALS TO SUPPORT TACHYCARDIA AND BRADYCARDIA DETECTION.


Undergraduate Thesis

Sebastián Andrés Cajas Ordoñez
María Alejandra Landínez García

Project director

PhD. Diego Mauricio López.

**University of Cauca**
**Faculty of Electronic Engineering and Telecommunications**
**Telematics Engineering Department**
**Group in Telematics Engineering – GIT**
**e-Health Research line**
**Popayan, August of 2019**

# MECANISMO PARA LA CARACTERIZACIÓN Y ELIMINACIÓN DE ARTEFACTOS POR MOVIMIENTO DE BAJA INTENSIDAD EN SEÑALES PPG PARA SOPORTAR LA DETECCIÓN DE TAQUICARDIAS Y BRADICARDIAS.


Trabajo de Grado

Sebastián Andrés Cajas Ordoñez
María Alejandra Landínez García

Director del Proyecto

PhD. Diego Mauricio López.

**Universidad del Cauca**
**Facultad de Ingeniería Electrónica y Telecomunicaciones**
**Departamento de Ingeniería Telemática**
**Grupo de Ingeniería Telemática - GIT**
**Línea de investigación en e-Salud**
**Popayán, Agosto de 2019**

# Table of contents

# List of tables

# List of figures

# List of abbreviations

| | |
|---|---|
| **WHO** | World Health Organization. |
| **ICER** | Incremental cost effectivity rate. |
| **QALY** | Quality adjusted life year. |
| **ECG** | Electrocardiogram. |
| **PPG** | Photoplethysmogram. |
| **OR** | Operation Room. |
| **ER** | Emergency Room. |
| **PACU** | Post Anesthesia Care Unit. |
| **ICU** | Intensive Care Unit. |
| **STFT** | Short Time Fourier Transform. |
| **CWT** | Continuous Wavelet Transform. |
| **DWT** | Discrete Wavelet Transform. |
| **MRA** | Multi Resolution Analysis. |
| **WT** | Wavelet Transform. |
| **VT** | Ventricular Tachycardia. |
| **VFB** | Ventricular Flutter Fibrillation. |
| **ABP** | Arterial Blood Pressure. |
| **SQI** | Signal Quality Index. |
| **TPR** | True Positive Rate. |
| **FPR** | False Positive Rate. |
| **TNR** | True Negative Rate. |
| **NFEEMD** | Notch Filtered Ensemble Empirical Mode Decomposition. |
| **EMD** | Empirical Mode Decomposition. |
| **IMF** | Intrinsic Mode Function. |
| **BOW** | Better or Worse. |
| **MAF** | Moving Average Filter. |
| **PMAF** | Periodic Moving Average Filter. |
| **LPC** | Linear Predictor Coefficient. |
| **TP** | True Positives. |
| **FP** | False Positives. |
| **TN** | True Negatives. |
| **FN** | False Negatives. |
| **NLMS** | Normalized Least Mean Square. |
| **LMS** | Least Mean Square. |
| **RLS** | Recursive Least Squares. |
| **IEEE SPC** | IEEE Signal Processing Cup. |
| **MSE** | Mean Squared Error. |
| **RMSE** | Root of the Mean Squared Error. |
| **MAD** | Mean Absolute Deviation. |
| **PSNR** | Peak to Signal Noise Ratio. |
| **WPPG** | Wrist PPG. |

# CHAPTER 1. INTRODUCTION.

## 1.1. PROBLEM STATEMENT.

### 1.1.1. Definition.

An arrhythmia can be defined as an increment or decrement of the regular heart electrical impulses sequence, forcing the organ to beat faster, slower or erratically. This irregularity affects other essential organs of the body, such as the lungs or brain, due to poor blood perfusion. There are several types of arrhythmias which can damage different regions of the heart, either auricular or ventricular. These can cause premature contractions, or make the heart works faster or slower than it usually does [1].

According to the World Health Organization (WHO), heart diseases have been the leading cause of death all over the world, prevailing in this ranking from 2000 to 2015. An increase in their incidence has been reported, from 6.8 to 8.5 millions of affected people in a period of one year. The consequences of asymptomatic arrhythmias and non-clinically recognized are arising, especially in the elderly population [2].

Ischemic heart disease and stroke added up to 15 million deaths, approximately 26% of the total deaths worldwide (56.4 million). The following leading cause of death are obstructive respiratory chronic diseases, lung cancer, diabetes, and dementia. Through advances in the pharmaceutical handling of the disease, a 30% reduction of these numbers is expected between 2010 and 2020. However, in low-median income countries there are limitations to the availability and accessibility to proper medication (aspirin, statins, and blood pressure reducing agents) [2].

Usually, an arrhythmia does not lead to an immediate change in health or lifestyle of the patient, but it produces a high increment in morbidity and consequently in mortality. In addition, the risk of other illnesses like thromboembolism or strokes is increased in older adults, who are more likely to suffer from arrhythmias. Furthermore, the disease is linked to other diagnostics, such as diabetes, hypertension, chronic fatigue, among others [3].

The incidence rate[1] has increased in the last decade, in nearly 2.000 cases per 100.000 inhabitants, both men and women. The worldwide estimate of the prevalence rate[2], which increases with age, has been noted to be 0.4%, reaching up to 2.5% in people over 60 years [4]. At least 2.7 million of U. S. citizens have manifested not being told of having some of these arrhythmias until finally going to a specialized medical check-up. This specially because some of them are asymptomatic and the person does not expect to have it. Therefore, it is essential that these disorders are correctly identified, as they can cause significant damage when left unnoticed and uncorrected, thus requiring numerous and expensive treatments [2, 5]. In

[1] Prevalence rate is defined as the proportion of persons in a population who suffer a particular disease or attribute over a specified period of time. This measure contains all the cases, both new and pre-existing.

[2] Incidence rate is defined as the occurrence of new cases of diseases in a population over a specified period of time.

Colombia, around the decade of 2000-2010, the cases of a particular type of arrhythmia, atrial fibrillation, increased by 10.4%, ascribing 1.995 deceases. The costs of this arrhythmia are distributed along the treatment, between hospital care and emergency visits, due to chronic decompensations. Continuous acquisition of anticoagulants like warfarin and other oral anticoagulants as dabigatran are required. This last anticoagulant demonstrates an increasing record of 0.37 years of life gained for the 150 mg dose and an incremental cost-effectivity rate (ICER) of $23.078.506 COP per year of life adjusted to quality (QALY) [6].

The primarily used method for arrhythmia detection is the electrocardiogram (ECG). Other diagnostic methods are heart palpation or pulse examination, yet their detection criteria are purely subjective, and thus less effective [3]. This kind of cardiac monitoring has been made through methods that include ECG, using specific instruments for this labor (e.g., the Omron HCG-801 and Merlin ECG event recorder), or adapting devices for this task (blood pressure monitors e.g., WatchBP). These two methods showed similar results regarding the accuracy of the measurements and the correct diagnosis for people who suffered from a particular kind of arrhythmia. Specifically, the sensitivity of both ECG meters and the adapted blood pressure monitor is 94.5%, 93.5%, and 94.9%, respectively [3].

Therefore, numerous detection methods with ECG signals and studies have been proposed. These methods allow early detection of arrhythmias and thus, the development of applications that grant a classification of this arrhythmias by type [7]. In some cases, the creation and later training of algorithms is decided using datasets extracted from open databases such as Physionet or MIT-BIH arrhythmia database. Nevertheless, in practice, the continuous use of electrocardiography monitors or holters[3] is too expensive and uncomfortable as the patient needs to remain as still as possible to avoid damaging the electrodes. The use of electrocardiography is also time consuming as the electrodes are placed strategically by cardiovascular health professionals. Therefore, the use of new techniques becomes necessary as they allow continuous monitoring of the patient with equivalent, or even enhanced accuracy of execution [8].

One of these new techniques is photoplethysmography (PPG). PPG is a non-invasive, electro-optic method that enables sensing of peripheral zones of the body. This technique generates a time signal, which represents the changes of blood volume flowing through the veins. This signal is obtained by illuminating specific areas of the body: earlobes, fingertips or the forehead, with a LED emitted light on a given wavelength. There are two possible modes of sensing in PPG: transmission or reflection. The obtained signal is the wave reflected by all of the tissue within the measuring zone, therefore it carries both DC and AC components. The direct current or DC component reflects all of the solid tissue like skin, muscle, and bone, which remain unaltered. The alternating current or AC component corresponds to the volumetric changes of blood [9].

The arrhythmia detection could be possible through PPG. The process of detection is commonly made through a comparison of the results of the R peaks detection algorithms for an ECG signal, and the M peaks detection of PPG signals. As it will be seen later in subsection 2.1.4., the M peak is the peak value of a PPG signal, and it is related to the ejecting moment

[3] A Holter is a device used for cardiologic diagnose tests, which consist of ambulatory monitoring of the electrocardiographic record during a long time, generally twenty-four hours in a moving person.

during systole. Intervals between M peaks are highly correlated to those between R peaks in the ECG signal [9]. The R peak is the peak value that the ECG wave reaches along the QRS complex. These intervals between R peaks are called R-R intervals, and refers to the time elapsed during a heartbeat.

These results have shown that the algorithms for the detection of specific arrhythmia through ECG monitoring are less efficient than those developed with PPG algorithms either in terms of sensitivity, precision, or specificity [8]. The PPG signals algorithms have improved over time. Currently, several advances are being developed in many abnormalities' detection. This is because PPG provides large sets of information not only at a cardiac level but also at a breathing, blood pressure and oxygen saturation level. The main PPG algorithms found in state of the art reviews were: ASM Algorithm (Adaptive Segmentation Method), the IBI algorithm (Inter-Beat-Intervals), and the peak detection algorithm [9,10,11].

One of the main limitations for the arrhythmia detection with PPG signals, is the lack of identification, classification, and handling of data with real patients, leaving as the only dataset option the one found in Physionet, which implements Atrial Fibrillation (AF). This limitation is stated clearly in Solosenko's work. He uses this challenging situation as the reason for the modeling of the PPG signal from an ECG signal sequence, taking as the main feature, the R-R intervals [12,13].

Additionally, the PPG signals are not immune to noise, they might even be more prone to be affected by it due to the places in the body where the measurements take place. The type of conditions that create a distorted or unclear signal are called artifacts, and its removal is troublesome since its high-frequency components often overlap with those of the PPG signal [14]. Other types of artifacts are: Environmental artifacts (electromagnetic interference, light interference, thermic noise, shot noise, etc.); Experimental error artifacts (those which are not controllable, such as the patient movement or the friction of the sensor with the finger), and physiological artifacts (eye movement or small signals caused by muscle tension or brain signals) [15].

This work focuses on experimental artifacts, more specifically the movement artifacts of the patient. These artifacts are the most challenging to extract from the signal since these cannot be found in a narrow band as the artifacts' spectrum overlaps with the signal's spectrum [15]. As a consequence, we have found in the literature many algorithms that allow cleaning the signal. However, it has not been set a specific distinction about the context where the person encounters in most of the cases. Therefore, an accurate movement description cannot be made about a specific artifact or corrupted signal. The main algorithms present in the literature are the Notch Filtered Ensemble Empirical Mode Decomposition (NFEEMD), filter banks, adaptive filters, neural networks for attribute selection, and the transformations in the frequency spectrum with wavelet processing [16,17,18,19,20,21].

In the same way, it is also challenging to process and manipulate corrupted signals with noise movement due to lack of public databases that can be used to do this. Just one database that meets the most basic conditions has been found, because we sought to delimitate the noise caused by specific kinds of movements.

In *Chapter 2*, the works focused on the detection of motion artifacts in PPG signals [17, 21, 29], varying in their intensity depending on the activity or movement that was made. Most of these

studies do not focus on the detection of arrhythmias of any type but are dedicated to determine the heart rate from the corrupted signal. The heart rate determination has been performed in two ways. The first is removing some intervals of the signal that exceed the threshold of noise allowed by them. The second is manipulating and processing the signal mathematically, achieving a new one that has a clear form and can be readable to future algorithms.
The definition of the word arrhythmia is very wide. We will put a particular focus on two cardiopathies which have a relationship with the heart rate, a feature that can be easily identified in the middle of the Inter Beat Intervals (IBI) in ECG or the Inter Pulse Intervals (IPI) in PPG (Further definitions for IPI and IBI will be given in *Chapter 2*). Bradycardia and tachycardia are heart rate disorders. In either case, the heart rate slows down too much, or it increases unexpectedly and reaches an abnormally fast rhythm [32, 33]. There are three types of tachycardia, the difference lies on the location of the cardiac cavity, or the way the cardiac frequency is affected [35]. Therefore, the proper diagnostic of these two conditions, bradycardia and tachycardia, is needed. This is made so that a broader range of heart diseases can be localized.
A single study has been found where the aim is to detect episodes of bradycardia or tachycardia over movement conditions (free-living conditions) [19]. However, its performance its comparatively low, reaching a sensitivity of 85% for bradycardia and 89% for tachycardia compared to [26] using a fuzzy entropy approach. According to Bonomi, low-intensity movements generate readings that can be masked as pulses, generating a misinterpretation of the increase of the heart rate. As a consequence, these misinterpretations are classified as tachycardia, increasing the number of false positive diagnosis. In the case of bradycardia, when the patient is at rest, it is possible that several beats are lost, converging to an overestimation of episodes of bradycardia.

### 1.1.2. Research question.

Considering the current state of the art of the arrhythmia diagnosis, our research question is defined as: ″¿How to improve the sensitivity in the bradycardia and tachycardia detection with PPG signals in the presence of low-intensity movements?"

As a hypothesis, it is set that the sensitivity of bradycardia and tachycardia detection in PPG signals can be improved, through the construction of a mechanism for the description and subsequent artifact removal of corrupted signals due to low-intensity movement artifacts in PPG signals.

### 1.1.3. Justification.

Due to the asymptomatic nature of arrhythmias, it is necessary to make an opportune diagnostic of them. In this way, we avoid the existence of more severe consequences.
Although there are various techniques for diagnosing, such as the ones exposed in *subsection 1.1.1.*, these generate an elevated cost for the patient, given that the devices used for ECG measurements are usually fragile and the acquisition costs for them are high. As an example, we can name holters or multi-derivation ECG measurement devices.

## 1.2. OBJECTIVES

### 1.2.1. General Objective

Propose a mechanism for the characterization and elimination of artifacts due to low-intensity movements in PPG signals to support the tachycardia and bradycardia detection.

### 1.2.2. Specific Objectives

**1.** Construct a simulated dataset of bradycardia and tachycardia, with artifact annotations due to low-intensity movements.
**2.** Create an assembled mechanism through the selection and evaluation of different algorithms, for the removal of low-intensity noise and the detection of bradycardia and tachycardia events in the constructed dataset.
**3.** Evaluate the sensitivity of the proposed mechanism in the detection of tachycardia and bradycardia with PPG signals in the presence of low-intensity movements experimentally.

## 1.4. METHODOLOGY

**Phase 1:** We used CRISP-DM as methodology for the specific objectives 1 and 2 [27]. The following steps are taken:

**1. Business understanding:** In this stage are included the objectives of the investigation, the success criteria, restrictions and costs. This stage is required as a proper outline of the problem is needed. In this way, a structure for the preliminary plan for the project is made.

**2. Data understanding:** This stage is focused on collecting the data, analyzing it, and finding preliminary ideas of the possible associations (dimensions, fields, or columns). In the same way, it seeks to make a description and verify the quality of the minable data.

**3. Data preparation:** In this stage, the data for the dataset generation is cleaned and prepared. Here, the attribute selection is made, where a connection between different attributes can be found, or interesting subsets can be identified.

**4. Modeling:** During this stage, the modeling techniques are selected, and the design of the model to implement is made. In this case, various algorithms are modeled in order to process and manipulate a dataset using different perspectives or techniques.

**5. Evaluation:** After having built the models which seem to reach high quality from a data analysis, it must be evaluated and review if it achieves the objectives. Then, the convenient tests are made, and the better alternative is chosen.

**6. Implementation:** during this phase is intended to implement the system/mechanism/ model; in its entirety as the solution to the business problem initially stated. This system/mechanism/ model will count with the steps before mentioned. Thus, the data can be treated using the selected process and obtain the best possible results.

## 1.5. STRUCTURE OF THE MONOGRAPH

The present monograph has been structured in the following chapters:

**Chapter 1:** Introduction. Here, the statement of the problem is presented, the justification, the objectives of this work, and the methodology.

**Chapter 2:** State of the art. Here, the essential concepts for the understanding of this work are described. Then it's exposed the state of the art about the methods for artifact noise reduction over PPG signals and the scenarios of the measurements.

**Chapter 3:** Characterization of Movement artifacts. Here, the state of the art about current methods for movement artifact noise characterization on PPG signals and another type of PPG-resembled biomedical signals is presented. Then, our methods for noise characterization are proposed and evaluated.

**Chapter 4:** Tachycardia and bradycardia detection mechanism. Here we present the state of the art about the most used techniques in the denoising of PPG signals and a review of the methods for tachycardia and bradycardia detection from the Physionet/CinC 2015 Challenge. Then, the most prominent denoising techniques are retrieved to be used over the chosen tachycardia and bradycardia detection algorithms under the induced noise conditions.
A comparison of the performance of the mechanisms with and without noise conditions is made, using the denoising techniques studied. The best results for each mechanism are chosen. Finally, the analysis of the general improvements for tachycardia and bradycardia detection is made.

**Chapter 5:** Conclusions and outlook. Here, we present the conclusions and contributions of our study, and give an outlook for future research.

# CHAPTER 2. STATE OF ART

This chapter presents an analysis of related works about methods for movement artifact noise reduction and the scenarios where the measurements were taken. This review is done through a systematic review of articles in bibliographic databases. We seek to define the gaps of the conducted studies, taking into account elements such as the conditions and the devices used in the measurement taken. We also look to define the type of artifacts considered in these studies, as well as the techniques they use to remove them or evade the negative effect of them over each work's purpose. First, medical concepts about PPG signals are exposed. These are important for the development of the proposed mechanism in this thesis. Then, we explain the selection criteria for the articles, the methodology of the review and the data analysis.

## 2.1. THEORETICAL FRAMEWORK

This section explains briefly the most relevant clinical concepts for the development of this thesis work. This will ease the comprehension of subsequent chapters and sections.

### 2.1.1. Circulatory system and cardiovascular system

Circulatory system is one of the main systems of the human body. It's conformed by the cardiovascular system and the lymphatic system. The first one conducts and irrigates the blood through all the body, and the second one conducts the lymph unidirectionally towards the heart [36].
The cardiovascular system is conformed by a central propulsion organ: the heart, and a closed-circuit of tubes. This closed-circuit is formed by the arteries (which conduct the blood from the heart to the organs); the capillaries and sinusoids (where the interchange of water, solutes, and gases is made between the system and the tissues); and the veins (that return the blood towards the heart) [37].
The main functions of the cardiovascular system are [38]:

- The fast transport of the nutrients and waste products.
- Hormonal control through the transport of hormones.
- Body's defense through the transportation of immune cells, antigens, and other mediators.
- Regulation of body temperature by blood irrigation between the center of the organism and the skin.

*Figure 2.1* exposes a schematic view of the primary and secondary circulatory system. The primary or arterial system (in red) transports oxygen-rich blood and the secondary or venous (in blue) transports oxygen-poor blood.

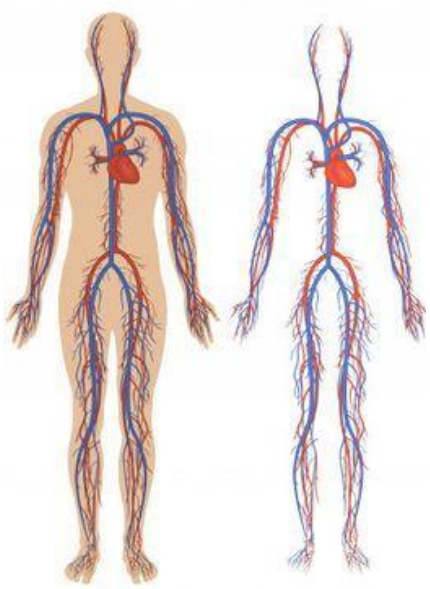


**Figure 2.1. Primary and secondary cardiovascular system.**

### 2.1.2. The heart

The heart is a muscular organ, the most important in the cardiovascular system. It's located in the interior of the thorax, above the diaphragm and slightly at the left of the sternum. It's a hollow, self-controlled muscle and has the shape of a cone resting on its side. It functions as a double bomb, an aspirant one and an ejecting one. It's made up of four cavities: two superiors and two inferiors. Superior ones are called atriums (or auricles), and inferior ones are called ventricles. [39]

The heart impulses and returns the blood through the systole and diastole movements. Systole is a contraction movement either from the auricles or the ventricles, and it's the one that ejects blood to the body. Diastole is a relaxation movement of the heart for receiving the blood coming from the rest of the body. [40]

The structure of the heart is composed of three layers: epicardium (exterior and thin layer), myocardium (central and thick layer) and endocardium (interior and thin layer) [41]. There are four valves that control the blood flow through the heart [42]:

- The tricuspid valve controls the blood flux between the right auricle and the right ventricle.
- The pulmonary valve controls the blood flux from the right ventricle to the pulmonary arteries. These pulmonary arteries transport blood to the lungs for their oxygenation.
- The mitral valve allows the rich oxygen blood coming from the lungs to pass from the left auricle to the left ventricle
- The aortic valve allows the rich oxygen blood to pass from the left ventricle to the aorta (the biggest artery in the body). This artery transports blood to the rest of the body.

The locations of these valves and mentioned cavities are exposed in *Figure 2.2.*

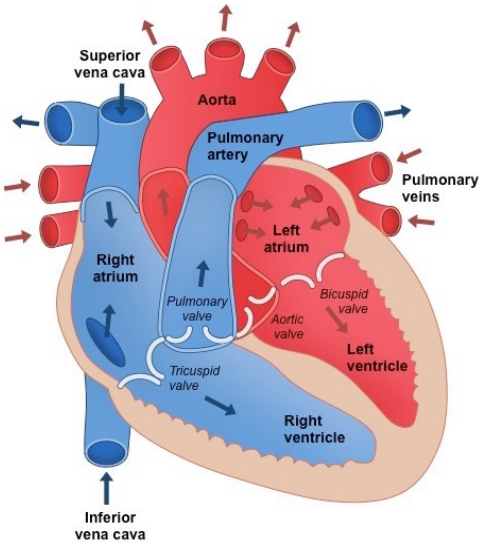



**Figure 2.2. The internal structure of the heart *[42]*.**

### 2.1.3. Cardiac cycle

The cardiac cycle is the series of all electrical and mechanical events that occur during each heartbeat. This includes the contraction and relaxation of the auricles and ventricles, the open and close actions of the valves and other minor actions [39].
The term systole refers to the contraction phase and the term diastole refers to the relaxation phase. Each cardiac cycle consists of auricular systole and diastole, and ventricular systole and diastole. In each cycle the auricles and ventricles contract and relax alternatively, making the blood flow from areas of lower pressure to areas of higher pressure [39]. The phenomena that occur during each cardiac cycle can be scheduled as follows [43]:
Systole period includes:

- Atrial contraction.
- Isovolumic contraction.
- Ejection.

Diastole period includes:

- Isovolumic relaxation.
- Rapid ventricular filling.
- Slow ventricular filling (diastasis).

A most detailed explanation of each one of the phases of the cardiac cycle can be seen in [44].

### 2.1.4. Photoplethysmography (PPG)
#### 2.1.4.1. Introduction

Photoplethysmography is a technique described for the first time by Alrick Hertzman from the Department of Physiology at St. Louis University School of Medicine, on 1937 [45]. PPG

waveform is the base technology for the pulsioximeter and thus, the ABP signals. This wave is shown and used in the monitors throughout the critical care areas of a hospital (operating room (OR), emergency room (ER), post anesthesia care unit (PACU), intensive care unit (ICU), etc.) [47]. However, it is rarely recorded or analyzed in comparison with other signals such as ECG. Along the years, it has been studied and analyzed, concluding in various useful applications. Some of these applications are: venous assessment, evaluation of aging and vascular disease, assessment of orthostasis and more lately, diagnosis of several cardiac arrhythmias [46].

#### 2.1.4.2. Photoplethysmogram

Photoplethysmogram is the graphic representation of the photoplethysmography measurements. Photoplethysmography is a simple optical technique used for discovering volumetric changes of the blood in the peripheral circulation. It has a low cost and it's a non-invasive method because it makes measures in the skin surface [48]. A coupled photodiode and phototransistor are used for the signal acquisition from the capillaries of the patient. [52]
When the heart beats, capillaries expand and contract based on blood volume changes. Light is emitted from PPG sensor photodiode and it is reflected onto the skin to accurately and continuously measure weak blood flow signals [49].
PPG technique uses low intensity infrared green light. When light travels through biological tissues it is absorbed by bones, skin pigments and both venous and arterial blood. Given that light is more strongly absorbed by blood than the surrounding tissues, the changes in blood flow can be detected as changes in the intensity of light by PPG sensor phototransistor [52].
There are various forms of taking the measurements of the PPG signal as shown in *Figure 2.3.*, these are transmission mode and reflection mode.
In the transmission mode, light goes through all tissue, and this is received with several changes in a photoreceptor at the other side of the area. The reflection mode has both the transmitter and photoreceptor on the same side; thus it receives the reflected wave by the tissue [47]. The first one of these modes is the most used currently. An extra mode called optical fiber mode is also named [52] but never referred in the reviewed studies.

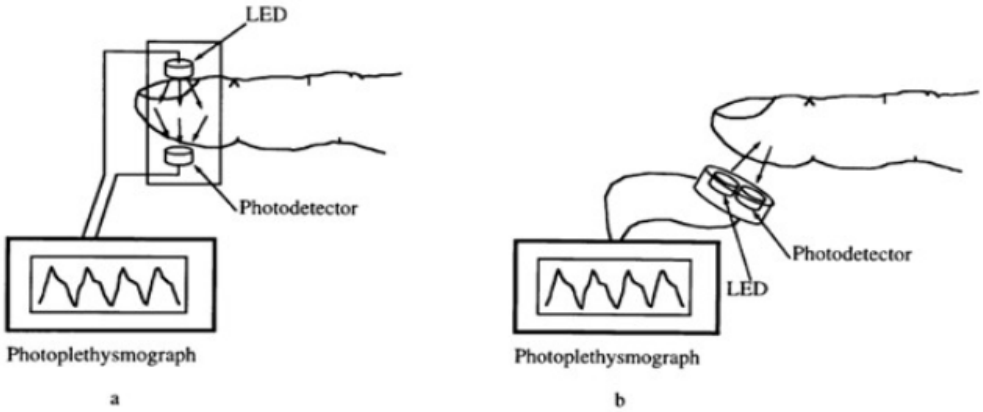



**Figure 2.3. Transmission and reflection mode for PPG signal measurements *[52]*.**

*Figure 2.4* shows the AC and DC component from PPG signal. AC component corresponds to variations in blood volume due to heartbeats and DC component refers to the steady light absorption of the tissue [47]. DC component may show minor changes with respiration, and the AC component is superimposed over the DC steady part [48].

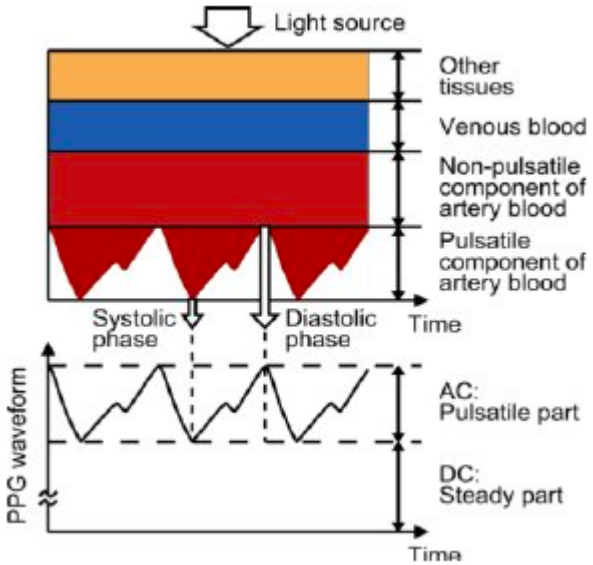



**Figure 2.4. AC and DC components of the PPG signal *[150]*.**

The voltage from PPG is proportional to the quantity of blood flowing through the blood vessels. Even small changes in blood volume can be detected using PPG [49]. Despite the apparent simplicity of the waveform, several body responses can be found. For example, heart rate, respiration, oxygen saturation, blood viscosity, arterial blood pressure or even posture changes [53].

It has been stated that this signal is very sensitive to artifacts and that many factors can affect the reproducibility of these measures, an important characteristic for PPG since it gives confidence in different measure environments. These factors include the method of probe attachment to tissue, probe-tissue interface pressure, pulse amplifier bandwidth, subject posture and relaxation, breathing, wakefulness, room temperature and acclimatization [45].

#### 2.1.4.2. Parts of the photoplethysmogram

*Figure 2.5* shows a typical PPG signal with its parts. In the horizontal axis, we have the time and in the vertical axis, we have the voltage of the signal. It should be annotated that the shown waveform in *Figure 2.4* is inverted on the vertical axis (in other words, it's backward). The one obtained below is the result of this inversion.

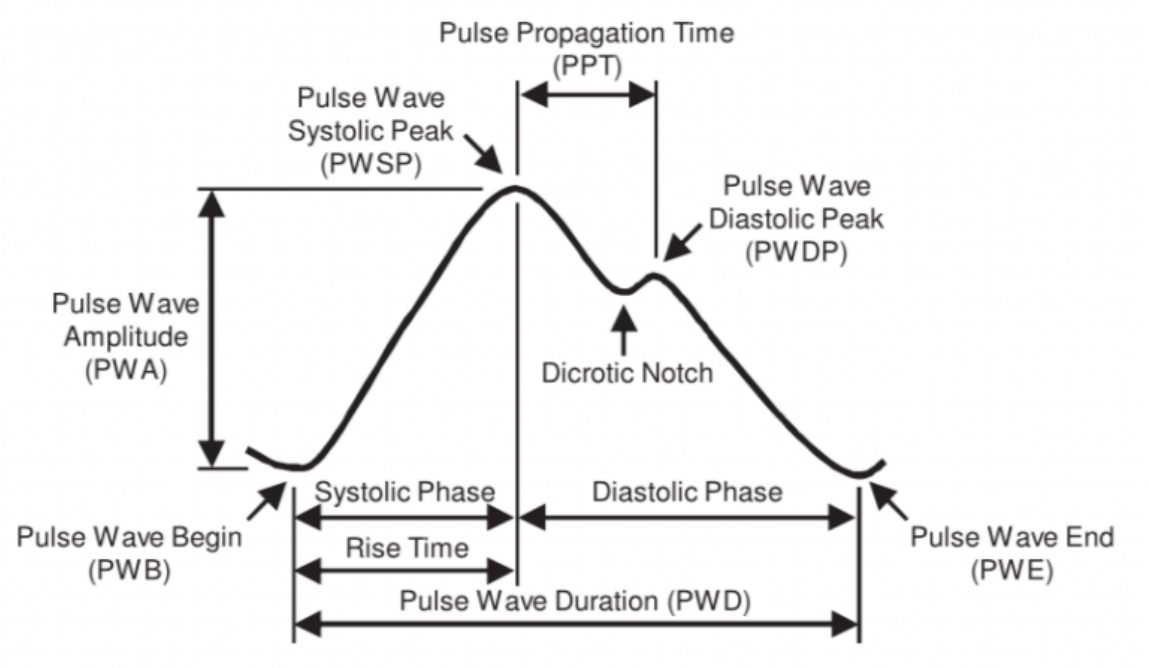



**Figure 2.5. Parts of a typical PPG waveform *[151]*.**

PPG waveform morphology has two main characteristics described by Hertzman and Spealman in [45]:

**Anacrotic (systolic) phase:** refers to the rising edge of the pulse, this phase is concerned with systole.

**Catacrotic (diastolic) phase:** refers to the falling edge of the pulse and is concerned with the diastole and wave reflections from the periphery. In this phase, a dicrotic notch can be sometimes identified. Dicrotic notch is related to vascular tone, according to [50]. A high arterial vascular tone is associated with the notch occurring early and high up on the downward diastolic curve. This notch is usually seen in subjects with healthy compliant arteries. The presence of this notch depends greatly on the person whose measures are being taken, and it's for this reason that sometimes this notch isn't noticed.

Peaks in PPG waveform are named different than in ECG signals. According to [9], peaks are named S, M, P, and Q.

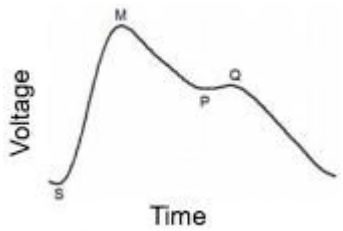



**Figure 2.6. Peaks of a typical PPG waveform *[9]*.**

Then, the correspondence of peaks is:

- **S peak:** corresponds to the pulse wave beginning.
- **M peak:** corresponds to the systolic peak.
- **P peak:** corresponds to the dicrotic notch location.
- **Q peak:** corresponds to the diastolic peak.

The S-M segment corresponds to the anacrotic phase and the segment after M peak until the end of the pulse wave corresponds to catacrotic phase. Other important segments or intervals in PPG waveform are described as follows:

**Systolic amplitude:** is the maximum volume level detected over the area that is being measured. *Figure 2.5* calls it as pulse wave amplitude. It can also be seen as the amplitude difference between S-M peaks in *Figure 2.6*. It corresponds to the systolic phase or anacrotic phase, where the blood is ejected [53].

**Diastolic amplitude:** is the amplitude difference between the Q peak and the end of the pulse wave. This is related to the diastolic or catacrotic phase where the ventricles fill with blood. After the M peak, blood volume starts to decrease. However, sometimes Q peak is not noticeable [53]. PPG waveform can also display venous pressure characteristics, which are often first detected by the presence of large peaks during diastole [51].

**Pulse width:** is the time elapsed since the signal level surpasses 50% of the M peak level until the signal level decreases lower than it. This value is related to the cardiovascular system resistance [53].

**Inter beat interval:** is the time elapsed between two successive M peaks. This value is correlated with the RR time interval taken from ECG signal. It can be measured between the S peaks too, because this interval value is usually identical to IBI. However, its use is only preferred when M peaks are too difficult to recognize [53]. It's also called inter pulse interval (IPI).

The part of the body where the measures are being taken is an important consideration. Sensitivity to changes in the sympathetic system is greater with measures from the fingertip, compared to other areas such as the earlobe [47]
The PPG blood volume pulse has similar changes with the blood pressure pulse occurring in vascular diseases, such as damping or loss of pulsatility [46].

### 2.1.5. Cardiovascular risk and cardiovascular diseases

Cardiovascular diseases can refer to a number of conditions that affect primarily the heart arteries and blood vessels and therefore, the rest of the body. The main organs affected by cardiovascular diseases are the brain, the kidneys, the eyes and lower limbs [54] [55]. There

are more than 70 types of cardiovascular diseases [56], but the main and the more important are the myocardial infarction and cerebrovascular accident [54].
Cardiovascular risk can be defined as the probability that a subject has to suffer a cardiovascular disease during a certain period of time. This depends mainly on a number of risk factors presented by the subject [54]. A risk factor is a characteristic or condition of a subject or population that is present early in life and it's associated with a greater risk of developing a future illness [57]. It can be a behavior or a habit like smoking or sedentarism; a hereditary trait or family history; or a paraclinical variable like high cholesterol level [59]. They are divided into non-modifiable risk factors such as sex, age, race or family background; and modifiable factors like arterial hypertension, increase in cholesterol level, overweight and obesity. A detailed list of the main risk factors can be found in [58] and [60].

#### 2.1.5.1. Arrhythmias

Arrhythmia refers to any change from the normal sequence of the heart electrical impulses. These impulses may happen too fast, too slow or erratically, causing the heart to beat irregularly. If the heart doesn't pump blood effectively, other main organs may be severely compromised [1].

##### 2.1.5.1.1. Bradycardia

Bradycardia is a type of arrhythmia characterized by a slower heart rate than normal. "Too slow" or "slower" is relative for each subject and depends on the age and physical condition. Usually, older people are more prone to suffer from bradycardia, and for adults, a resting heart rate of fewer than 60 beats per minute qualifies as bradycardia [33].
There are some exceptions when the subject is at deep sleep, heart rate may fall below 60 beats per minute, as it also happens for physically active adults (and athletes), who have a resting heart rate slower than 60 beats per minute [33].

##### 2.1.5.1.2. Tachycardia

Tachycardia is another type of arrhythmia characterized by a faster heart rate than normal. Once again, how "too fast" is defined may depend on the subject's age and physical condition. For adults, a heart rate higher than 100 beats per minute is considered "too fast" [61].
There are various types of tachycardia, namely: atrial or supraventricular tachycardia, sinus tachycardia and ventricular tachycardia [61].

#### 2.1.5.2. Prevention of cardiovascular diseases

According to the World Health Organization [59], there are many behavior changes regarding lifestyle that could help to lower the risk of cardiovascular diseases. Reduction of tobacco intake, body weight, blood pressure, blood glucose and blood cholesterol all have a beneficial impact on major biological cardiovascular risk factors. Some of the most important reductions

or modifications are over tobacco, diet, alcohol, physical activity and therefore body weight. There are also some psychosocial factors that might affect the subject who suffers from cardiovascular diseases. Depression, anxiety, lack of social support, social isolation and stressful conditions at work are some of them [59].
Prevention of cardiovascular diseases is focused on the reduction of these factors from the beginning. It's much better to avoid these detrimental conducts, rather than drug treatments, which are usually more expensive to the health care system and might end up in adverse effects [59].

## 2.2. REVIEW OF MOVEMENT ARTIFACTS IN DIFFERENT MEASUREMENT SCENARIOS

### 2.2.1. Review methodology

#### 2.2.1.1. Databases

This state of art research was made in order to identify the scenarios and the noise artifacts for PPG signal-based applications, that use mainly as a reference the ECG signals. The above was made taking into account that these applications are made under different activities that induce movement artifacts. Several techniques of pre-processing and filtering are also analyzed and exposed.
State of the art in this occasion was outlined by 13 different articles having in account databases from EBSCO, Science Direct, Google Scholar and Scopus.

#### 2.2.1.2. Selection criteria

The following search strings were used to provide the information compiled in *Table 2.1.*

(holter) AND ((tachycardia OR arrhythmia OR AFM OR atrial) AND fibrillation) AND (ppg OR photoplethysmographic OR photoplethysmography) AND (movement OR artifacts OR noise)

(holter) AND (tachycardia OR arrhythmia OR atrial OR Fibrillation) AND (ppg OR Photoplethysmography OR photoplethysmographic) AND (movement OR artifacts OR noise)

#### 2.2.1.3. Data analysis

Inclusion and exclusion criteria were applied to the results of these strings. The main subject was the use of methods for decreasing artifacts as well as the distinction of the scenario where the measures were taken. It was possible to identify 22 articles through this process. These last articles were evaluated in a more rigorous way. The factors chosen were the arrhythmia desired for analysis, the population used for each one or the database type that was used (if it was created by themselves or it was taken from another site). Consequently, we declare the 12 articles found in *Table 2.1*., and the gaps corresponding to each one of them.

| ARTICLE NAME | DEVICES USED FOR MEASUREMENT | ARTIFACTS | REFERENCE SIGNAL/ GOLD STANDARD | PATIENTS | PERFORMANCE PARAMETERS | GAPS | YEAR |
|---|---|---|---|---|---|---|---|
| Automated Atrial Fibrillation Detection Algorithm using SmartWatch Technology [22]. | -AliveCor ECG. -Apple Watch. | None. | ECG-Holter. | 100 patients from 57-79 years. 17% of the subjects were women. -Measurements are made in the waiting room at rest. **They use their own dataset** | -Sensitivity: 93% -Specificity: 84% -K coefficient: 0.77 | - The algorithm isn't applicable to PPG signals, but to ECG signals. - PPG signals aren't used in the study. Low specificity results are shown. - It doesn't make any focus on noise remotion or artifact identification. | 2018. |
| Classification of the quality of WristBand-Based Photoplethysmography signals [23]. | **Wearable:** Empatica E4 wristband **Holter:** General Electric Seer Light Extend Holter. This article focuses on eliminating movement artifacts | Movement artifacts in daily life activities. | PPG without noise. | 15 healthy participants for 24 hours with wearable. **They use their own dataset.** | -PPG sensitivity: 98.3% -PPG specificity: 98.3% | - Data retrieved from Holter isn't used in this study to generate a comparison. - It isn't made processing on PPG signal, they take windows of short duration where artifacts don't exist. - They don't consider any type of arrhythmias. | 2017. |
| Detecting Episodes of Brady- and Tachycardia Using Photoplethysmography at the Wrist in Free-living Conditions [19]. | -**Wrist wearable PPG:** CM3 Generation 3 Philips (Optical PPG sensor and accelerometer sensors). -**Holter:** (H12+, Mortara). | Movement artifacts present in free-living conditions. | ECG-Holter. | 20 healthy subjects, where Male=55%, between 52-80 years. BMI= 28.1 ∓20% - ECG and PPG recordings for 24 hours. **They use their own dataset.** | -Tachycardia prediction: (Sensitivity: 89%, Specificity: 99%) -Bradycardia prediction: (Sensitivity: 85%, Specificity: 99%) -Detection on normal sinus rhythm (Sensitivity: 99%, Specificity: 85%). | - Scenarios that conform to daily activities aren't specified. It's desired a better specification of what could generate certain artifacts. - Low sensitivity on tachycardia detection due to low intensity movements. - Low sensitivity on bradycardia prediction due to premature beats at rest. **- They use their own dataset and it isn't public.** | 2017. |
| Validating Features for Atrial Fibrillation Detection from Photoplethysmogram under Hospital and Free-living Conditions [24]. | -**Wrist wearable PPG:** CM3 Generation-3, Wearable Sensing Technologies, Philips, Eindhoven). -**Holter:** Actiwave Cardio (CamNtech Ltd., Cambridge, United Kingdom). | For testing dataset: no motion artifacts. For testing dataset: movement artifacts under free-living conditions. | ECG-Holter. | PPG measurements for 18 patients before and after cardioversion, and for 16 patients (4 with AF) during 24 hours. ECG and Holter measures were also taken as reference. **They use their own dataset.** | -Sensitivity: at hospital 92.3%; under daily life conditions 71.6% -Specificity: at hospital 60.7%, under daily life conditions 84.9% -Accuracy: at hospital 78% | - The movements made by each patient during measurements are not identified. | 2017. |
| Gyroscope vs. accelerometer measurements of motion from wrist PPG during physical exercise [25]. | Shimmer 3 Sensor. | Movement artifacts: in the wrist. Activities: Riding bicycle, treadmill. | ECG-Holter. | 8 subjects. **They use their own dataset.** | - Performance parameters are not mentioned. | - Arrhythmia prediction isn't made, but it's developed a heart rate estimation algorithm. - Performance parameters are not mentioned. They just focus on proving that gyroscopes are more | 2016. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | effective than accelerometers, by reducing noise with each one of them.<br>- It doesn't focus on arrhythmias either. | |
| Fuzzy Entropy based Detection of Tachycardia and Estimation of Pulse Rate through Fingertip Photoplethysmography [26]. | - They make their own device for the cardiac pulse measurement in the fingertip. | Power line interference, baseline wandering. | PPG | -15 healthy people.<br>- Also analyzed the technique using the Capnobase database.<br>**They use their own dataset.** | - Accuracy for taken measurements: 0.9949.<br>- Accuracy for the database (noise + no noise): 0.993. | - A wearable use isn't made, the measurements are taken with a pulsioximeter.<br>- PPG isn't contaminated with movement artifacts. They consider another type of noise. | 2015 |
| Optimized PPG-based wearable acquisition unit for massive analysis of heart rhythms [20]. | -MEMs sensors. | Movement artifacts, the capture of various movement patterns. | PPG | -Tested by 4 persons for 24 hours.<br>**They use their own dataset.** | -Maximum accuracy at night hours while they sleep, with 97.5% of valid data.<br>- Specificity or sensitivity are not determined. | - Arrhythmia prediction isn't made.<br>- Movements are not specified.<br>- A commercial wearable isn't used. They propose a new one with HW and SW modifications. | 2017 |
| Photoplethysmography -Based Method for Automatic Detection of Premature Ventricular Contractions [21]. | - Not used. | -None. | PPG | - MIMIC and MIMIC II databases. | The obtained sensitivity and specificity values for both considered PVC types were 92.4 / 99.9% and 93.2 / 99.9%. | - Detection of the artifacts is not very wise, since the used databases (MIMIC and MIMIC II) are measured from patients in ICU. For this reason, movement artifacts couldn't be appreciated. | 2015 |
| Heart rate estimation using wrist-acquired photoplethysmography under different types of daily life motion artifact [29]. | Pulse Sensor. | Movement artifacts. | PPG | - From 16 subjects, they choose 5 for information retrieval.<br>**- They use their own dataset.** | -Errors located within the range of 95%, between -7.6 and 6.6 bpm. | **-** Performance parameter analysis isn't made. However, they mention error ranges in bpm (beats per minute).<br>- Arrhythmia detection isn't made. | - |
| Description of a Database Containing Wrist PPG Signals Recorded during Physical Exercise with Both Accelerometer and Gyroscope Measures of Motion [30]. | Shimmer 3 GSR+ unit. | Movement artifacts on 4 different kinds of activities: walking, running, riding a bicycle with low resistance and bicycle with high resistance. | ECG | - 26 subjects. | - Not specified. | **-** Arrhythmia detection isn't made. | 2016 |
| A Robust Dynamic Heart-Rate Detection Algorithm Framework During Intense Physical Activities Using Photoplethysmographic Signals [17]. | Wrist wearable that implements 2 pulsioximeters in the wrist and 1 accelerometer. ECG signal is measured with sensors connected to the chest. | Movement artifacts:<br>-Training dataset: low and medium intensity activities (resting with wrist movements- walking- running)<br>-Testing dataset: high intensity activities such as (boxing, intense arm movements, running). | ECG | -23 datasets: 12 for training and 11 for testing. (2015 IEEE Signal Processing Cup Competition). | - Parameters aren't mentioned, although it compares with previously made algorithms regarding the absolute error and the mean absolute error. | -The developed algorithm allows to eliminate artifacts during high intensity exercise, but it doesn't focus on arrhythmia detection.<br>- Wearables are not used, but a pulsioximeter. | 2017 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Motion and Noise Artifact Detection and Vital Signal Reconstruction in ECG/PPG based Wearable Devices [31] | - It depends on the study | Movement artifacts depending on the activity: (walking, running, boxing, arbitrary movement) | ECG | -23 datasets (2015 IEEE Signal Processing Cup)<br>-10 datasets from Chon Lab. | -It depends on the study.<br>- Parameters are not mentioned, although it compares with other algorithms (TROIKA, JOSS) | - The algorithm allows to depurate PPG signal with movement artifacts and posterior heart rate estimation.<br>- It doesn't make arrhythmia detection. | 2015 |

**Table 2.1. Summary of analyzed works identifying scenarios where motion artifacts are induced.**

### 2.2.2. Results and analysis of the gaps

The main gaps found in the previously mentioned works are summarized below:

A. The majority of the studies oriented towards noise detection caused by movement artifacts are solely focused on the heart rate estimation [17][23][29][31], the data acquisition from PPG corrupted signals [20][30] or just the comparison of other accompanying signals besides from PPG [25]. From those which make arrhythmia detection, there are two works [26][19] that analyze possible bradycardia or tachycardia and also determines performance parameters, obtaining sensitivity and specificity measures. However, the first [26] is made with their own created device and don't consider movement artifacts but powerline interference. The second [19] obtains lower sensitivity results than [26] due to low intensity movement artifacts. The author of [19] states the increment of false positives in tachycardia cases along with the missed bradycardia events.
B. Better sensitivity results in [19] are expected by having a greater rigor at the time of the movement characterization made by the subject. In this case, the low movement artifacts stated as the limitation in this study, could be annotated in the used dataset. This with the purpose of eliminating this kind of artifacts more efficiently.
C. None of the works shown in the table seeks to characterize a specific arrhythmia using different attributes or additional to the PPG wave nature, the mathematical treatment of the PPG signal, sex or age of the subjects. Neither has been made a correlation of the arrhythmia detection taking into account a population that suffers from a secondary illness or common within this population, such as diabetes, hypertension, metabolic disorders, etc.

### 2.2.3. Conclusions

A state of the art review was made for the specific conditions mentioned above in *subsection 2.2.1.3.* As a result, three gaps could be identified and summarized in the research question stated in *chapter 1.* This work intends to improve sensitivity as a performance parameter and then, avoid the false arrhythmia detection. The first focus is the characterization of low intensity movements, so the clean signal can be distinguished from the one generated because of low intensity movement artifacts. Then, the second focus will be centered on the improvement of the detection of true arrhythmia events under the characterized low movement artifacts.

As a conclusion, these two gaps complement each other and give as a result one problem. This can be resolved using signal processing and the appropriate classification techniques over PPG signals.

# CHAPTER 3. APPROACH TO A MOVEMENT ARTIFACT NOISE CHARACTERIZATION.

## 3.1. INTRODUCTION

As it was stated in *Chapter 2,* some of the developed arrhythmia detection algorithms don't consider motion artifacts [21][26]. Others make it, but the collected datasets by them are not publicly available, because they were acquired from patients who gave an informed consent [22][24]. Given that the focused arrhythmias are tachycardia and bradycardia, it is sought to obtain a dataset that contains both arrhythmias and movement artifact noise. The closest approximation to a realistic arrhythmia scenario with movement artifact noise is given by Bonomi [19]. The authors try to detect episodes of bradycardia and tachycardia on PPG signals under free-living conditions. However, their dataset is protected under informed consent, which makes it unavailable to the public. This clearly evidences a limitation for the arrhythmia detection with PPG signals, leaving as the only option the use of datasets retrieved from Physionet or another kind of repositories.

Although, these repositories containing arrhythmias do not reflect the real conditions of the movement freedom in wearable devices. For example, when a person is moving, walking or moving the most common measurement area, such as wrists or fingers. Solosenko states this limitation in his paper, using this challenging situation for adjusting the modeling and simulation of the PPG signal. He achieves this PPG model from an ECG signal sequence, taking as the main feature, the R-R intervals [12].

Bonomi uses a threshold coming from the accelerometer to discard the portion of the PPG that contains a considerable amount of movement and therefore, movement artifacts. Bonomi et. al. state that this procedure could lead to the loss of valuable information for the posterior arrhythmia detection. Even so, they expect to have a better classification, because they don’t use any preprocessing method over the PPG signal (which they presume was motion artifact free) [19]. An advantage was stated by this signal cropping process, exposing that this helped to embetter sensitivity in the tachycardia detection. Nevertheless, a pitfall was also outlined by expressing that even the low intensity wrist movements distorted the PPG signal, increasing the HR estimation and therefore leading to a wrong diagnosis (increased tachycardia estimations and missed bradycardia episodes). We suspect this could be occurring because of a scarce identification of the activities carried out by the patients. Because of the optical nature of the signal and the peripheral areas of the body where this is acquired, it is more likely to be affected or distorted by several kinds of noises. Since the dataset used by Bonomi is unavailable, we hope to recreate this scenario. First, characterizing the motion artifacts from another database found. This database contains noise due to low and medium intensity activities, well defined in time intervals. Zhang’s webpage [74] provides this database, which has been used for the IEEE Signal Processing Cup (IEEE SPC) and within other studies to assess various algorithms for movement artifact reduction [17,31,74].

The first procedure to simulate a scenario like Bonomi's (and then try to improve the sensitivity over tachycardia and bradycardia detection) is the creation of a simulated PPG signals dataset

with annotations of noise artifacts due to low-intensity movements. In this way, an approximation to this noise (caused by specific low-intensity activities) is developed. Then, it will be attached to a new dataset that includes the chosen arrhythmias for investigation.
This chapter exposes first a brief theoretical framework, where the noise types that affect PPG signal are specified. In the second part, a review of some algorithms or techniques for movement artifacts noise modeling is made. In the third part, the proposed methods for the characterization approaches are explained. Tests are done over these noise characterizations, computing their performance parameters for a classification task. Lastly, the structure of the chosen noise summed to the new dataset is shown.

## 3.2. THEORETICAL FRAMEWORK: NOISE IN THE PPG SIGNAL

### 3.2.1. Introduction

Any measure of phenomena consists of two parts. One is the interest data or the valuable content of the signal, which allows making further analysis. The other part is the noise, which usually seems as random information that is overlying over the interest data [62].
It is impossible to find noise-free data in real-life experiments because of different events that generate noise during measurement. These events can be either thermodynamic, electric, or quantum. Not to mention the experimentally induced errors that are interpreted as noise in the signal as well [62].
PPG signal is also affected by noise because the measurements are taken from peripheral body parts which generate greater movement quantity, such as the wrist or fingers. The conditions that avoid a correct understanding of the signal can be denominated as artifacts.
It is possible to classify these artifacts in different types [15]. The detection, characterization, and subsequent removal of these artifacts make the biomedical signal understandable and manageable for any diagnostic that is desired.

### 3.2.2. Types of artifacts

According to [15], three types of artifacts can be identified:

- **Environmental artifacts**: these artifacts are caused by environmental conditions, such as electromagnetic interference, light interference, thermal noise, or shot noise. In the specific case of the PPG signal, because it is an electro-optical sensing method, it is primarily affected by light interference [15]. For example, it is not the same to make measurements under the sun's light, in the night or under lamplight. However, wearable devices which make this type of measurements, generally suppress this kind of noises during measurement. That is the reason why there is no big focus on these types of artifacts.

- **Experimental error artifacts**: these artifacts are not manageable during the data acquisition and are caused because of the patient's movement. The movement generates a friction effect between the skin and the sensor, translating it into noise [15].

The removal of these artifacts is difficult, since they cannot be found in a narrow frequency band. Oppositely, their frequency bands often overlap with the ones of the signal. A greater focus is made on this type of artifacts. It has been encountered in the literature, that they are the focus objective for removal using machine learning techniques or filtering algorithms.

- **Physiological artifacts:** these artifacts are produced by variables of the human body itself, such as eye movement, little signals caused by muscle tensions or brain signals. This kind of artifacts generate a noise which can be considered as unnoticeable [15], and for this reason, it is not made a high focus on them.

#### 3.2.3. Movement artifact noise

Motion artifacts are the focus of most of the studies and papers, according to the review later in *Chapter 4*. These are the most harmful for the signals and the most influent when the information from the signal is extracted.
These movement artifacts are especially hard to remove because of the overlapping of their bandwidth with the signal's valuable information. *Table 4.1* exposes some of the proposed methods, such as adaptive filtering or filter banks, wavelet decomposition, smoothing filters, or enhanced versions of already developed algorithms.
The PPG signal concentrates its information widely on the 1-7 Hz frequency band, but we could even consider that the real concentration of the information is between 1-4 Hz (a very narrow band). The above can be seen from the frequency analysis in *Appendix A*. The information signal overlaps with the noise frequency component inside this interval, making difficult the noise extraction [23]. The main consequence of these so mentioned artifacts over the signal is the baseline drift, introduced principally because of the skin friction with the sensor. The movement of the extremities causes this low-frequency noise. Another effect of these artifacts is the high-frequency noise. It can be witnessed as very slight and frequent peaks, sometimes imperceptible, which follows the curve of the signal.
With this in mind, we tried to recreate these artifacts, dividing them into the frequency components mentioned above (high frequency and low frequency). State of the art will give us an idea of already used techniques for this noise characterization and modeling.

## 3.3. STATE OF ART: REVIEW OF ALGORITHMS FOR PPG SIGNAL NOISE CHARACTERIZATION

#### 3.3.1. Review methodology
#### 3.3.1.1 Databases

Three different articles outlined the state of the art for noise modeling, extracted from searches in the following databases: EBSCO, Google Scholar, and Scopus. This search was made in order to identify the scenarios and the movement artifact noise induced on PPG signals

datasets. Also, we expect to find some noise characterization methods or techniques that allow modeling movement artifact noise either using mathematical models or statistical approaches.

#### 3.3.1.2. Selection criteria

The following terminology is used in the search strings to find articles that could contain techniques for artifact extraction and noise modeling over PPG signals:

(artifacts OR noise OR movement) AND (characterization OR model OR modeling) AND (ppg OR photoplethysmography)

Applying exclusion criteria to this search strings, where the central theme was the extraction or characterization of the noise caused by movement artifacts on the PPG signal; it was possible to identify the articles [63, 64, 12]

*Table 3.1.* Shows an identification of the techniques and gaps found in the articles above that were useful to our approximation.

| Article | Techniques used that could be valuable for our work | Main gaps encountered | Study date |
|---|---|---|---|
| Noise cleaning and gaussian modeling of smartphone photoplethysmogram to improve blood pressure estimation [63] | In this article, a way of modeling the PPG signal starting from a Gaussian model is proposed. Here, the author refers to the PPG as the sum of two Gaussian functions, arguing that a mathematical model can ensure a better signal realization.<br>This technique gives us the idea of modeling a clean PPG signal, thus subtracting this clean signal from the contaminated signal of the dataset, and therefore, obtaining a form of motion artifact. | The central gap that was found in the application of this technique as an artifact collecting idea was that due to the uncertain form of the PPG pulses, a mathematical model does not give a particular idea for all of the possible types of PPG waveform.<br>In other words, one of the PPG cycles could differ from another in terms of its morphology. Amplitude, width, peak height, or peak distance are considered as morphology features.<br>If only one general model is created, it would not fit all the possible morphologies for all PPG cycles. Instead of cleaning the signal, this could distort it. | 2015 |
| Modeling of Motion Artifacts in Contactless Heart Rate Measurements [64] | In this article, it has implemented a statistical approach over the signal. The author argues that the ECG signal could be understood as the sum of valuable data, Gaussian noise, and artifact noise (both superimposed over the signal). Employing histogram analysis, it has established the most probable distribution for this movement artifacts and then a spectral density analysis. This last analysis gives a | This idea has been considered as a good one in terms of practical use, given that it is much easier to create noise from a probabilistic model, that also generalizes the behavior of the artifacts.<br>However, the creation of artifacts has been analyzed for ECG signals instead of PPG signals. An in-depth analysis of this paper allows us to see that the obtention of the probabilistic distribution of the artifacts basing on amplitude histograms could be incomplete. Even when the simulated | 2013 |

| | | | |
|---|---|---|---|
| | piece of information that allows "creating" a probability mass function of the named artifacts. | results are similar to the real artifacts of the capacitive ECG signal, this idea could not work for small sets of data as ours. Additionally, the author mentions that the extracted intervals are not perfectly fitted to real data because they may also contain QRS complexes and other not modeled noise sources which might influence the spectrum accordingly. | |
| Modeling of the Photoplethysmogram During Atrial Fibrillation [12] | In this article, it has noticed the need for a model of the PPG signal for further arrhythmia diagnosis, given the lack of free online datasets that contain this type of information.<br>Two kinds of information are given here: first, a model of the PPG signal based on two Gaussian and one log-normal waveform; second, it has mentioned a generation of motion artifacts employing filtering white noise according to spectral properties of these artifacts. | A similar case to the first article is seen in the first suggestion; the gap can be justified in the same way and, moreover, the mathematical expressions and procedures for the obtention of the model are complicated and require a more in-depth analysis on other parameters and expressions.<br>About the other suggestion, it could be taken until this point, but the spectral properties were not specified, thus making it impossible to recreate. | 2016 |

**Table 3.1. Summary of analyzed works for noise characterization in specific PPG signals.**

The literature review allows to conclude that very few studies have been focused towards noise recognition and posterior characterization in PPG signals. Modeling of the noise over these signals would allow to train and validate algorithms for noise reduction, which is an important part of their development process. Additionally, it should be considered that the real-data acquiring is a time-consuming task and generally involves the signing of an informed consent, for ethical reasons.

Another search string was made, without being limited by the specific signal we are working on. We considered that these other signals, such as EEG or ECG, must have a resemblance to PPG signals in terms of quasi-periodicity, correlation, statistical, and temporal features. Thus, we expect to obtain articles that could contain some technique of movement artifact characterization over another kind of signals.

The primary purpose of this new search is to analyze and compare techniques that would allow us to extract the main noise features from contaminated PPG-resembled signals.

Taking into account the beforehand mentioned, we limited the search as follows:

(artifacts OR noise OR "movement artifacts" OR "motion artifacts" OR "motion noise" OR "Gaussian additive noise") AND (characterization OR model OR modeling) AND (ECG OR EEG OR "periodic signal" OR "non-stationary")

The resulting range was limited by year of publication, subject area, and document type. The results are contained in *Table 3.2.* were eight works were found, shown as follows:

| Article | Techniques used that could be valuable for our work | Main gaps encountered | Year |
|---|---|---|---|
| Motion artifacts reduction from PPG using cyclic moving average filter [65] | A cycling moving average method is proposed, which averages the values of each sample by separating the cycle of the PPG signal. If there are some motion artifacts in continuous PPG signal, disjoin the signal based on a cycle. Moreover, then, these signals are fitted to have the same cycle by coordinating the number of samples. | An overlapping problem is presented. Also, it filters only the high-frequency component, leaving the low-frequency component unchanged. Therefore, the efficiency of the filter is limited by the noise amplitude and the order of the moving average filter. Finally, this article gives an idea of how to filter a signal by smoothing it, but also suggest a way of obtaining the high-frequency noise component through MAF. | 2014 |
| Improved EEG Segmentation Using Non-linear Volterra Model in Bayesian Method [66] | In this paper, the cascade of linear predictive coding (LPC) and non-linear Volterra filter is employed for modeling of noise in EEG signal, and this methodology is applied to the procedure of change-point detection. LPC filter is initially used as an estimator of the noise (correlated and uncorrelated noise). Then, a Volterra filter is used to find the non-linear relation between the noisy signal and noise. In this way, they extract the existing noise in the signal. | The technique for an LPC filtering is taken because this can help us in the same estimation of noise. Regarding Volterra nonlinear filter, it is argued by the author that this technique could be computationally expensive. The more the Volterra coefficients are, the longer is the computation time. | 2018 |
| Single channel EEG artifact identification using two-dimensional multi-resolution analysis [67] | A multi-resolution time-frequency analysis is proposed, using techniques such as 2-dimensional wavelets and curvelets, then, extracting statistical features from the signal — this entire process to classify several activities which induce different motion artifacts during EEG measurement. | It is known that wavelets allow a better signal analysis, giving spectral information as well as energy concentration over time. The author argues that spectral analysis by itself may help, but they do not provide temporal information about a specific event.<br>It is for this reason that it is considered the wavelet as an information extraction technique. Also, the recommendation for using not only frequency features but time features is taken. | 2017 |
| Detection and separation of EEG artifacts using wavelet transform [68] | Once again, the wavelet transform is used as a technique for denoising EEG signal. Thus, the author | A similar approach as [67] is made. Even though the wavelet importance is recognized, this study does not provide a noise characterization | 2018 |

| | | | |
|---|---|---|---|
| | exposes wavelets as a decomposition of the signal. Results are evaluated in terms of time, power, and number of cells utilization for DWT denoising. | technique, but a noise removal one. In addition to this, the study seemed to have more focus on computational requirements for processing than noise removal itself. | |
| Real-Time EEG Signal Enhancement Using Canonical Correlation Analysis and Gaussian Mixture Clustering [69] | A real-time removal artifact algorithm in EEG signal is proposed. Based on three main techniques, Canonical Correlation Analysis (CCA), the feature extraction, and Gaussian Mixture Model (GMM).<br>CCA method exploits the fact that the autocorrelation of muscle activity is weaker than that of brain activity. This approach solves the problem by forcing the sources to be maximally auto-correlated and mutually uncorrelated.<br>GMM is used for unsupervised learning, assuming that all data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. | The same case as [68] is encountered, given the fact that it is a noise removal algorithm, but not a noise characterization one. Despite this, it is observed that the GMM exposed may be an applicable approximation, analog to PPG signals.<br>Feature extraction cannot be afforded because of the features difference between the signals. Also, because it requires labeled data (it is a supervised learning method), which we do not own.<br>However, the CCA method also exposes a good approximation jointly with GMM. | 2018 |
| Two-stage wavelet shrinkage and EEG-EOG signal contamination model to realize quantitative validations for the artifact removal from multi-resource bio signals [70] | An EEG-EOG signal contamination model is designed to remove artifacts from EEG using a two-stage shrinkage wavelet method with the decomposition of the undecimated wavelet transform (UDWT) method. For this, a semi-artificial EEG-EOG contamination dataset is created using real data and white noise, in order to model and to quantitatively validate artifact removal from EEG and thus, enabling to compare this study with others, since it is tested numerically and systematically by reconstructing the EEG signal. | The proposed reconstruction method was validated using a numerical method, which demonstrated that the first stage pursued abrupt changes with high amplitudes provided by assumed EOGs, and the second stage provided the EEG frequency spectrum as observed in the original signal. This result explains why when there are higher amplitudes; the conventional shrinkage parameters must be personalized to real values, in order to obtain the correct outcomes. | 2019 |

| | | | |
|---|---|---|---|
| Noise estimation in the electroencephalogram signal by using Volterra series coefficients [71] | Here, the author employed the Volterra model to find the nonlinearity relation between electroencephalogram (EEG) signal and the noise. It is exposed as a novel approach to estimate noise in the EEG signal.<br>Although it is found as a complex method computationally, the author proposes a new approximation that could help to reduce this. | An important issue in implementing the Volterra model is its computation complexity, especially when the degree of nonlinearity is increased. The author suggests that in many applications, it is urgent to reduce the complexity of computation. | 2015 |

**Table 3.2. Summary of analyzed works of noise characterization on PPG-resembled signals.**

#### 3.3.1.3. Summary of the state of the art analysis

It can be seen from *Table 3.1* that [63] proposes a PPG mathematical model for each S-S peak interval. The model gives an idea for noise extraction by creating clean models of the PPG signal and later join them, hoping that the subtraction of the clean signal over the noisy signal could give an estimation of the noise. However, since a constant morphology is assumed for all the intervals, this PPG signal model might not meet the entirety of possible morphologies that a noisy PPG signal has in each S-S interval. Therefore, the extracted noise by subtraction would not resemble the real motion artifact signal, and might even contain other information.
In [12], Solosenko models the PPG signal from the R-R interval series of ECG. He considers five possible templates for PPG signal, but at the same time states that this morphology varies considerably depending on age and medical condition. PPG signal modeled for this study is intended to model different arrhythmia cases (atrial fibrillation, atrial premature beats, and ventricular bigeminy), which is not the case for us, since what is needed is a clean signal model to subtract it from the noisy signal. Lastly, the mathematical modeling of this paper goes beyond the best of our knowledge, making it difficult to understand and exceeding the time scope for this project too. Even though it has used a set of five PPG template pulses, this selection biasing might not represent the totality of the possible noisy PPG waveforms. Besides, the ones in IEEE dataset vary significantly given that the heart rate increases and decreases as time passes. Because of this limitation, it cannot be used for noise extraction in this occasion. Although, one aspect to be rescued from this study is the use of a bandpass filtering method of white noise to simulate noise in these PPG signals. The author uses the spectral characteristics of the PPG signal. However, these are not mentioned, and therefore, a spectral analysis should be made in *Appendix A* to use this approach as a tool.
Wartzek proposes a good approximation through amplitude histograms [64]. This approach can be made when the number of realizations measured is much bigger than the ones we own. Also, it can be seen that the modeled artifacts belong to a capacitive ECG signal, which differs from our PPG signal. He suggests that even considering a probabilistic distribution, the

modeling of the artifacts is not perfect and will not be because it could contain many other sources of noise.

The second review suggests more techniques regarding motion artifact reduction or characterization, which means there might be a chance to adopt these techniques focused on PPG-resembled signals for making them useful in our approach.
Some of the found algorithms are more suitable for noise removal following the classification of another kind of phenomenon, as shown in [68,69]. Supervised learning methods or feature extraction methods exposed in [69] are majorly used in the context of classifying "artifact affected" or "non-affected" signal. In our case, we already know the signal is affected.
Volterra methods [66,71] are considered as a good suggestion. However, it has been recommended by the authors to consider the computational complexity they involve.
LPC exposed in [66] can be seen as an alternative for noise modeling, under the supposition that a linear predictive model can be implemented if it is fed with the high-frequency noise.
EEG-EOG signal contamination model [70] is proposed by adding white Gaussian noise with real patient's information data. This exposes gaussian noise as a common method for inducing noise over biomedical signals.
It can be seen in [72] that the PMAF method stacks quasi-periodic segments of the PPG signal, containing peak information within the segment. This analysis suggests a way for acquiring a more general high-frequency noise component that simulates having much more realizations than twelve. This procedure can be performed using moving average filtering over the signal, generating a higher correlation between adjacent samples.
These approaches give an idea on how to extract the high-frequency noise components from the overall motion artifacts, starting from a subtraction procedure and then manipulating the high-frequency noise signal as we consider.
[73] states imperatively to carry data analysis algorithmically rather than by visual analysis. Here, some high and low pass filter techniques are remarked, giving us an idea of dividing noise from its frequency components. If noise cannot be seen concentrated in a unique frequency band, it can be divided into low and high-frequency noise.

### 3.3.2. Motion artifacts modeling

Most studies centered on noise detection by movement artifacts are focused only on the estimation of the heart rate or the analysis of the quality of the PPG signal before and after the denoising process. Of the few studies focused on arrhythmia detection, only one [19] analyzes possible bradycardia or tachycardia events and also determines the performance parameters, obtaining the sensitivity and specificity. Sensitivity measure from this work is said to be affected by the presence of low-intensity movements. For this reason, tachycardia and bradycardia events must be studied under specific activities to describe the motion artifacts that are distorting the signal, focusing on low-intensity movements.
However, none of the works from *Table 2.1* is looking to characterize the specific arrhythmias we are working with under specific movement artifacts. The purpose of this chapter is to provide a simulated dataset, which allows recreating the specific movement artifact scenario for the

detection of bradycardias and tachycardias. The specific movement artifacts considered for the Bonomi's simulation scenario include hand, arms or wrists movements (low-intensity movements).
In order to reach these simulated conditions, noise from specific time-located activities and movements must be characterized. In this case, the external database which will be used for movement artifact extraction is the IEEE 2015 Signal Processing Cup (IEEE 2015 SPC) database [74]. Later, we expose five different approaches for noise modeling, capable of describing movement artifacts under the six different types of activities. The focus is made on the low-intensity activities, were two of them are within the six activities in the description of the IEEE SPC database.

### 3.3.3. Challenges and Contributions

The biggest challenge is to prove that our artificial noise is correctly characterizing the movement artifacts present on the six activities. To do so, we made tests under the supposition that after the subtraction of the modeled noise signal from the original PPG signal (or the noisy signal), the peak detection should improve. Then, the accurate extraction of the signal's peaks is needed. This will allow to make comparisons of the correct location of the peaks before and after the noise subtraction. ECG channel will be used as the gold standard to compare the accuracy of the peak detection in the PPG corrupted signal and the subsequent noise-subtracted PPG signal. Then, the artificial noise model with the best results in the tests in *subsection 3.4.3* is selected and added to the dataset, which contains tachycardia and bradycardia events. Lastly in *Chapter 4*, the aim is to clean the signals from the new dataset and try to increase the sensitivity in the detection of the focused arrhythmias.

Our contributions to artificial noise modeling are shown in the next points:

- Design five different noise approaches under various types of motion artifacts. Particularly, due to low-intensity movements which are within the six activities described in *subsection 3.4.1*.
- Experimental results for the noise models and the correct choice of the most suitable, according to tests.

## 3.4. METHODS FOR NOISE CHARACTERIZATION

### *3.4.1. Data collection methodology.*

This section describes the main features of the wristband-obtained PPG signal that serves as the provider of information for the later noise modeling. The data was obtained from the IEEE 2015 SPC database, available at Zhang's [74] webpage and used in his paper for noise reduction and posterior heart rate estimation.
The IEEE 2015 Signal Processing Cup implements 23 public datasets. Twelve training datasets were measured from people between 18 and 35 years old. They were asked to develop different kinds of activities, where can be found the low-intensity ones. Also, 11 testing datasets were measured from people between 19 and 58 years old. They developed a mix of low,

medium and high-intensity activities. Each dataset was obtained with a sampling frequency of 125 Hz and was measured simultaneously with the ground-truth of heart rate from the ECG signal. The BPM is taken from 8 seconds of windows size and overlaps every 6 seconds. However, this information was not taken in the tests in *subsection 3.4.3* because they barely corresponded with ECG gold standard signal information.
The first 12 training datasets were chosen for the noise characterization, because the testing ones did not have the same movement's structure or defined time intervals. Also because high-intensity activities were induced (such as boxing, jumping), which are not inside our scope.
Each of the 12 training datasets have two channels of PPG signal, three channels of 3-axis acceleration and one channel of ECG signals. Activities were classified into two types:

- **Type 1:** It includes 5 minutes of exercise with 6 types of activities along the entire time: 30 seconds at rest, 1 minute running at 8 km/h, 1 minute running at 15 km/h, 1 minute running at 8 km/h, 1 minute running at 15 km/h and resting the last 30 seconds. At rest activities, the subjects were asked to purposely move the hand with the wristband randomly to generate motion artifacts.

- **Type 2:** It includes 5 minutes of exercise with 6 types of activities along the entire time: 30 seconds at rest, 1 minute running at 6 km/h, 1 minute running at 12 km/h, 1 minute running at 6 km/h, 1 minute running at 12 km/h and resting the last 30 seconds. Apart from the above, other low-intensity movements were also mentioned: forearm and upper arm movements such as clapping, closing hands, pushing, shaking hands and stretching, so the resting activities are also artifact induced.

Due to similarities in the spectrum analysis of type 1 and type 2 activities, which differ only on the treadmill speed in activities type 2 (which are about 6 and 12 km/h), both types of activities are taken as equal. The main frequency components do not show big differences, and this mixture would allow a broader dataset to get a more general noise with a greater number of samples. Then, the final dataset considers the activities below:

1. Rest with hand movements (30 seconds).
2. Running 8 km/h (1 min) corresponds to 6 km/h in activity type 2.
3. Running 15 km/h (1 min) corresponds to 12 km/h in activity type 2.
4. Running 8 km/h (1 min) corresponds to 6 km/h in activity type 2.
5. Running 15 km/h (1 min) corresponds to 12 km/h in activity type 2.
6. Rest with hand movements (30 seconds).

The 12 signals disposed each with roughly 37500 corrupted samples, equivalent to five-minute recording time. Preliminary tests conducted in *Appendix A* allow us to see the spectral differences between a motion distorted PPG signal and a cleaned PPG signal (by our means). This allowed to estimate which components in the spectrum reflect the behavior of noise artifacts. The bandwidth where the largest amount of information is roughly estimated. Thus,

we expect to obtain a model which accomplishes with the conditions in the tests, because a perfect noise modeling is quite improbable.

### 3.4.2. Approximations to noise characterization

#### 3.4.2.1. The Savitzky-Golay noise model

The Savitzky-Golay smoothing filter is a type of digital filtering technique. The process fits successive segments of adjacent data with an adjustable degree polynomial regression, using the linear least squares criteria [75].
The main advantage of this filter is the preservation of the main characteristics of the original distribution of the signal, such as local maxima and minima. It also preserves the width of the peaks, which generally disappear with other averaging techniques.
For this reason, and in our case, it can accurately reduce the level of noise without biasing the morphology of the PPG signal. This filter allows to discriminate high-frequency noise, smoothing the signal and keeping low-frequency features such as baseline drift.

Let the high-frequency motion artifacts be denoted as $n(t)$, which is obtained from the subtraction of the Savitzky-Golay filtered signal[4] $s(t)$ from the original signal $x(t)$:

$$n(t) = x(t) - s(t) \tag{1}$$

Now, once we have the high-frequency noise components, we have to obtain the low-frequency noise component. Let the baseline drift noise component be denoted as $m(t)$, and $o(t)$ be the entire noise motion artifacts:

$$o(t) = n(t) + m(t) \tag{2}$$

As the low-frequency component $m(t)$ is difficult to track in the frequency domain (represented by an impulse near to zero), it is suggested to track it from the time domain. Following this reasoning, the baseline drift noise component (low-frequency noise) is characterizing the DC component generated primarily due to the patient movement. Because of the visible changes in each activity DC level, it is necessary to determine the baseline drift corresponding for each activity performed by the patient. Thus, two methods for obtaining the low-frequency noise component are described as follows:

A. **Mean Values:** this method finds the six most significant mean changes in the specific activity section of the original signal $x(t)$. Then, the six means are linked as a step

[4] A similar approach was intended using sym4 wavelet. This procedure allowed to extract high frequency noise using thresholding. This wavelet thresholding procedure removes noise by thresholding the wavelet coefficients of the detail sub bands while keeping the low-resolution coefficients unaltered [76]. The high-frequency noise component was extracted following the same procedure in equation (1). However, this approach was not used because the subtraction produced a really little amplitude noise. This noise did not resemble the desired features from high frequency noise artifacts *[Appendix A]*.

function. The output is a signal containing the estimate of the baseline drift over the entire length of the signal.

B. **Detrend Signal:** this method finds the trend of the original signal, using the number of desired breakpoints over the length of $x(t)$. These breakpoints are evenly spaced, and then they are linked by linear functions. The output is a signal containing an estimate of the baseline drift over the entire length of the signal.

Mean values function fits in a better way the error analysis in Savitzky-Golay noise model, and for this reason we decided to obtain $m(t)$ component through method A. Meanwhile, method B will be used for the obtention of $m(t)$ in both of the next models (linear predictor noise model and moving average noise model).
The Savitzky artificial noise model $o(t)$is then obtained using the equation $(2)$

#### 3.4.2.2. Linear predictor noise model

The linear prediction filter allows estimating future values of a discrete-time signal as a linear function of previous samples. For this, it determines the coefficients of a forward linear predictor by minimizing the prediction error in the least squares sense [77].
The linear prediction filter coefficients are computed for the high-frequency noise extracted from the approximation in $(1)$.
A FIR filter based on these configurable order coefficients is used to estimate the value of the predicted $n(t)$ signal, based on past samples of each point of the input signal. This generates a correlation between its samples. With this process over the Savitzky high-frequency noise, we expect to have a more general noise, simulating the case of having more than 12 realizations.
The estimated high-frequency component found by LPC is denoted as $\hat{n}_{LP}(t)$.

$$\hat{n}_{LP}(t) = LPC\{n(t)\} \tag{3}$$

Next step is obtaining the low-frequency component, following the guideline in $(2)$. For this purpose, method B from *subsection 3.4.2.1.* was used. By detrending the mean values of the corrupted signal, it is possible to obtain a DC level that is similar to the one obtained with the method A. So, if we name $d(t)$ to this low-frequency component, the equation$(2)$ can be rewritten for this specific noise approximation as

$$o(t) = \hat{n}_{LP}(t) + d(t) \tag{4}$$

#### 3.4.2.3. Moving average noise model

In this model, our aim remains in generalizing the Savitzky noise model in order to simulate more realizations from which this noise model was obtained. For this, the high-frequency noise component from Savitzky is used as an input signal to the moving average filter.

The moving average approximation uses the moving average filters (MAF). These compute a low pass filter FIR, commonly used for determining the average values of a signal within a period *T*. As the length of the filter increases, so does the smoothing of the signal and the sharp modulations become blunter. The moving average filter can also be understood as a convolution that averages peaks or outliers on many symmetric neighboring points concerning a reference point on the input signal [72]. *Appendix A* provides more detailed information about this technique.

The high-frequency noise from Savitzky is seen as a plain averaging, where extrema points may bias the total outcome of the noise. This approach will simulate the case where more realizations were considered by setting the proper window sizes. Thus, a more general noise model can be designed, hoping that it accurately describes motion artifacts for the activities explained previously.

We noticed that when the heart rate is slow, more waveform features are visible, whereas when the heart rate is fast, only the main peaks of the waveform remains due to the fast increase on each waveform slope. This is the main reason why it is expected to have a smaller window size for the MA filter when the heart rate is slow because it will produce a smaller smoothing action. This smaller window simulates the case where the averaging is made between much more components, without biasing or reducing in a considerable measure the high-frequency noise. On the other hand, in running activities when the heart-rate is fast, there is less distortion over the signal due to the high slope at each PPG waveform onset. This high slope fades high-frequency motion artifacts and consequently, they end up with a lower amplitude. This explains why in the case of jogging or running, the size of the window should be wider. The softness will be greater and then the steep peaks that do not characterize high-frequency noise in this case will be cut.

To obtain window sizes, we take into account the length of an M-M interval for a particular activity. The window size can be chosen between zero and the result of dividing the length of the signal in that activity by the number of peaks found in that section. This will set the denominator coefficients in the transference function of the filter. Since the number of beats changes for each activity and each realization, outcomes showed an interval for the window sizes.
Therefore, the optimal window sizes were computed empirically using an iterative script which loops over steps of length 5 for both windows (one for resting activities and another for running activities). The iterative algorithm found that the window size has an optimal value of 45 for rest activities and 70 for major intensity activities. This makes sense because in resting physical activities, heart rate is slow and as explained before, the window's size should be smaller in order to remove outliers without biasing the signal features. In the opposite case, for running activities, a greater window size sample will consider a larger number of neighboring samples. Therefore, it will decrease outliers greatly, presenting a slow adaptation over significative fluctuations during short periods.

This process will average every sample within the selected window, creating our high-frequency noise component.
The moving average filter is fed with the Savitzky high-frequency noise, which is extracted from the approximation in $(1)$**.**

$$\hat{n}_{MA}(t) = MA\{n(t)\} \tag{5}$$

Adding the low frequency noise $d(t)$, the expression $(2)$ for this specific noise approximation can be rewritten as:

$$o(t) = \hat{n}_{MA}(t) + d(t) \tag{6}$$

#### 3.4.2.4. The dynamic variance moving average noise model

This model is fed with the high-frequency noise component from the moving average model $\hat{n}_{MA}(t)$ described in $(5)$. A dynamic character is desired in the noise, that is, a major number of artificial noise models simulating new realizations. Therefore, a change in the noise variance was proposed to simulate this artificial noise. This can be done by multiplying the sample standard deviation by a real value, as shown in the following equation. In this way, sample variance is indirectly changed.

$$\hat{n}_{base}(t) = E[X_k(t)] + \sqrt{V[X_k(t)]}.F \tag{7}$$

Where $X_k(t)$ represents all of the high-frequency noise components from the 12 realizations of the dataset, stored as a collection. Then $E[X_k(t)]$ would correspond to an average of all high-frequency noise components. This term can be seen as a sample mean, and therefore, the unbiased sample variance can be computed from the moving average noise model and saved in $V[X_k(t)]$.
Factor $F$ changes the variance and creates a new artificial noise model for each iteration. This factor can take any desired value, but we considered values from (-3, 3) since broader factor values may result in extreme values that would not characterize the noise of the low intensity activities. Higher factor values may represent artificial noise models which represent motion artifacts with higher intensity levels of movement, which are out of our scope.
The resulting high-frequency noise component with the new variance is then filtered with a passband filter. This is made to emphasize the high-frequency components from 2 to 26 Hz following Solosenko suggestion [12]. The analysis of these spectral noise components can be found in *Appendix A.*

$$\hat{n}_{DVMA}(t) = PBF\{\hat{n}_{base}(t)\} \tag{8}$$

Finally, the low-frequency noise component is added to the high-frequency noise, producing the entire DVMA noise model.

$$o(t) = \hat{n}_{DVMA}(t) + d(t) \quad (9)$$

#### 3.4.2.5. Band-limited gaussian noise model

This model is fed with the high-frequency noise component from the moving average model $\hat{n}_{MA}(t)$ , described in $(5)$**.** The low-frequency components will be treated separately and will be added to set the total artifact noise model.

We considered each one of the signals of this dataset as a realization of a stochastic process. A stochastic process or a random process is a mathematical tool that joins a collection of random variables that model a system in time or space following probabilistic laws. [78].
Due to digital quantization during the digital conversion, it is said to be a discrete-time stochastic process. However, it will be treated mathematically as a continuous-time stochastic process, given that cardiac and blood volume signals are continuous over time.
This model will assume that every one of the samples of each high-frequency noise signal is a random variable $X(t)$. The collection of these $T$ random variables will conform to a random vector. Then, each realization can be seen as a random vector $\{X_k(t)\}, t \in T$**;** where $k$ represents the specific realization, and $T = \{1,2,3,\dots,N\}$. $N$ is the final sample for each realization (in other words, $T$ refers to the length of the high-frequency noise realizations).

It will be assumed a normal distribution to model every random variable $X(t)$ in all the realizations$\{X_k(t)\}$,because this probability distribution relates almost every noise process known in nature. As far as our knowledge goes, it's also the most used process for every noise approximation in the literature [79].
Thus, for the estimated high-frequency noise, each random variable $X(t)$**,** will create a noise value for each $t \in T$**.** These random variables will be affected by the unbiased sample mean and variance computed from the base noise. The result is a high-frequency noise estimate which obeys in a certain way the behavior of the 12 realizations of the base noise. This base noise is the moving average high-frequency noise $\hat{n}_{MA}(t)$. and would be the primary root for our Gaussian basic noise model.

High-frequency noise was found to be non-stationary, because of the variation of the sample mean and sample standard deviation over time. These results can be seen in *Appendix A.*
It is also shown that every one of the $\{X_k(t)\}$**,** are autocorrelated**.** Given that the autocorrelation is denoted as:

$$R_{XX}(t_1, t_2) = E[X(t_1) * X(t_2)];\ t_1 \in T, t_2 \in T \quad (10)$$

Correlation matrix was computed for the 12 realizations of the process. *Appendix A* provides this matrix*,* where the results showed that every $X(t)$ that composes the random vectors are correlated with the rest of themselves. Then we know that the random variables are not independent of each other. Therefore, we assume that every$\{X_k(t)\}$must have a memory feature**.**

The Gaussian Noise model is obtained from a normal distribution random seed generated for each sample along the entire noise signal length. There will be a different model each time it is retrieved from the function. *Appendix A* shows the generation of the random variable for each sample.

$$\hat{n}_{gbase}\,(t)\; = E[X_k(t)] + \sqrt{V[X_k(t)]}.S_i\;;\; t\;\in\;T, i\;\in\;T \tag{11}$$

Where $S_i$ is the random variable acquired in the i-th loop iteration and will vary according to the range of values obtained in a standard normal distribution. About 95% of the obtained values are located within two standard deviations of the mean and has more probability of being chosen [80]. However, by generating the samples in this way, it does not assure that the samples are correlated. For this reason, passband filtering is proposed after this process.

In Solosenko research, noise estimation for distorting PPG signals is made by filtering out white noise [12]. Here, it was stated to filter this white noise according to the spectral features of the PPG signal, which were not mentioned. The design of a simple passband filter with cut frequencies $f_{C1} = 2.5Hz$ and $f_{C2} = 26\;Hz$ was made.

$$\hat{n}_G(t) = PBF\{\hat{n}_{gbase}(t)\,\} \tag{13}$$

After passband filtering these models, the low-frequency noise was added, obtaining a rewritten expression for $(2)$

$$o(t)\;=\;\hat{n}_G(t) +\;d(t) \tag{14}$$

### 3.4.3. General statistics of characterization approximations

The following tests will be executed based on a peak detection assessment. It is expected that the artificial noise models allow a better peak detection after subtracting them from each one of the signals used to create them. In this way, we prove the viability of the resemblance with the well-defined activities named before. Then, the following procedure was performed with the same training signals in the IEEE Processing Cup 2015 database for each one of the models:

1. Extract $o(t)$ for each activity in each realization.
2. Calculate the number of peaks in $x(t)$ and in $y(t)$for each activity, where:

$$y(t) = x(t) - o(t)$$

3. Extract the number of peaks in the ECG channel for each activity in each realization. Given that the ECG signal was affected by motion artifacts too, these were removed using the denoising described in a more detailed way in *Appendix A.*

4. Calculate two experimental errors having as real values the peaks found in the ECG signal and as experimental values, the peaks in $y(t)$ and $x(t)$. Then compare if the experimental error from $y(t)$ has improved about the one obtained with $x(t)$, resulting in a logical value which was defined as BOW (better or worse) index.
5. Compute regression type errors like MSE (mean squared error), RMSE (root of the mean squared error), and MAD (mean absolute deviation), using the number of peaks from the reference signal ECG, $y(t)$ and $x(t)$.
6. Obtain classification errors by synchronizing both ECG and PPG signals. If the performance parameters obtained with the noise-subtracted signal $y(t)$ improve concerning the ones obtained with the original signal $x(t)$, this indicates a better model.

The results of this process are exposed in the summarized tables in *Appendix B*. Here we have the number of peaks found in $x(t)$, the number of peaks found in $y(t)$, the number of peaks found in ECG and the BOW index. Classification and regression errors computing are contained as well.
For the first tests, the relative experimental error is calculated following the next equation:

$$e_{rel} = \frac{f_m - f_r}{f_r} \tag{15}$$

These two experimental errors will allow us to obtain an index which estimates numerically if the model improves the peak detection in several peaks. This index is called the BOW (Better or worse) index, and it is calculated with the coefficients below. These coefficients are given by the distance between two points in one dimension:

$$coeff_1 = |\, peaks\ in\ ECG - peaks\ in\ y(t)\,|$$

$$coeff_2 = |\, peaks\ in\ ECG - peaks\ in\ x(t)\,|$$

As both experimental errors are made with respect to ECG signal, the comparison can be made only with the distance between the points. The $f_r$ value in the denominator of the equation $(15)$ will always be the same for both experimental errors. Then, the BOW index is given by a logic comparator which outputs 1 if the value of $coeff_2$ is higher than the value of $coeff_1$, and 0 in the opposite case.

| Noise models | BOW index values | |
|---|---|---|
| | Class 1: the model ensures a better peak detection | Class 0: the model does not ensure a better peak detection |
| Savitzky Golay | 79.17 | 20.83 |
| Linear Predictor | 70.83 | 29.17 |
| Moving Average | 72.22 | 27.78 |
| Dynamic Variance[5] | 79.17 | 20.83 |

**Table 3.4. Comparison of BOW index values for each noise model**

*Table 3.4* allows to see that apparently, Savitzky Golay along with the Dynamic Variance noise approach are the ones that reduce the distance between real and estimated values after the denoising of the signal, followed by Moving Average and LPC model. With this in mind, regression errors were computed from the number of peaks detected in each signal.

| Signals | Regression error values | | |
|---|---|---|---|
| | MSE | MAD | RMSE |
| Original Signal (PPG w/noise) | 40.25 | 3.72 | 6.34 |
| Savitzky Golay | 19.82 | 2.62 | 4.45 |
| Linear Predictor | 21.27 | 3.02 | 4.61 |
| Moving Average | 38.25 | 3.72 | 6.18 |
| Dynamic Variance | 33.75 | 3.33 | 5.81 |

**Table 3.5. Comparison of regression errors for each noise model.**

Savitzky Golay stands in the first place, while the Dynamic Variance and Moving Average methods stand third and fourth.

This seems to favor Savitzky Golay model, but the regular results for Dynamic Variance model can be explained by the fact that errors like MSE and RMSE emphasize the magnitude of the errors, from the estimated value towards the real value. We noticed that the Dynamic Variance

[5] For this test purposes and given that the dynamic variance model varies according to the factor that is provided, the best possible factor was computed for the sensitivity and specificity tests, throwing an optimal *F* value of -0.53 *(Appendix A)*. The regression tests were made with this value, but we understand these values could change depending on the factor that's used. In this occasion, we show the results obtained with the best of the possible values.

method obtains a higher number of instances where there were better peak estimations than Linear Predictor and even Savitzky Golay. However, in the failed estimations, its error magnitude value reaches higher values. In other words, the number of PPG peaks differ in a higher quantity from ECG peaks, but the times this occurs are minimal. Although, the number of times it gets errors is the lowest amongst all the other methods. We take into account a tolerance of 1 beat, because these signals haven't been synchronized yet. Then we make a simple count of how many times this squared error is greater than the tolerance. Dynamic Variance method obtains the lowest percentage (43,06%), while Savitzky Golay, Moving Average and Linear Predictor show a greater number of errors (47,22%; 50%; 54,17%, respectively), but with a lower magnitude. The above can be seen in the tables in *Appendix B.*

Now, we cannot decide by just having in account previous results. The number of peaks counted before and after denoising the signal is not enough to measure the performance of the noise model. In order to set an evaluation under the same conditions, these PPG peaks must be detected in the same locations or nearly enough from the peak locations of the ECG after subtracting the characterized noise. If the number of PPG peaks is closer to the number of peaks in ECG, but the position of these two does not match, it means that those are fake PPG peaks. These could be either fake positives (additional peak) or fake negatives (missing peak). This could, in turn, imply that the noise subtracting process would be introducing more errors than the ones in the original signal. In other words, the noise would not characterize in a right way the features of the signal.
For this third and last analysis, performance parameters such as sensitivity, specificity, and accuracy were measured before and after the noise subtraction. A more detailed view to the classification consideration can be seen in *Appendix A*. ECG and PPG signals were aligned, since these signals share a correspondence but their main peaks occur at different times. The lag occurs because of the delay of the blood pump from the central part to the distal part of the body. If the performance parameters improve after denoising the signal, it means that this model fits the requirements and features of the twelve set of realizations. These tests were also made with the original signal, because this would allow us to see which model is really improving the peak detection from the noisy conditions.

| Signals | Classification performance parameters | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 70.01 | 71.39 | 70.70 |
| Savitzky Golay | 69.00 | 69.54 | 69.27 |
| Linear Predictor | 68.93 | 69.36 | 69.15 |
| Moving Averages | 71.56 | 72.74 | 72.15 |
| Dynamic Variance | **71.98** | **72.76** | **72.37** |

**Table 3.6. Comparison of performance parameters for each noise model.**

*Table 3.6* shows that not every model improves the results from the original signal. Just two of them overcome the results from the original PPG signal. This can be explained by the number of false positives obtained in the other models, which decreases the classification quality. It might suggest that even when the number of peaks was nearer to the real value, they were not M peaks. The morphology of the PPG signal makes it more prone to commit mistakes in the peak detection, given that the Q peak might sometimes mask the valuable M peak, acting like one when actually, it is not.

However, it appears that some of the noise models could work better for noise characterization in certain activities. We computed these performance parameters for each activity in all models and compared them to the ones obtained in the original PPG. This would give us greater certainty of the chosen noise. For future work considerations, we propose the attempt to recreate a hybrid kind of noise with the better results obtained.

| Signals | Classification performance parameters for activity 1 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 81.58 | **83.87** | **82.72** |
| Savitzky Golay | 71.77 | 73.1 | 72.44 |
| Linear Predictor | 69.61 | 72.33 | 70.98 |
| Moving Average | 79.60 | **82.65** | 81.13 |
| Dynamic Variance | **83.11** | 82.32 | **82.72** |

**Table 3.7. Comparison of performance parameters for each noise model in activity 1.**

| Signals | Classification performance parameters for activity 2 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 64.19 | 63.1 | 63.65 |
| Savitzky Golay | 64.71 | 63.27 | 64 |
| Linear Predictor | 65.18 | 62.47 | 63.82 |
| Moving Average | **71.79** | 70.32 | **71.07** |
| Dynamic Variance | 71.39 | **70.66** | 71.03 |

**Table 3.8. Comparison of performance parameters for each noise model in activity 2.**

| Signals | Classification performance parameters for activity 3 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 65.04 | 66.58 | 65.81 |
| Savitzky Golay | 64.64 | 64.47 | 64.56 |
| Linear Predictor | 65.13 | 64.65 | 64.89 |
| Moving Average | **65.47** | **66.85** | **66.16** |
| Dynamic Variance | **65.47** | 66.13 | 65.80 |

**Table 3.9. Comparison of performance parameters for each noise model in activity 3.**

| Signals | Classification performance parameters for activity 4 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 73.5 | 76.53 | 75.01 |
| Savitzky Golay | 72 | 73.65 | 72.82 |
| Linear Predictor | 70.71 | 72.46 | 71.58 |
| Moving Average | 74.12 | 76.69 | 75.41 |
| Dynamic Variance | **75.08** | **77.51** | **76.30** |

**Table 3.10. Comparison of performance parameters for each noise model in activity 4.**

| Signals | Classification performance parameters for activity 5 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 70.12 | 71.58 | 70.85 |
| Savitzky Golay | 69.9 | 71.3 | 70.6 |
| Linear Predictor | **70.31** | **71.92** | **71.11** |
| Moving Average | 70.11 | 71.67 | 70.89 |
| Dynamic Variance | 70.22 | 71.77 | 71.00 |

**Table 3.11. Comparison of performance parameters for each noise model in activity 5.**

| Signals | Classification performance parameters for activity 6 | | |
|---|---|---|---|
| | Sensitivity | Specificity | Accuracy |
| Original Signal (PPG w/noise) | 79.28 | 80.2 | 79.74 |
| Savitzky Golay | 80.12 | 80.06 | 80.09 |
| Linear Predictor | 81.22 | 81 | 81.12 |
| Moving Average | 81.02 | **81.42** | 81.23 |
| Dynamic Variance | **82.25** | 81.22 | **81.75** |

**Table 3.12. Comparison of performance parameters for each noise model in activity 6.**

From the results obtained in the previous tables, it seems that none of the noise models makes a perfect characterization of the noise. Even the Dynamic Variance which was the only one with the best results in *Table 3.6*. However, we noticed that in over half of the activities, this model improves peaks detection. This model obtains a better close up to the desired results, specially in the low-intensity activities.

For filtered Gaussian noise model, only classification tests were made due to the complete randomness of the process. We obtained 57 different models, which were tested to obtain a range of sensitivity and specificity results. An interval for the sensitivity of (51.40 - 58.99) was reached, as well as an interval for the specificity of (62.04 - 69.81). This shows that this model might not be resembling the specific motion artifacts within this dataset, but a more general case. Although it is the one with a more solid mathematical base, it is left for future works given the limited number of realizations that are disposed of.

## 3.5. RESULTS AND ANALYSIS

### 3.5.1. Discussion

The main objective of this chapter was the obtention of a noise model from IEEE 2015 Signal Processing Cup datasets. Although the results obtained are positive, they show that noise characterization techniques still can be improved for specific kind of movements. Logically, most of the authors are centered on noise elimination for a posterior classification of cardiovascular events.

However, the opinion from Solosenko about the lack of annotated datasets cannot be discarded. His statement could be the right motivation for other experimental works on PPG signals where it is needed to contaminate clean signals, or in the case that a stress test is not feasible.

Solosenko [12] and Wartzek [64] propose approaches having in account spectral and statistic features, respectively. Those suggestions were considered for our movement artifact noise

modeling, as well as the Moving Average from Han [72] and linear prediction coefficients from Hassani [66].

The Savitzky-Golay method was the base for the other noise models, but it exposed a poor performance at the classification error level. The good results at regression error are emphasized too. This seems to indicate that it makes a proper approximation to the real values of peaks but induces a high number of false positives with it. This might suggest that a plain averaging of high-frequency noise realizations is just not enough to characterize the each one of the realizations in a balanced way. Further analysis should be considered, as the one which was made with the Moving Average model.

According to the tests conducted in *subsection 3.4.6.,* the Dynamic Variance noise model is the most suitable representation of motion artifacts for the majority of the activities, reaching a general increment of 2% for sensitivity, specificity, and accuracy. The best results for Dynamic Variance model can be seen in *tables 3.7,3.8,3.9,3.10* and *3.12*. This exposes the adaptability of the dynamic variance moving average noise model, specially for low intensity levels of movement. Thanks to the optimal window sizes computed, it could be analyzed that when heart rate is slow, more waveform features are visible. Any small change over the waveform due to a particular random noise can abruptly change the visibility of M, P, and Q characteristic points in the PPG signal. This can be seen from the results in this model in all activities concerning resting activities 1 and 6. When the heart rate is fast, that is, for activities 2-5 only the main peaks of the waveform remain due to the fast increase on each slope's waveform. In other words, the heart pumps blood with higher pressure as the intensity of the activities arises, thus making the sensor to graph steeper slopes, which are hardly affected by high-frequency noise components (these high-frequency components throw an insignificant contribution on a steep slope, making them harder to notice and damage the signal). We should expect more distortion over the time domain if high-intensity movements were considered.

Gaussian Filtered noise model shows an advantage, because of its entire randomness. However, we consider that the noise sample size is small for making generalizations as a stochastic process. Twelve noise signals are not enough to make characterizations about all possible noise features, as seen in [64].

With these results, the Dynamic Variance method would be the most suitable for the majority of activities, since sensitivity tests were surpassed. This model did not make an excellent performance at regression tests, but it can be seen that the number of false positives and false negatives it introduces after denoising is the lowest one of all models. In addition to this, we can generate an infinite number of noises for this model, with the moving average model as a particular case. *Tables 3.8* and *3.9* show that in some cases, the Moving Average obtains better results than the Dynamic Variance model. This appears to suggest that the DV model is a versatile method, which could obtain a better noise characterization for each activity by just adjusting the variance factor, thus setting a proper model for a specific activity. This model has also a randomness feature, given the infinite values that *F* could take. A boundary between a model that characterizes the signal, but at the same time demonstrates randomness is noticed. We intended to overcome these limitations through an approach that shows the possibility for a movement noise approximation. The best obtained model exposed suitable results for the noise characterization even if the improvement of general classification results was slight.

As a conclusion, our data suggest that this extensive line of noise modeling should be followed for future works. Others could take as base the models created here or combine them with other techniques.

### 3.5.2. Conclusions

In this chapter, a noise model that resembles motion artifacts caused by well-defined activities with low and medium intensity levels is proposed. The best results can be obtained through the dynamic variance noise model, which allows characterizing properly the motion artifacts caused by specific activities and then improving the peak detection assessed by the reference ECG signal. These outcomes show that it is possible under certain conditions to obtain a movement artifact noise model for different activities as long as these are well defined and repeatedly executed. The lack of realizations disposed was an undesired but inevitable pitfall in our case. The recreation of a movement artifact noise model enables the possibility to create new datasets under realistic conditions for simulating physical activity situations. In our case, the physical activities of our interest are the wrists and hand movements presented in 1 and 6 activities.
The model could be seen as a contribution for future works on noise removal algorithms validation and posterior arrhythmia detection. As the time intervals of the movements are known, an annotated noise dataset can be considered as one of the first contributions to PPG noise modeling.
Further analysis for the band-limited Gaussian noise model is suggested. We could obtain a more general motion artifact representation for low and medium intensity levels of movement. Although, due to its randomness, this model was not able to fit the requirements for these specific tests, seeming like this is a more general approach to motion artifacts modeling. This is stated as a future work because of time scope limitations.

## 3.6. GENERATION OF THE DATASET WITH NOISE.

The chosen noise model will be summed to each one of the signals of the new dataset. The structure of this new dataset will be explained with more detail in *subsection 4.2.3.* Since all scripts used to retrieve the dataset records have the same structure, the sum of noise is easily done by loading it to the main file and then summing it. This noise must be denormalized to the same size as the new database signals, as shown in the scheme of *Figure 3.1.* Due to the chosen Dynamic Variance model, each subset was then corrupted with noise coming from different *F* factor values. The choosing of these *F* values can be seen in *Appendix D* and the results for the different F factors choosing are stated in *Appendix C*. Of course, pre-processing filters included in the original code are removed from the classification algorithms, to simulate in a proper way Bonomi's conditions.

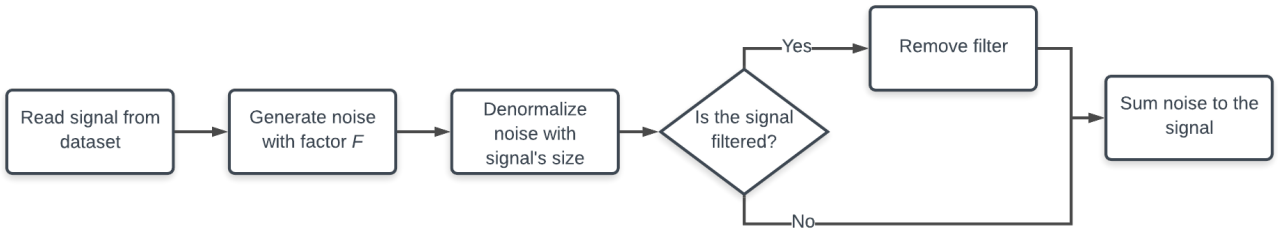



**Figure 3.1. Schematic view of the noise addition process**

*Figure 3.2* shows different segments of one of the signals retrieved from the main loop with and without noise.

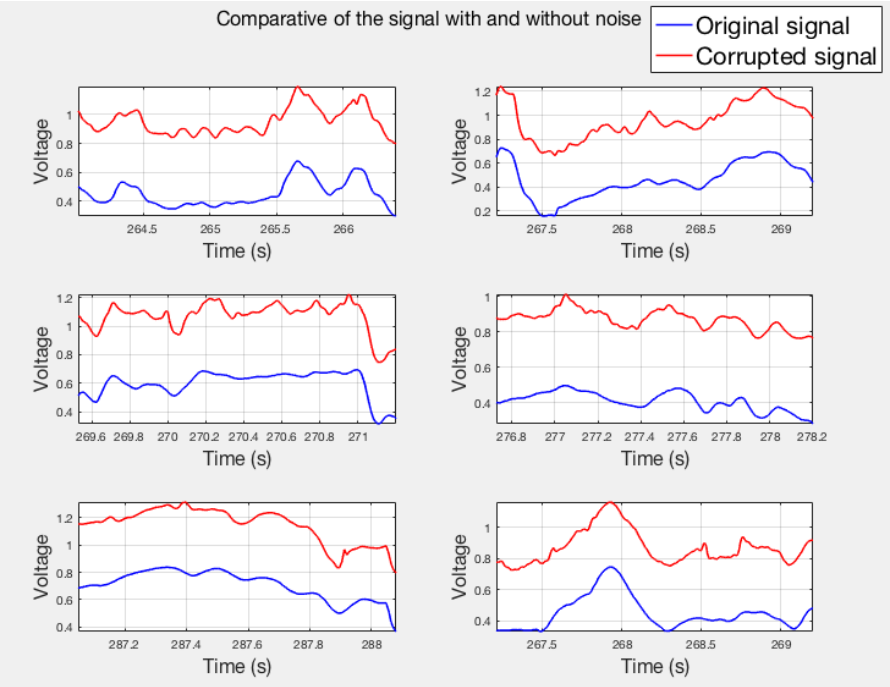



**Figure 3.2. Comparative of the original pulsatile signal and the corrupted pulsatile signal.**

A high-frequency artifact noise is visible, as well as the baseline drift, which is induced depending on the activity that is being executed. High-frequency noise induced by the modeled noise gives the original signal a noisy aspect, which could traduce as fake peaks in the future brady- and tachycardia detection. *Figure 3.4* shows the oscillations of the high-frequency.

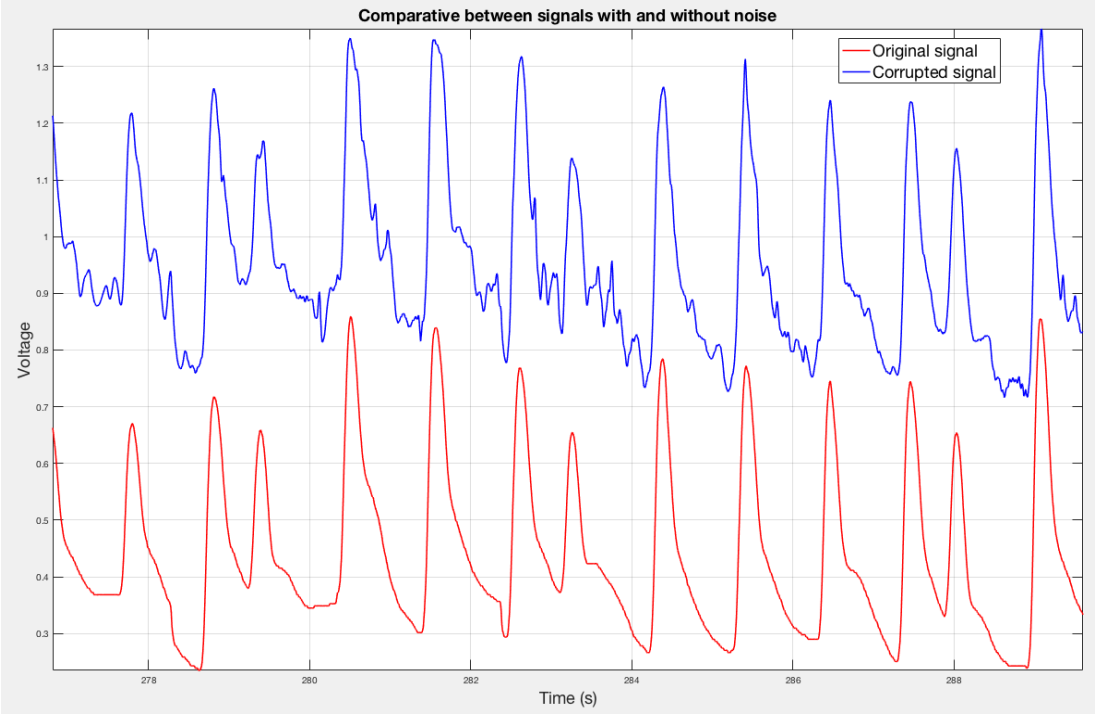



**Figure 3.3. High-frequency noise present in the corrupted pulsatile signal.**

# CHAPTER 4. TACHYCARDIA AND BRADYCARDIA DETECTION MECHANISM

## 4.1. INTRODUCTION

This chapter aims to demonstrate the improvement of several brady- and tachycardia detection mechanisms, using noise removal techniques. We expect to obtain better results from a mechanism with proper denoising processes for the arrhythmia detection rather than a mechanism which receives a noisy signal. This work could provide clinicians with a piece of reliable evidence for them to notice how motion artifacts of even low intensity can change the diagnosis outcomes for brady- and tachycardia detection.

Our approach includes the following tasks: 1) Add noise to the clean arrhythmia dataset, 2) Select denoisers to be used over the arrhythmia detection dataset and select arrhythmia detectors for testing, 3) Implement the denoisers-detectors combination over the newly created dataset, 4) Compare results with and without noise conditions using the denoiser that suits in the best way each detector.

In the first task, we perform the adjustment of the data used in the detection of brady- and tachycardia events. This task creates a simulated dataset of PPG corrupted signals by adding the low intensity movement artifacts over a new dataset which served as the primary arrhythmia data provider. This procedure has already been performed in *subsection 3.6.* The new dataset mentioned is the Physionet CinC Challenge 2015 dataset provided by Physionet [81]. The contents and structure of this dataset will be later explained.

In the second step, the denoising methods are chosen through a review of the literature in bibliographic databases. The selection of the arrhythmia detection techniques is made as well. These are extracted from the challenge outcomes. The Physionet 2015 Challenge seeks to reduce false alarms of arrhythmia detection in ICU. As the outcomes of this challenge, numerous algorithms and techniques for brady- and tachycardia detection have been proposed. A review of the articles that describe each work is performed, in order to choose the appropriate detection methods to be coupled with the denoising phase. The selection of arrhythmia detection techniques is made according to the established selection criteria.

The third step comprehends the implementation of the group of selected denoisers over the group of selected arrhythmia classifiers. This implementation tests will also have the results obtained with the original implementation from each author and the results obtained with the noise simulated conditions.

Lastly, in the fourth step, the results are compared for each denoiser-detector combination, which thrown the best results in each case. This process will allow us to see the improvement over noisy simulated conditions, then answering our research question stated in *Chapter 1.*

This chapter is divided into four main parts: the first part is a conceptual framework where some concepts about arrhythmia detection and the new dataset information are presented. In the second part, a review of noise removal techniques is shown and the denoising techniques are chosen. In the third part, the review for brady- and tachycardia papers from the challenge outcomes are presented. This is made for choosing some of the arrhythmia detection

techniques to create different mechanisms. In the fourth part, the tests for each mechanism are run with and without noise. Results showing the influence of the best case denoiser over the performance parameters of each selected mechanism are presented and analyzed along with the conclusions at the end of this chapter.

## 4.2. CONCEPTUAL FRAMEWORK

### 4.2.1. Arrhythmia detection under noise conditions

Noise conditions can induce distortion to the signal, making the information unrecognizable and therefore leading to a wrong clinical diagnosis. For this reason, several studies have proposed different mechanisms to clean up the signal. Most of them do it by removing the entire portion of the signal that contains the greater corruption of movement artifact noise [19,26]. Other studies opt for denoising mechanisms, so they don't lose information that might be valuable for arrhythmia detection [14,17,29,74].

In the case of PPG signals, tachycardia and bradycardia detection under noise conditions has not been widely performed yet. The lack of open databases with arrhythmia events taken under movement of the patient is the main reason. Another option is the data collecting by our own means, but due to time constraints of informed consents, this could not be performed. The closest approach found was Bonomi's research. He proposed a technique for brady- and tachycardia events detection carrying out their normal routines (in free-living conditions), using a wrist wearable device for measuring PPG and acceleration data. ECG measures were also taken with another device, the total measure time was 24h [19]. We could infer from this paper that the signals used were affected mainly by low-intensity movement artifact noise. The reported sensitivities were 85% and 89% for bradycardia and tachycardia events detection, respectively. Therefore, we expect to improve the conditions for this arrhythmias' detection under the reported low intensity artifact noise (wrists movements, according to Bonomi). It is also expected to obtain better sensitivity results than Bonomi. However, this comparison cannot be made because we don't own the same database as Bonomi, but a simulated scenario. The main objective is not about the comparison with his study but about the improvement that might be induced by choosing an appropriate denoising technique. The main objective will be the choice of the best mechanism for these two arrhythmias' detection. Then we expect that they maximize sensitivity and specificity results. Sensitivity is chosen because it is seemingly the most affected variable due to low intensity artifact noise, according to Bonomi's results. Specificity was also chosen because the Challenge purposes the aim to maximize this variable as well.

### 4.2.2. The 2015 Physionet/Computing in Cardiology (CinC) Challenge: Reducing False Arrhythmia Alarms in the ICU

The dataset that we used as the information base corresponding to specific arrhythmias (tachycardia and bradycardia) was extracted from Physionet, on its 2015 Challenge version: reducing false arrhythmias in the ICU. This dataset contains information from a population of real patients in the ICU, who present five different types of arrhythmias. The participants

elaborate scripts for identification and posterior detection of the five arrhythmias, but we will focus on the two before mentioned.

#### 4.2.2.1. Purpose of the challenge

The Physionet/CinC Challenge 2015 was set as an approach to reduce the incidence of false alarms in the ICU. They encourage the use of different algorithms that involve machine learning or digital processing techniques [81]. False alarms in the ICU can lead to a care disruption through several noise disturbances causing an impact over clinical staff and patients, such as sleep deprivation [82,83,84], inferior sleep structure [85,86], stress for both (patients and staff) [87,88,89,90] and depressed immune systems [91].
These rest disruptions have a substantial effect on recovery and length of stay [92,87] due to elevation in cortisol levels (reflecting increased stress) [89,90]. ICU false alarms have reported rates as high as 86% [93] and only 2% to 9% of alarms are essential for patient management [94]. With this in mind, organizers of the challenge propose a classification task, where each five-minute record must be classified as a true or a false alarm. It should be highlighted that this task differs from Bonomi's one because of the type of assessment done in each one. The challenge considers each record as one instance of classification, while Bonomi considers time intervals within each record or measurement for brady- and tachycardia detection.
Two different kinds of events were proposed by the challenge authors: the real-time event and the retrospective event. Both of these obey to the classification task explained above. The difference between them is the length of the recordings. The real time event assumes that there isn’t any data beyond the alarm triggering, then the classification is made with just the five-minute information of the record. The retrospective event gave 30 seconds after the alarm of information for the participants to decide if use it or not. Both of these events are considered in this work, since they don’t differ greatly in the classification purpose.

#### 4.2.2.2. Tachycardia and bradycardia detection

Bonomi's research is an essential reference for brady- and tachycardia detection. His technique uses the IBI intervals in PPG signals for detecting both arrhythmia events over time.
Brady- and tachycardia sensitivity results obtained in his study could be improved by employed proper techniques over these artifacts. Portions of data were discarded if the movement artifact noise increased beyond a threshold, because these arrhythmias should be detected at rest conditions [33, 35]. But even in this kind of environment, when the person is not moving from one place to another, wrist movements induce low-intensity artifacts, as declared by the authors. Then, we aim to induce the low-intensity noise exactly in the moment of the arrhythmia occurrence. With this, we can evaluate the influence of this artifacts.
After this, our main intention is to improve the sensitivity and specificity results obtained in the arrhythmia detection approaches from the challenge outcomes. We expect this scenario can simulate the arrhythmia detection scenario with low intensity movement noise, which was the one stated by Bonomi as the causative of the detection errors. Thus, it has been proposed a

comparison of the results under this noise conditions and then, applying denoising techniques retrieved from *subsection 4.3*.
The guidelines have been given by the challenge organizers to recognize each one of the following arrhythmias, namely asystole, extreme bradycardia, extreme tachycardia, ventricular tachycardia, and ventricular flutter/fibrillation. According to the challenge documentation, these arrhythmias provoke the trigger of an alarm at the end of the recording (approximately 16 seconds before the alarm). The competitors are encouraged to use the information before the alarm triggering to determine the real state of the alarm, either if it was a true or a false alarm. Specially the last seconds, because within this time interval would be encountered the cause for the alarm triggering. Due to signal length issues stated in *Appendix D*, the noise provoked by activities 1 and 6 from the IEEE SPC database are located in the same time frame as the arrhythmia event.
This requirement fits in a proper way the contamination of the signal, because the portion of the signal which is used to make the determination is contaminated with the low intensity movement artifacts.
The following definition of the five alarms types is used in this challenge [81]:

- **Asystole:** no QRS for at least 4 seconds.
- **Extreme bradycardia (EB):** heart rate lower than 40 bpm for 5 consecutive beats.
- **Extreme tachycardia (ET):** heart rate higher than 140 bpm for 17 consecutive beats.
- **Ventricular tachycardia (VT):** 5 or more ventricular beats with a heart rate higher than 100 bpm.
- **Ventricular flutter/fibrillation (VF):** fibrillatory, flutter, or oscillatory waveform for at least 4 seconds

Our focus will be done over the cases of extreme bradycardia and extreme tachycardia, given that these definitions agree the most with those given in the theoretical framework of *Chapter 2.*

## 4.3. NOISE REMOVAL TECHNIQUES ON PPG SIGNALS

### 4.3.1. Review Methodology

An approach to a review methodology taken in this part of our work is the PRISMA Checklist for scoping reviews (ScR). This checklist is taken as a reference for the next explained steps and for this reason, only the most relevant aspects are taken.

#### 4.3.1.1. Rationale and objectives

This review is made in the context of searching the noise removal methods that are more frequently used to clean PPG data corrupted by movement artifacts. The objectives of this review are the retrieval of some of these methods for a further implementation along with the arrhythmia detection algorithms. The background of this review has already been explained in *section 4.2.*

#### 4.3.1.2. Selection criteria

Articles were selected under a few eligibility criteria such as the year of the publication, the place of the PPG measurement stated in the study (preferably wristband) and the fact that the denoising algorithms were executed in software tools after the measurements (not made in any embedded component).

#### 4.3.1.3. Databases

*Table 4.1.* exposes the results of the review from searches in EBSCO, Scopus, and Google Scholar databases.

#### 4.3.1.4. Search string

The following search strings were used to provide the information compiled in *Table 4.1.*

((artifacts OR noise OR motion AND artifacts) AND (ppg OR photoplethysmography) AND (denoising))

#### 4.3.1.5. Data analysis

The review for noise removal techniques is summarized as follows, using the search string provided above. It was possible to identify 41 articles and reduce them to 13 different articles, thanks to the eligibility criteria established. This allows declaring the 13 articles found in *Table 4.1*. Finally, we explain the techniques that are of interest for our work, and provide an analysis of the advantages and disadvantages of each one of them.

| Paper | Description of the Approach | Analysis of the Results, advantages, and disadvantages | Study Year |
|---|---|---|---|
| Motion Artifact Reduction from Finger Photoplethysmogram Using Discrete Wavelet Transform [96] | - Decompose noisy signal at different decomposition levels in order to analyze different frequency bands.<br>- Soft thresholding method was used to remove the noisy components.<br>- Different wavelet functions (Daubechies, Symlet, Coiflet) and soft thresholding methods ('rigrsure,' 'heursure,' 'sqtwolog,' etc.) were used to denoise the corrupted PPG signal. | - The mother wavelet 'db6' and 'rigrsure' soft thresholding method showed the best result using 5 level decomposition. This method could be an approach to a denoising process.<br>- Statistical parameters are taken instead of classification ones. | 2019 |
| Improved Heart Rate Estimation from Photoplethysmography during Physical Exercise Using a | - Preprocessing is made by simple passband filtering techniques (0.4 - 3.5) Hz.<br>- The remotion of motion artifacts is made using a cascade and parallel combination | - This approach seems to be a simple and effective technique. As a result, proper performance parameters are found.<br>- However, some inputs (like | 2019 |

| | | | |
|---|---|---|---|
| Combination of NLMS and RLS Adaptive Filters [97] | of NLMS and RLS adaptive filters which have as the input PPG and tri-axis accelerometer data.<br>- The denoised signal from NLMS and RLS filters are combined through convex combination.<br>- Average Absolute Error: 0.96 BPM<br>- The standard deviation of Absolute Error value: 1.21 | acceleration data) for this filtering are unavailable for the arrhythmia dataset we have created, since we made an estimation of the movement artifacts only for the PPG signal. | |
| Robust Heart Rate Monitoring for Quasi-periodic Motions by Wrist-Type PPG Signals [98] | - In this paper, a robust HR monitoring scheme for different quasi-periodic motions using wrist-type PPG is proposed, which consists of dictionary learning techniques.<br>- Also it is used a human motion recognition technique for the current motion recognition and dictionary selection. Additional sparse representation-based MA elimination is used for denoising, and spectral peak tracking for HR-related spectral peak tracking. | - The proposed scheme is robust to MA caused by different motions and has high accuracy.<br>- Dictionary learning and human movement recognition techniques are too complex to be implemented, therefore are out of the scope of this work. Mathematical bases are described in the paper, but open scripts or examples about these techniques are not provided.<br>- The same disadvantage as [97] is noticed, given that they necessarily use accelerometer data from their database. | 2019 |
| Denoising of PPG signal by wavelet packet transform [99] | - PPG signals are denoised with wavelet techniques. Then, the signal is reconstructed and the denoised signal is obtained.<br>The results are compared to adaptive filter techniques like Least Mean Square (LMS), Normalized Least Mean Square (NLMS), Recursive Least Square (RLS) and Savitzky-Golay filter. | - The experimental result reveals that wavelet packet transform based thresholding gives better performance than adaptive filtering techniques and Savitzky-Golay filter.<br>- The Mother wavelet family for signal decomposition is not specified. | 2018 |
| VLSI Wavelet-Based Denoising of PPG Signal [100] | - They consider power line noise (60 Hz) for the performance evaluation of VLSI Wavelet-based denoising of PPG signal.<br>- Different kinds of Wavelets such as db4, Coif1, Haar for denoising are considered.<br>- Reconstruction is made with thresholding methods.<br>- db5 denoising seems to show better results compared to other wavelets used in the denoising the PPG signal. | - Wavelet methods seem to be the most used.<br>- A heuristic approach is shown in this paper by comparing different kinds of mother wavelets, then allowing to conclude that there is not a predefined wavelet for a determined problem. | 2015 |
| Evaluation of Wavelets for Reduction of Motion Artifacts in Photoplethysmographic Signals [101] | - Daubechies (db10), biorthogonal (bior 6.8), reverse biorthogonal (rbio6.8), symlet (sym8) and Coiflet (coif 5) wavelets are used in the evaluation.<br>- db10 wavelet has the best results for retrieving respiratory signal without | -The Daubechies was established by the authors as the most preferred wavelet for pulse-oximetry applications and additionally, the results showed a better respiratory signal resorting | 2010 |

| | | | |
|---|---|---|---|
| | disturbing the signal in the attempt of removing movement artifacts. | while removing motion artifacts (96% of the motion artifacts making bending movements and 92% of the motion artifacts making vertical movements). | |
| PREHEAT: Precision heart rate monitoring from intense motion artifact corrupted PPG signals using constrained RLS and wavelets [102] | - In this paper, a constrained recursive least-squares (cRLS) based denoising technique is presented.<br>- A coupled, Wavelet-Fourier based frequency estimation technique is used to estimate HR more precisely.<br>- To separate the clean PPG, the MA corrupted composite PPG signal is decomposed using the ensemble empirical mode decomposition (EEMD) based signal separation technique.<br>- Mother Morlet wavelet is used.<br>- When EEMD fails to separate the clean PPG from the MAs effectively, the new cRLS based denoising is used.<br>- They use the same dataset as Zhang [74]. | - The RLS denoising technique is used. However, authors declare in their state of the art that some of the former approximations done before with RLS have limitations (may not work well universally for all windows of the PPG signals). Moreover, they propose the cRLS method for dealing with these limitations.<br>- cRLS technique uses the accelerometer signal for the first artifact reduction task, which is a limitation in our case.<br>- Wavelet techniques are used for another purpose (heart rate estimation). | 2017 |
| SVM-based spectral analysis for heart rate from multi-channel WPPG sensor signals [103] | - A mixed approach called Mix-SVM is proposed. It uses the same dataset as Zhang [74]<br>- The proposed method is comprised of five main stages, of which three of them are for denoising: preprocessing, initial motion artifact reduction, and sparse signal reconstruction model. | - Although SVM is used for heart rate estimation, the denoising processes are done by adaptive filtering of PPG signals (LMS filtering) using a “reference” motion artifact signal which was extracted from Principal Component Analysis (PCA)<br>- Reference motion artifacts extraction is made through PCA of the accelerometer, which is a limitation in our case.<br>- The average absolute error in the approach is 1.01 bpm<br>- The Pearson correlation was 0.9972. | 2017 |
| Unobtrusive heart rate estimation during physical exercise using photoplethysmographic and acceleration data [104] | - The proposed algorithm consists of four stages, of which the first two were for denoising purposes: a wavelet-based denoising and an acceleration based denoising.<br>- A Meyer wavelet was chosen as mother wavelet in all DWTs.<br>- For the proposed wavelet-based denoising system, a three-level decomposition-reconstruction tree is | - Results are satisfactory for this study, considering that they used the same dataset as Zhang [74] and obtained similar results regarding the mean absolute error and a smaller standard deviation.<br>- Even when it has already been said that we do not have accelerometer data, the approach using wavelet decomposition, and | 2015 |

| | | | |
|---|---|---|---|
| | realized, taking into account that the sampling frequency is 125 Hz.<br>- Results obtained:<br>Mean absolute error: 1.96 bpm<br>Standard deviation: 2.86 bpm<br>Correlation: 0.98 | specifically the decomposition-reconstruction level, is considered for our work. | |
| Photoplethysmogram (PPG) signal analysis and wavelet denoising [105] | - Besides showing the disadvantages of other types of widely used denoising techniques such as Fast Fourier Transform (FFT), adaptive filters, Periodic Moving average filter (PMAF), and LMS adaptive filter; they propose to analyze a wavelet denoising method through various mother wavelets.<br>- Three mother wavelets are studied, and the performance of them is evaluated based on the cross-correlation with the original signal. Haar, Daubechies (db4) and symlet (sym3) wavelets are used. | - The signal reconstruction results are more accurate with db4 mother wavelet. Other wavelets had a lower cross-correlation with the original signal, showing that they are less effective for their data. | 2014 |
| Dual-tree complex wavelet transform motion artifact reduction of PPG signals [106] | - Dual-tree complex wavelet transform (DT-CWT) and the inverse of DT-CWT are presented. This method performs two wavelets procedure in parallel using the real and imaginary part followed by the common high and low filtering banks.<br>- db10 wavelet is used to compare it with the proposed method. | - This research proposes a method called "dual-tree complex wavelet transform" as a mechanism mixed with a thresholding procedure. It merely mentions its outcomes comparing them to the Daubechies 10 mother wavelet. However, the authors do not provide the code.<br>- Using real and imaginary part of the wavelet transform is taken as an interesting method, but a challenging one though, since working with both parts increases algorithm complexity.<br>- Daubechies wavelet is taken as a good suggestion for a wavelet decomposition denoising approach. | 2012 |
| An effective method to denoise EEG, ECG and PPG signals based on Meyer wavelet transform [107] | - This approach uses Meyer Wavelet decomposition to extract the motion artifact, which is subsequently utilized as the reference input of an adaptive filter for noise cancellation. | -Many different techniques are mentioned for each signal separately, they do not mention the use of any wavelet technique for PPG, while they do for ECG or EMG.<br>- The authors broadly explain what is the purpose of the wavelet decomposition for noise extraction. However, the authors do not clearly discuss the results nor the improvement provided by the denoising process. In addition, they do not specify which method | 2018 |

| | | | |
|---|---|---|---|
| | | was used.<br>- As a conclusion, discrete Meyer wavelet with soft thresholding is said to be effective for artifact removal, but it is not clear over which signal it is effective. | |
| Denoising of EEG, ECG and PPG signals using the wavelet transform [108] | The proposed algorithm makes a peak detection in all three signals, and then a new one after the wavelet decomposition. The signals used in the process are not explicit referred by using just the notation for ECG signals. | - Wavelet transform may be carried out for EEG, ECG, PPG denoising. The authors state that wavelet packets can be used as well, and provide higher consequences. They provide results for each denoising process of the signal and SNR is calculated. SNR results are better for EEG, PPG, and ECG in descending order.<br>- The Mother wavelet used for the decomposition techniques is not specified. | 2018 |

**Table 4.1. Summary of analyzed works for noise elimination techniques in PPG signals**

#### 4.3.1.6. Analysis of the review

*Table 4.1* highlights important information about these denoising techniques. First, the majority of the used techniques are those who perform a wavelet decomposition (either by wavelet packet transform or discrete wavelet transform) [96, 99, 100, 101, 104, 105, 106,107, 108]. In the second place, stand the adaptive filtering methods, namely LMS, NLMS, and RLS. However, paper [99] suggests that wavelets show better results than adaptive filtering or smoothing techniques (where Savitzky-Golay filtering is briefly mentioned). Authors consider performance measures such as MSE and PSNR, to make a comparison. Techniques as the ones presented in [98] could be overly advanced for what is needed, which is a fast and effective denoising of the signal so that it does not affect the internal processing performance of the given arrhythmia detection algorithms.

EEMD method is mentioned in [102] for artifact separation from the signal, just as another reviewed approximations like [17]. Also, in [44], it is defined as an effective method for artifact removal on systems that have only a single channel. It is stated that these types of methods are usually utilized in the personal healthcare domain due to the desire for low hardware costs and operational setup complexity. Some limitations over this method are the empirical bases where it comes from, as explained in *subsection 4.4.3.*

Methods like adaptive filters are seen as desirable techniques. However, implementation searches carried out showed that although these are methods available in the documentation from the tool, they need a reference signal so the filters can adequate their transference function (according to the criteria they use to minimize error). As this approach is supposed to simulate a scenario where only the corrupted signal is available, it can't be afforded because one entry for this filter would be missing.

*Figure 4.1* shows that more than half of the studies mention wavelet denoising, where all methods use thresholding to recover the appropriate signal information for reconstruction.

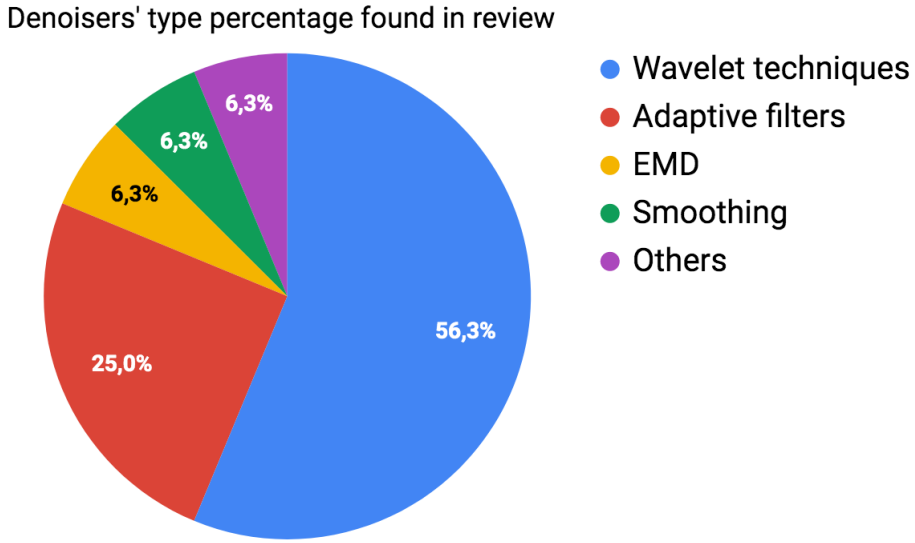



**Figure 4.1. Percentage of denoisers types found in review from *Table 4.1.***

For our work, various techniques are taken into account and then, tests will be made for each one of these them, following the proposed model in [100]. We decided to give a heuristic focus towards wavelets. Wavelets have an enormous quantity of possibilities respecting the mother wavelet that is used, the vanishing moments, and so on. But this could also be seen as a disadvantage, since it does not exist a criterion that allows specifying which mother wavelet is the one that better couples the given application, neither it provides a rule for deconstruction/reconstruction levels.
As a solution, a comparative test is planned, with different types of proposed wavelets which will be chosen following more detailed research on the desired features of these wavelets. This theoretical suggestions plus the results of this review allow to select the wavelet families. Suggestions from the listed studies will be taken for the mother wavelet choosing. For example, the suggestion of [101] for using Daubechies wavelets since they are the preferred ones for pulse oximetry applications. Regarding the reconstruction coefficients, various approaches with fixed reconstruction levels are proposed, and an approach that proves the soft thresholding reconstruction is proposed as well. A final decision on the selected techniques can be seen in *subsection 4.4.5.*

### 4.3.2. Challenges and Contributions

Our main challenge in this chapter is to select the most effective techniques for denoising the PPG signal. This implies that these techniques allow improving the sensitivity and specificity parameters obtained over noisy signals conditions. This task is challenging because there is not a unique denoising method (especially regarding wavelets) for each application that is

needed. Therefore, we propose to make a comparative evaluation of noise removal techniques with each one of the arrhythmia detection algorithms provided.
Our goal is to configure a mechanism consisting of a denoising process and arrhythmia detection process, which is able to provide good classification performance if the input is a noisy signal. We are also trying to identify the best denoising method for each detection algorithm, which is evaluated later in *subsection 4.6.*

## 4.4. APPROXIMATIONS TO NOISE REMOTION TECHNIQUES

### 4.4.1. Wavelet-based denoising

#### 4.4.1.1. General concepts

Wavelets are a new problem-solving tool, which has been applied to several areas such as mathematics, physics, and engineering [126]. Unlike Fourier series which base functions are sinusoids, wavelets create a family of base functions. Depending on the desired properties, these base functions can or cannot be orthogonal (however, it is more likely to seek that they are orthogonal) [125]. Fourier analysis offers a complete frequency domain, but it loses time domain; even with techniques as STFT (by selecting just one segment of the signal), the method is limited due to the fixed size of the window. If a full window is selected, the frequency resolution is weak, as it is stated by Heisenberg principle of uncertainty (frequency, and duration cannot be measured for a determined precision degree) [127].

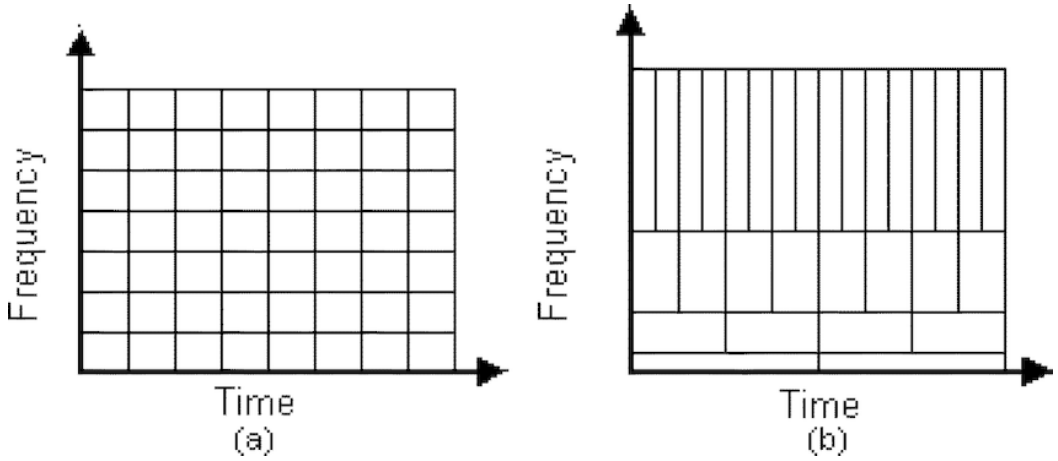



**Figure 4.2. A grid is showing how STFT and WT split the time-frequency domain *[153]*.**

Wavelets allow a time-frequency analysis through two processes: scaling and translation [128]. Both procedures are stated as a parametrization, which is made in the mother wavelet expansion equation. Every function $f(t)$ can be expressed in terms of a double parameter system, which orthogonal base is the mother wavelet and it's characterized by expansion coefficients. These coefficients are denominated as discrete wavelet transform [125]. A wavelet family has subclasses, often given by the number of vanishing moments it possesses (these vanishing moments are better recognized by the number aside the family of wavelets that is mentioned), and this number is related to the number of wavelet coefficients [128].

While Fourier establishes a unique base function for all applications, in wavelets, it is decided which mother function or orthogonal basis will be used for a specific application [125]. Although this would suppose an infinite or extensive set of different wavelets for the same application, we decide to choose those who comply with the desired conditions for multi-resolution analysis (MRA). More detailed information about why wavelet analysis is effective is referred in [125] and more detailed information about the advantages of wavelets can be referred in [125,128].
Wavelet transforms are mainly divided into CWT (continuous wavelet transform) and DWT (discrete wavelet transform). The discrete analogous of WT arises in the multi-resolution analysis context, which pretends the calculation of approximation and detail coefficients using a tree structure, called the Mallat algorithm. This algorithm bears considerable resemblance with a high and low pass filter bank.
Almost every useful wavelet system satisfies the MRA requirements [125]. However, there exist mother wavelets which are not computationally feasible and therefore implementable. The multi-resolution analysis allows analyzing the signal in different sized time and frequency windows. The translation of the wavelet allows localizing time events, while the scaling of the wavelet gives a detail representation or resolution [125].

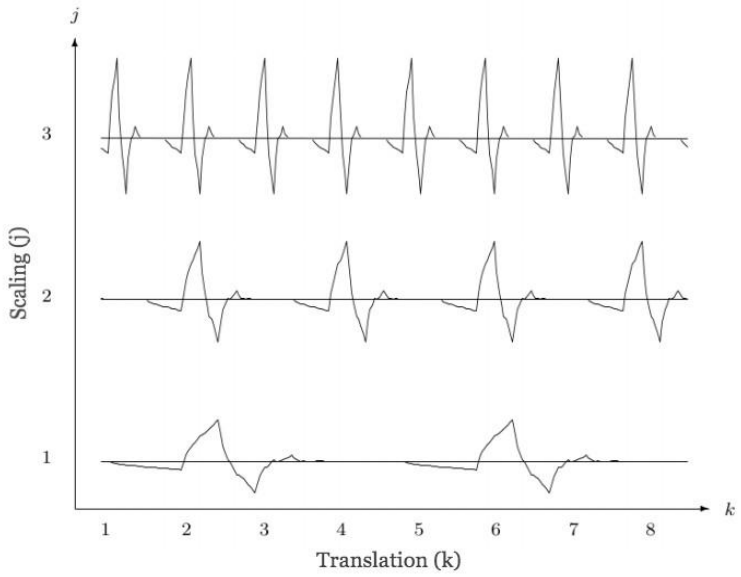



**Figure 4.3. Translation and scaling processes of a mother wavelet *[125]*.**

Not all of the wavelets have the same shape. The shape can change as the wavelet system is designed, and it depends on the desired application. In 1986 Meyer and Mallat demonstrated that orthonormal wavelet decomposition and reconstruction could be implemented in the context of signals multi-resolution analysis [130]. For this reason, there exist desirable features for the wavelets used in this analysis, such as orthogonality and compact support. Those wavelets which accomplish these properties are the most used. Besides the Haar wavelet, there were not any other orthogonal wavelets with compact support, until Daubechies, in her work [129], generated a family of functions which complied with both conditions and also mentioned other features as regularity. Haar, Daubechies, Coiflet, and Symlet wavelets are

included within this orthogonal, compact supported, and computationally feasible wavelet set [131].

Stephane Mallat produced the Mallat algorithm in 1988. This algorithm calculates the approximation and detail coefficients successively, by decomposing the signal spectra in smaller bands each time [128]. Approximation coefficients are the ones who contain the 'smooth' or the coarse information of the signal, and the details coefficients are the ones which contain the detailed information at each decomposition level, as it can be seen from *figure 4.4.* This decomposition can be done in the required levels, having in mind that each time it goes lower in the decomposition tree, more detail is made in smaller frequency bands and less approximation or low frequency information is left.

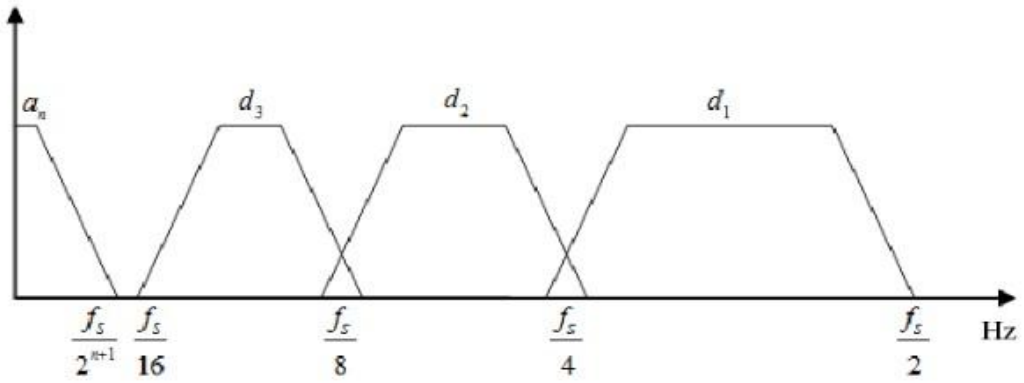



**Figure 4.4. Spectra division in the Mallat algorithm *[152]*.**

#### 4.4.1.2. Noise reduction by wavelet decomposition

The reconstruction and posterior denoising through Mallat's algorithm are done again by a filter bank, as it can be seen in *Figure 4.5.* Those coefficients which characterize the signal and avoid noise are chosen. Therefore, a differentiation of the frequency bands which are not needed is made, and then, the coefficients which do not contribute much to the signal's information are suppressed. In this way, by reconstructing the signal with the desired coefficients, also is being performed denoising [108].
The reconstruction can be made through fixed coefficients (if recognition of the frequency distribution is made) or employing thresholding. This last is a system where it is globally reviewed, which are coefficients where a higher noise concentration exists. If the noise components surpass a threshold (defined by the type of thresholding made), then the coefficient where the noise is located is suppressed [128].
In *Figure 4.5*, it can be seen the decomposition process as a filter bank system. It is desired that this filter bank satisfies the condition of the quadrature mirrored filters, which are commonly used in signal processing [149]. In other words, it is desired that the wavelet family satisfies a perfect reconstruction of the signal, although the idea of ideal filters is impossible. Under certain conditions, some wavelet families like Daubechies and others created by the same Ingrid Daubechies have demonstrated to provide a perfect reconstruction of the signal.

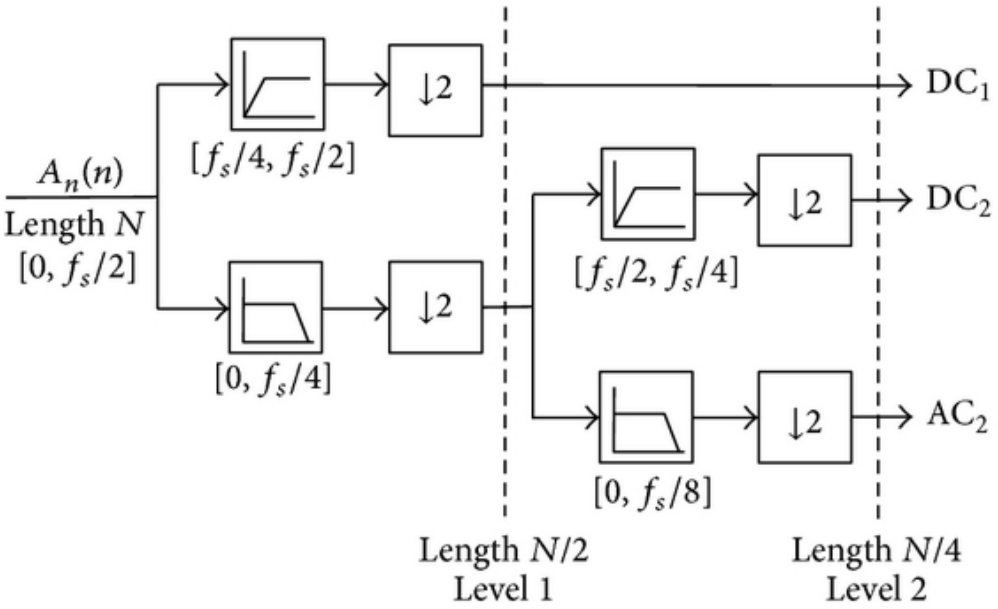



**Figure 4.5. The decomposition process is seen as a filter bank *[154]***

As stated in *subsection 4.4.1.1,* and in [140], orthogonality and compact support are features desired in wavelets for multi-resolution analysis (MRA). Also, the computational feasibility of these wavelets used for the MRA is desired, mainly because the filters used for this analysis are FIR filters. Some wavelet families are not feasible as FIR filters, but approximations of them are made with IIR filters.
The main wavelets which accomplish these three conditions are: Haar, daubechies, symlet and coiflet. These features have been extracted from MATLAB tool documentation since it is the software used for this signal processing task, by using the *waveinfo* command.

*Figure 4.6.* presents the absolute frequency of the types of wavelet used in the literature review of noise reduction approaches. It can be seen that the more used wavelets are daubechies, coiflet, and symlet, in descending order. Daubechies wavelets seem to present the best results in terms of noise reduction in all the studies which used them in the review. Also, considering the documentation available for their implementation in [140], symlet and daubechies wavelets are the preferred wavelets for denoising signals. This last consideration puts daubechies as the preferred ones, followed by symlet and coiflet wavelet families.

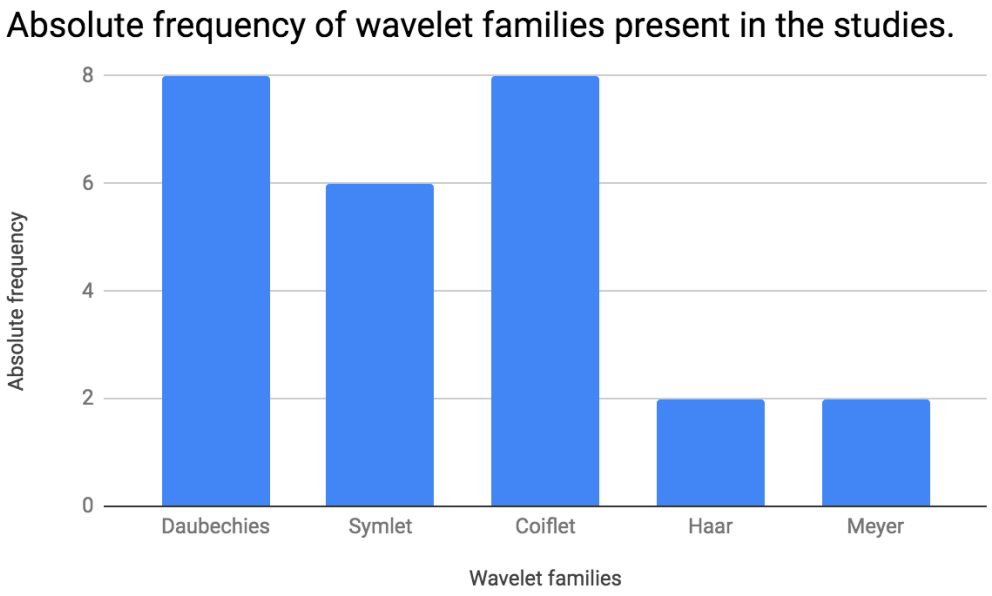



**Figure 4.6. Absolute frequency of wavelet families used in the literature review.**

Lastly, Daubechies and symlet wavelets are chosen for this analysis, although the coiflet family could have been chosen. In the following subsections, a description of the specific decomposition levels is made for each wavelet family, taking into account the results from the previous review.

#### 4.4.1.3. Daubechies wavelet family

The Daubechies wavelet family is one of the most mentioned in the review of *Table 4.1*, where it can be seen that it has been used in several studies obtaining the best results in terms of MSE, PNR, MAE, SNR, MAD, among others, varying in the study [96,100,101,105,106]. According to [108], this is the preferred artifact noise reduction technique for pulse oximetry applications. In our experiments, the first test will use the wavelet configuration presented in [96], which corresponds to db6 with five decomposition levels. Then, the next two tests are with db6 from [96] and db10 from [101,106], which demonstrated good outcomes. The decomposition levels for these two approaches will be 4, since we consider it is an excellent approach for the PPG signal frequency distribution (0 - 3.9) Hz. Mallat's algorithm structure in *Figure 4.4* was taken into account, throwing the decomposition/reconstruction levels, which are closer to the PPG frequency distribution (3 and 4).
Also, as proposed by [104], db8 mother wavelet is used with three level decomposition/reconstruction. This decomposition level is expected to provide good results as well as in the study, because the bandwidth retrieved from a three-level reconstruction is reasonably near to the one from the PPG signal. Calculations throw a bandwidth of (0 - 7.81) for this three-level decomposition, by dividing the sampling frequency of the signal between $2^{N+1}$, where N is the decomposition level.

Also, we will take the approximation coefficients for all wavelets approach, given that it is already known that the PPG signal information is more concentrated in low frequencies rather than in high frequencies.
Summarizing the approaches named, we have proposed four daubechies cases for denoising, based on the results of the literature review along with heuristic proposals, from which we expect good results. The analyzing mother wavelets possess a scaling function, which is the one that gives the approximation coefficients, and the most known wavelet function, which is the one that gives the detail coefficients. In *Figure 4.7*, an example of the wavelet functions for wavelets with different vanishing moments from the Daubechies family are shown.

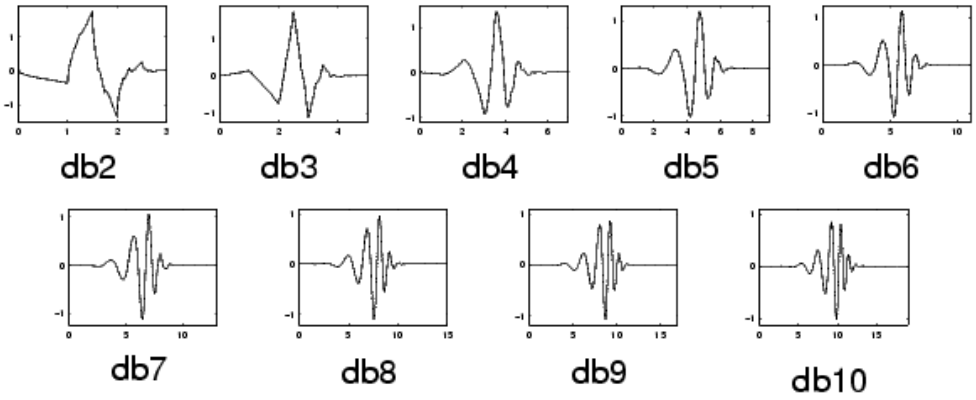



**Figure 4.7. Wavelet functions from the Daubechies family *[131].***

#### *4.4.1.4 Symlet wavelet family.*

Symlet family wavelets own their name because of "symmetrical wavelet," even when they are almost symmetrical. They were proposed by Ingrid Daubechies, as a modification to the daubechies wavelet family. In [96] it has used a wide range of symlet mother wavelets, from sym1 to sym5, also in [105] and [101], they use sym3 and sym8, respectively. The decomposition/reconstruction level calculation mentioned above is used in both of the approaches for this wavelet. In this way, sym4 as in [96] with three levels decomposition/reconstruction will be used. In addition, we will test sym6 with three levels of decomposition since its fixed-bandwidth in the frequency domain sets an appropriate interval for retrieving PPG signal information. In the *Figure 4.8* below, the wavelet functions for the symlet family with different vanishing moments are shown.

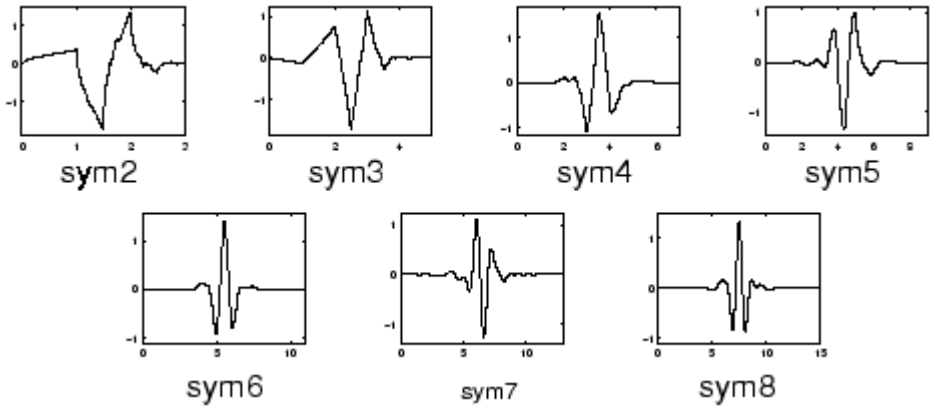



**Figure 4.8. Wavelet functions from Symlet family *[131].***

### 4.4.2. Noise reduction by smoothing filters

As it was stated in the review from *Table 4.1.*, more specifically in [99], a comparison of wavelet denoising technique is made against other techniques, namely LMS, NLMS and RLS and Savitzky-Golay filtering. Despite the results obtained, which favored the wavelets (as expected), other approaches like Savitzky Golay smoothing filters (which are inside this study) are considered.

In terms of the frequency components, a smoothing process can be seen as a low pass filtering operation. It eliminates the high-frequency components, and let pass low-frequency components with small changes [133]. The purpose of smoothing techniques roots in the fact that a signal varies slowly and simultaneously is also being affected by random noise. Thus, an estimate of a sample can be calculated from the adjacent ones (since their location do not vary greatly), reducing the noise level without biasing the value obtained [134].

Savitzky-Golay filtering method, also called as least squares smoothing filters or DISPO (Digital Smoothing Polynomial) filters; is based in a polynomial adjust over a data window. Unlike other smoothing methods as averaged windows, the task in this method is to approximate signal segments inside a window through a superior order polynomial, which is usually done from quadratic to quartic orders [133]. However, in comparison with averaged window methods, it is a little less effective for noise reduction but better at retaining the signal's original shape [133]. How big this window should be is not answered in the same way in every case, and it depends on what we want to do with the signal. The wider the window, the more points it will have for the smoothing, and therefore, it is more likely to distort the shape of the signal [133].

In other cases, smoothing filters as the median filters are more helpful in cases as spikes or outliers. Since Savitzky-Golay tries to maintain the shape of the signal, it could be more prone to keep these spikes, which is not a problem for median filters. These kinds of filters are effective if the spike is one or a few points wide (equal or less than the window size of adjacent points considered) [133]. This is the reason why we propose this filter for the task, although we know there is a wide set of smoothing techniques available for this situation.

However, median filters act poorly at preserving the edges, where the samples start to scarce, although this limitation is often more seen in digital image processing than signal processing.

In addition to this, at computational practice, the window size is truncated at the endpoints when there are not enough samples, and the median is calculated over the elements that fill that window [135].
More detailed information about this smoothing algorithm and their features can be seen at [133][134].

### 4.4.3. Noise reduction by EMD algorithms

Empirical Mode Decomposition (EMD) method is mentioned in some papers as a suitable method to recover the signal from noise components, such as in [17,102]. Other authors use this technique in combination with other techniques to enhance noise elimination, for example, combining it with DWT [136], using it for thresholding along with wavelets [137] and with smoothing methods like the Savitzky-Golay [138]. This technique along with the Hilbert Spectral Analysis (HSA) is a fundamental part of a major technique called Hilbert-Huang transform (HHT). EMD decomposes the signal in a finite and usually small set of several signals (denominated IMF), which range from the finest temporal scales (high frequencies) to the coarsest temporal scales (low frequencies) [138], through an iterative process called sifting [137]. This approach makes it comparable to methods such as Fourier transform or a Wavelet transform.
It has highlighted its independence of using fixed-functions based on a base function (as in wavelets or Fourier), because EMD is able to reconstruct a signal from the IMFs, rebuilding only those that do not have noisy content, through a comparison process. If the energy of a noise-only IMF is known, this is compared to the energy of the IMF obtained from a signal that contains both noise and valuable information. Then, a comparison is made, looking at the discrepancy between these energy values. If it has retrieved a high value, the IMF can be maintained, in the opposite case, it is discarded [137].
First implementations of this method were proposed by Flandrin et al. and Wu et al. [137], and ever since then, various improvements have been proposed. However, EMD lacks mathematical theory, and up until now, it is only described by algorithms, besides the fact that HHT[6] (Hilbert-Huang transform) theoretical bases are purely empirical, as Huang in [139] emphasized it. Another limitation are mode mixing problems and signal end effect presented when calculating HHT. These could be mitigated by EEMD (Ensemble Empirical Mode Decomposition), but this technique will not be referred to in this work, because it hasn't been encountered a resource which can provide with an implementation of this ensembled technique.

### 4.4.4. Conclusions

The state of art review from *subsection 4.3* shows a variety of techniques which are applicable for artifact noise elimination in PPG signals. We have selected seven wavelet-based techniques, one EMD and two smoothing filter techniques.

[6] Hilbert-Huang transform refers to the combined method of empirical mode decomposition and Hilbert spectral analysis. This issue has only been empirically tested and validated.

The wavelet approaches include the daubechies family with vanishing moments: 6, 8, and 10 with the fixed recomposition methods found in the literature. Symlets at orders: 4 and 6 with fixed recomposition (one of them considered in the review) and also a thresholding technique. This thresholding will be implemented over all the wavelets mentioned previously, because it is also wanted to analyze which one of the reconstruction methods fits in a best way this application. The thresholding denoising which gives the best results over all cases will be the one which will be compared with the fixed reconstruction methods. Thresholding method will be "soft" since it is the most used of all articles from the review in *Table 4.1.* Regarding the EMD method, an open source script will be used, which was previously analyzed. Thus the functioning and sifting method are understood. The two smoothing filters to be implemented are Savitzky-Golay and median filters, which are tested with different windows, calculated experimentally.
All of these methods will be accompanied by a detrending process that will help to reduce low-frequency artifacts, as some of these methods deal with high-frequency noise only. Then, the integrated implementation of these two techniques will conform to each denoiser method. More information of the implementation of these denoising methods can be seen in *Appendix D.*
Positive results in the improvement of brady- and tachycardia detection are expected, through the comparison of these ten techniques. The evaluation of the methods performance will be part of the evaluation of the accuracy of arrhythmia detection classifiers described in the next section.

## 4.5. REVIEW OF TACHYCARDIA AND BRADYCARDIA DETECTION MECHANISMS

### 4.5.1. Review methodology

The review performed for choosing the algorithms that will be used for the detection of brady- and tachycardia is based on the results obtained by the Physionet/CinC Challenge scoring system. These are currently published and available in the Computers in Cardiology journal.
We made this review to find the most suitable arrhythmia detection algorithms for their further integration with the selected denoise techniques proposed in the last section.

#### 4.5.1.1. Selection criteria

In this short review, various selection criteria were taken into account and later used for choosing the arrhythmia detection algorithms which are adapted with denoisers. As the first and most important criterion, we considered the fact that these algorithms use the PPG or ABP signals (either uniquely or mainly) in arrhythmia detection. If the arrhythmia detection is only based on the ECG signal, or the PPG signal is used as the last classification resource, the classifier is discarded. This because we intend to improve the classification in PPG signals contaminated with artificial movement noise added to the PPG signals. Therefore, if the classifier only uses the ECG signal, we do not expect an effect of the denoising process.
As the second criterion, we take into account that the developed algorithms try to stick as much as possible to the base structure that was provided for the competitors because these set of

scripts suggest from the beginning an arrhythmia detection based on pulsatile waveforms resampled at the same sample frequency as the generated noise. The competitors were given total freedom to change scripts, create new functions, or make everything from scratch. For this reason, we look forward to use scripts with a similar structure as the base one, which will be one of the tests (the simplest case of arrhythmia detection).
Finally, some of the classification methods include machine learning algorithms, as specified by competitors in their corresponding papers. It should be mentioned that these algorithms sometimes are easy to understand in theory, but complex to understand in the code. For this reason, if it is the case, we will discard some of them because of the complexity of the functions used, since what we want is to understand the steps followed by them rather than blindly execute processes without knowing their internal operation.

#### 4.5.1.2. Data analysis

Nineteen contributions were the outcomes from this competition. The 16 articles shown in *Table 4.2* were selected after applying inclusion and exclusion criteria.

| Contribution | Signal used | Techniques and performance parameters or main features | Gaps |
|---|---|---|---|
| Decreasing the False Alarm Rate of Arrhythmias in Intensive Care Using a Machine Learning Approach [109] | PPG/ABP and ECG | -Random Forest classifiers were trained separately for every type of main arrhythmia feature. Five predictive models were created.<br>- Downsampled to 125 Hz.<br>- Uses 500 data from the training set, unannotated during the challenge<br>- Reported classification sensitivity is 75–99 % (average 93 %) on the training set with cross-validation and 22–100 %(average 90%) on the unrevealed test set.<br>- Reported classification specificity on the training and test set were 76–94% (average 80%) and 75–100% (average 82%), respectively. | Even though both the ECG and ABP/PPG are used, the approach already included a very robust algorithm, were many denoising processes are made along the machine learning process. This issue makes the identification of the place for including the denoising process very problematic.<br>The quality attributes selection may cause that many signals are discarded, specially ABP/PPG (if these are noised), leaving as the only signal the ECG. Moreover, one single computation of the main loop lasted at least 2 minutes, becoming then a low feasible method for testing every denoising case. |
| Enhancing the Accuracy of Arrhythmia Classification by Combining Logical and | PPG/ABP and ECG | - Combination of logical and SVM-based machine learning techniques.<br>- Sample frequency: 250 Hz<br>- The feature set includes R-R intervals, power spectrum density, autocorrelation plot, and standard deviation values. | -Fs is not specified.<br>-The primary disadvantage of the methods can be attributed to the non-usage of PPG/ABP signals for VT (ventricular tachycardia) and |

| | | | |
|---|---|---|---|
| Machine Learning Techniques [110] | | -They make signal preprocessing for subtracting low-frequency noise; it means baseline drift (bidirectional second order Butterworth filter)<br>- If any of the records had flat-lines or zigzag, then it was discarded. | VF (ventricular flutter).<br>-Sensitivity is not specified.<br>- As they consider the noise as a limitation for VT prediction, they use more methods over ECG signals, which could be a limitation. |
| Reducing False Arrhythmia Alarms Using Robust Interval Estimation and Machine Learning [111] | PPG/ABP and ECG | Classifiers include:<br>Binary classification trees (BCTs)<br>-Discriminant analysis classifiers (DACs)<br>-Support vector machines (SVMs)<br>-Linear and pseudo quadratic discriminant analysis classifiers.<br>-Preprocessing by resampling to 100 Hz, bandpass filtering with 1-30 Hz and normalization to zero mean and unit standard deviation is performed but not explained. | - Sampling frequency is not compatible with the one used in our noise.<br>- No quality indexes are used, or at least not those they mention. |
| Reduction of False Cardiac Arrhythmia Alarms through the use of Machine Learning Techniques [112] | PPG/ABP | This article intends to reduce the number of false arrhythmia alarms implementing decision trees, which have been shown to generate consistent results in reducing false cardiac arrhythmia alarms, because of its robustness to noise in the data and because its low computational complexity.<br>- Downsampled to 125 Hz.<br>-Sensitivity: real-time event (79%); retrospective event (84%)<br>-Specificity: real-time event (84%); retrospective event (89%) | - This study uses only PPG/ABP signals, which makes it suitable for our analysis.<br>-Alternative, machine learning techniques were considered for future works: neural networks, fuzzy logic, and support vector machines. |
| Reliability of Clinical Alarm Detection in Intensive Care Units [113] | PPG/ABP and ECG | - Analysis techniques and main features: High-pass filtering to remove baseline instability, scaling to normalize waveform amplitudes, detection of noisy and flat waveforms, differentiation to accentuate sharp waveform edges, beat detection, the timing between beats preceding alarm onset, and detection of alarm conditions<br>-Sensitivities: training (89%); Testing (88%).<br>-Specificities: training (38%); Testing (38%)<br>ECG data alone used for training can obtain an average sensitivity of 88%. After adding the arterial blood pressure and peripheral pulse:<br>-Sensibilities: Training (91%); Testing (84%).<br>-Especificities: Training (25%); Testing (29%) | - They first considered using algorithm without PPG/ABP data, showing a noticeable decrement of the performance parameters after adding these signals; for this reason, they consider using only ECG.<br>- Sensitivity improved (2% only in training) by including arterial blood pressure and pulse data, this improvement was offset by a reduction in specificity (of 13% in training and 9% in testing). Since this additional data creates new false positives; authors highlight the need for finding new analysis techniques to improve performance |

| | | | |
|---|---|---|---|
| | | | parameters. |
| Reducing False Arrhythmia Alarms in the ICU by Hilbert QRS detection [114] | ABP/PPG and ECG | -Removes high and low-frequency components by using two median filters for ECG channels. Missing values were replaced with zeros.<br>-Detects D-waves QRS complex using Hilbert filter from ECG channel.<br>-Signal Quality Index (SQI) was calculated from ABP and PPG signals<br>-Template matching was used to detect ventricular tachycardia (VT) and the power spectrum of ECG signals and identifying the VF frequency components employed to investigate ventricular fibrillation.<br>- Retrospective algorithm: 98% (TPR), 66% (TNR) with a score of 74.03.<br>Real-time algorithm: 98% (TPR), 65% (TNR) and a score of 69.92%. | - The main denoising and processing techniques are used for ECG signals, which could be a limitation for our work since ECG clean signal improves, and it is beyond our scope.<br>- The noise removal procedure is operated for QRS complex cleaning and detection in ECG channels using filters. |
| A Multimodal Approach to Reduce False Arrhythmia Alarms in the Intensive Care Unit [115] | ECG - ABP/PPG | - To find the heart rate from PPG-ABP, they use an OSC-ANF filter, which changes its central frequency to follow the instantaneous frequency of the signal. Then, this algorithm is extended to multiple input signals, and again, it is extended to the complex domain.<br>- Denoising is made to ABP/PPG signals at the beginning (sum of noise after this filter or remove filter), with eighth-grade lowpass Butterworth filter of cut frequency 5 Hz.<br>- In order to reach a more robust detection, algorithms are fed with smoothed versions of ABP/PPG signals by means of moving averaging.<br>-Sensitivities: real-time (94%) - retrospective (99%)<br>-Specificities: real-time (77%) - retrospective (80%) | - Resampling to 5 Hz makes it incompatible with our noise approximation. |
| Identification of ECG Signal Pattern Changes to Reduce the Incidence of Ventricular Tachycardia False Alarms [116] | ABP/PPG and EGC | - They obtain SQI for ABP/PPG, and it is compared with the threshold, which is different for each type of alarm.<br>- They make the detection with SQI because any other machine learning method is mentioned. It seems that the principal signal processing has been given to ECG signal in VT, with the five proposed hypotheses test.<br>- However, the algorithm used an additional threshold for the QRS complex shape feature, based on standard deviation.<br>- No testing information is provided, just for training, where: sensitivity: 92.5%, specificity: 74.5 % | - Even when the title suggests the opposite, the ECG signal is used to detect ventricular tachycardia, the other four arrhythmias are detected through PPG/ABP.<br>- However, by not having various methods for pulsatile signals, except for the thresholds, it is seen as something very similar to the base code. |

| | | | |
|---|---|---|---|
| Multi-modal Integrated Approach towards Reducing False Arrhythmia Alarms During Continuous Patient Monitoring [117] | ABP/PPG and EGC | - Five arrhythmias are treated as independent problems, so they use five subroutines for each one of them. Each subroutine has two main steps: peak detection and alarm verification.<br>- Signals are resampled to 125 Hz.<br>- For PPG: wabp routine from Physionet is used. If it is a PPG signal, it is scaled before introducing it to the routine.<br>- Quality coefficients from PPG and ECG are calculated.<br>-Sensitivities: real-time (89%); retrospective (89%)<br>- Specificities: real-time (84%); retrospective (87%) | - They use ABP/PPG signals in a good way in arrhythmia detection, even in asystole, where the signals are supposed to have the worst quality. |
| Multimodal Data Classification using Signal Quality Indices and Empirical Similarity-Based Reasoning [118] | ABP/PPG and ECG | - They do not make a preprocessing further than what it has provided in the base script. However, they obtain an SQI for each signal.<br>- Five additional features are retrieved, based on the heart rate.<br>- A correlation of which register belongs to which arrhythmia is made using the SQI and additional features. If the similarity increases more than a determined threshold, the record is taken as the arrhythmia case. | - No mention of the used algorithms for peaks extraction is made. Neither do they mention what type of signals they use to extract them. It could be inferred that they use ECG because of the reference to RR intervals, but this assertion cannot be made with total certainty in such an ambiguous case. For this reason, it is discarded, even though it is a good, novel, and different method. |
| Reducing False Arrhythmia Alarms in the ICU using Novel Signal Quality Indices Assessment Method [119] | ABP/PPG and ECG | - Considering the denoising result and time, they selected wavelet transform method to deal with the problem.<br>- After filtering the signals with the wavelet transform method, a novel method of signal quality assessment was developed based on modifying Townsend and Tarrasenko's methods to fuse signal quality indices of different types of data from multiple sensors.<br>- SQI from both ECG and ABP/PPG signals was retrieved: for ECG signals it was used K-means algorithm. For PPG/ABP signals it was used a combination of two algorithms: a beat-by-beat fuzzy logic-based assessment of features in the ABP waveform and heuristic constraints of each ABP pulse to determine normality.<br>- -Sensitivities: real-time (65%) - retrospective (65%)<br>-Specificities: real-time (82%) - retrospective (87%) | - Only ECG signal was denoised with wavelet decomposition (coif4 wavelet, with eight decompositions)<br>- Approaches such as fuzzy logic-based assessment are not within our knowledge, and for this reason, we decide it is better not to use it. |

| | | | |
|---|---|---|---|
| Reducing False Arrhythmia Alarms in the ICU [120] | ABP/PPG and ECG | They make a quality assessment of all signals since some of the channels might be damaged by artifacts. This assessment is made in the last 30 second of recording, for real-time line and detection of heartbeats will be performed using only signals that pass this quality check.<br>- Heartbeat annotation is made using multiple methods: for only ECG (I and II): gqrs algorithm. For only ABP/PPG except in VT case: wabp algorithm. For combinations of both ECG leads and ABP/PPG: they used their in-house algorithm. | - For VT they used ECG only because it seems to be an illness just noticeable by the changes of the QRS complex.<br>- Since the signals used for the real-time line are sometimes discarded because of the quality, it is quite possible that none of the noise contaminated signals can pass over this condition, making the results null. |
| Reduction of False Alarms in Intensive Care Unit using Multi-feature Fusion Method [121] | ABP/PPG and ECG | - A 5-40 Hz band-pass filter was used for filtering the ECG signals and 5-35 Hz band-pass filter for ABP/PPG signals.<br>- For ABP and PPG signals, pulse foot detection was performed using the wabp function<br>- Then some features over these IBI and IPI detections were retrieved, with different windows sizes. | - Different rules have been set for VT, which seems to be the most complicated arrhythmias overall.<br>- Although they use ECG signals, also count with PPG to give the scoring. |
| Algorithm for life-threatening arrhythmias detection with reduced false alarms [122] | Mainly ECG and then ABP/PPG. | -Resampled to 250 Hz.<br>- A more rigorous treatment is made to ECG signals, while in the ABP/PPG signal, just offset is taken off.<br>- Offset is retrieved from ECG signals, as they would share the same offset.<br>- They do not use machine learning techniques, just peak calculation through slopes. In ABP/PPG case, they do it with the reverted signal, to find the negative change of slope.<br>- Also, low-frequency noise is removed with zero phased eighth grade Butterworth filter with 40 Hz cut frequency.<br>- The non-causal median filter is applied to ECG (250 ms window) and ABP/PPG (350 ms window) to eliminate offset. | - The central gap noticed is the use of ABP/PPG signals as an alternative and not as the primary tool for arrhythmia detection.<br>- Sampling frequency does not match with the one of our noises. |
| False Alarms in Intensive Care Unit Monitors: Detection of Life-threatening Arrhythmias Using Elementary Algebra [123] | ECG | Data channels are first independently searched for invalid blocks and QRS complexes.<br>Using the QRS distribution and derived R-R information, each channel in the record is tested for regular heart activity. If any of the channels pass this test, a<br>false alarm is reported and the process ends. If this is not the case, a specific arrhythmia test is executed | -They do not use PPG signals as the primary tool for arrhythmia detection. |

| | | | |
|---|---|---|---|
| Heartbeat Fusion Algorithm to reduce false alarms for arrhythmias [124] | ECG | - They developed an algorithm based on global heartbeat annotations generated by fusing individual heartbeat detections from multiple physiological signals<br>- Support vector machines (SVM) with radial basis function kernels to classify signals as "noisy" or "clean" were trained for each signal type using libSVM | - The central gap is the consideration of ECG signals as the main referent, having ABP/PPG signals as an alternative only if ECG signals do not work. This structure is followed only for 3 of arrhythmias and for the others, it is strictly used ECG. |

**Table 4.3. Summary of analyzed works for arrhythmia detection techniques.**

#### 4.5.1.3. Summary of the reviewed algorithms

The Physionet/CinC Challenge 2015 present different results regarding classification techniques for brady- and tachycardia. All of the articles presented in the table are the best 16 scores obtained at reducing false alarms in ICU.

Approaches like [110,111,115,122] contain varied techniques. However, the incompatibility of the sampling frequency used by them and the one from the generated noise in *Chapter 3*, puts them out of the frame. If we use upsampling processes, these would generate a change in both spectral and time features of the noise. The same would occur if we use downsampling techniques over the pulsatile signals.

PPG signal is not used in articles [113,123] where they use only ECG, and [114,122,124] do not use pulsatile signals as a primary reference to the classification, but as the least resource instance, in case ECG does not work. Thus, these ECG signals would still be clean as the pulsatile ones would be easily discarded or even never used because of its noise condition.

In [118] the purpose of PPG signal for arrhythmia detection is not clarified, making the detection ambiguous in the sense that it is not known if the ECG signal or the PPG signal is being used. However, it is highlighted the usage of a method different from all the others.

Approaches like [109] and [119] are discarded because of the complexity of the used techniques, which surpass our knowledge as far as it goes. Also in [109], the computational time required was considerably long. The usage of the signal quality indices only in ECG signals can be seen in [120]. Therefore, the review of just these signals makes it indifferent to what happens in noise corrupted PPG signals. Lastly, algorithms like [116] use the basic decision threshold techniques which have been already established by the challenge organizations on the primary provided material. Since the base case will be one of the used for comparing denoisers, it becomes unnecessary to review another one which contains the same techniques without adding any additional feature over this same basic structure.

It is finally decided to choose articles [112,117, 121] and the base case, which is briefly explained in [132]. In the next section, a short explanation of the techniques used in the different chosen papers is made.

#### 4.5.1.4. Tachycardia and bradycardia detection techniques

The classification techniques in the three chosen scripts are similar to the one in the base script. Each one included improvements, always considering that tachycardia and bradycardia are

diseases that can be detected through a correct identification of the heartbeats. According to the conditions given by the challenge, an alarm is called as ‘Tachycardia' when the heart rate is greater than 140 bpm during 17 consecutive beats. A ‘Bradycardia' alarm can be identified by a heart rate lower than 40 bpm during five consecutive beats [81]. We presume that these conditions imposed by the challenge organizers are at rest, since the original dataset was taken from ICU patients. An additional search of the medical definition of these arrhythmias also puts them at rest conditions for both [33,35]. However, as it has been stated by Bonomi, even wrist movements at rest might make changes in the quality of the classification.
Even when all scripts use different techniques, classification criteria for each disease must follow the guidelines previously mentioned. In the Base script, an onset pulse detection identifies the starting point for each PPG/ABP signal to separate each waveform. For calculating the peak location within each period, they use thresholding techniques. Additionally, a quality index analysis over the pulsatile signal is carried, to see if the measures are acceptable.
In Miguel Caballero’s script [112], a simple decision tree is implemented through *if-else* statements. This decision tree increases the granularity in the decision making, especially for ventricular tachycardia and ventricular flutter, those which are not our primary focus. However, new thresholds for bradycardia decision making are implemented.
In the scripts made by Chengyu [121] and Ansari [112], a multi-feature analysis is performed for ECG and ABP/PPG signals, which generates a highly robust algorithm, as it can be seen in the scoring results. The more variables exist for classification, the more precise it will be. Given that these two participants use ECG and PPG signals in an equal proportion, an analysis of the improvement of the results can be made by changing the denoising method only for the pulsatile signals. As the contamination for ECG signals is not within our scope, the filters for these signals are removed and then put again, combined with the denoising proposes over pulsatile signals.

### 4.5.2. Structure of the dataset

Physionet Challenge 2015 supplies a total of 1250 recordings. These were obtained from four hospitals in the USA and Europe, each containing an arrhythmia alarm**.** The training set consisted of 750 training datasets, and the testing sets consisted of 500 recordings. The testing set is hidden from participants for scoring purposes and is currently unavailable, then the results obtained will be the ones from the training set. A team of expert annotators reviewed each alarm and labeled it either 'true', 'false', or 'impossible to tell'. However, the challenge included only records that were reviewed by at least two annotators, of whom a two-thirds majority agreed if the alarm was either true or false [89].
All signals had a sampling frequency of 250 Hz and had been filtered with an FIR band-pass filter with a passband from [0.05–40] Hz and powerline notch filters to remove noise. Each recording contains two ECG leads and one or more pulsatile waveforms (the photoplethysmogram PPG and arterial blood pressure waveform ABP). Pulsatile channels can suffer from slight movement artifact, sensor disconnects, and other events (such as line flushes or coagulation in the catheter) [95]. Overall the records from this channel, the most common

situation is the sensor disconnect, as we analyzed the records correspondent to brady-tachycardia. Some of the authors within the selected detection algorithms deal with these problems, but some of them leave the records as they come from the dataset. We try to stick as much as possible to the considerations taken by the authors of these codes, and it is for this reason that we will not edit any of the contributions of the challenge beyond the addition of noise and implementation of the denoisers. In this way, we can suggest that some of the scripts might be prone to misclassification because of these details.
Event organizers idea was to allow the participants to edit a simple base algorithm which checks the quality of the ABP waveform, and (if it is acceptable) uses it to calculate the heart rate. If no ABP is available, then the same process is performed on the PPG waveform. If the specific heart rate calculated for each arrhythmia is outside the range indicated by the provided guidelines (considering a tolerance level), then the alarm is suppressed and marked as false.
Participants could use this approach as the base to more sophisticated techniques, as organizers recommended it, or could make their own from scratch.

## 4.6. EVALUATION OF THE MECHANISM

This section exposes the steps to complete the evaluation process for the proposed mechanism. These steps are taken from the “Experimentation in software engineering” book [155] and can be seen in *Figure 4.9*. First, the definition of the evaluation is exposed. Then the planning and design of the mechanism evaluation is stated. Here, we mention the selected evaluation method, variables, and criteria in the arrhythmia classification context. We follow with the operation stage and lastly, the analysis and interpretation stage. Finally, the results are explained through analysis and the conclusions in the last subsection.

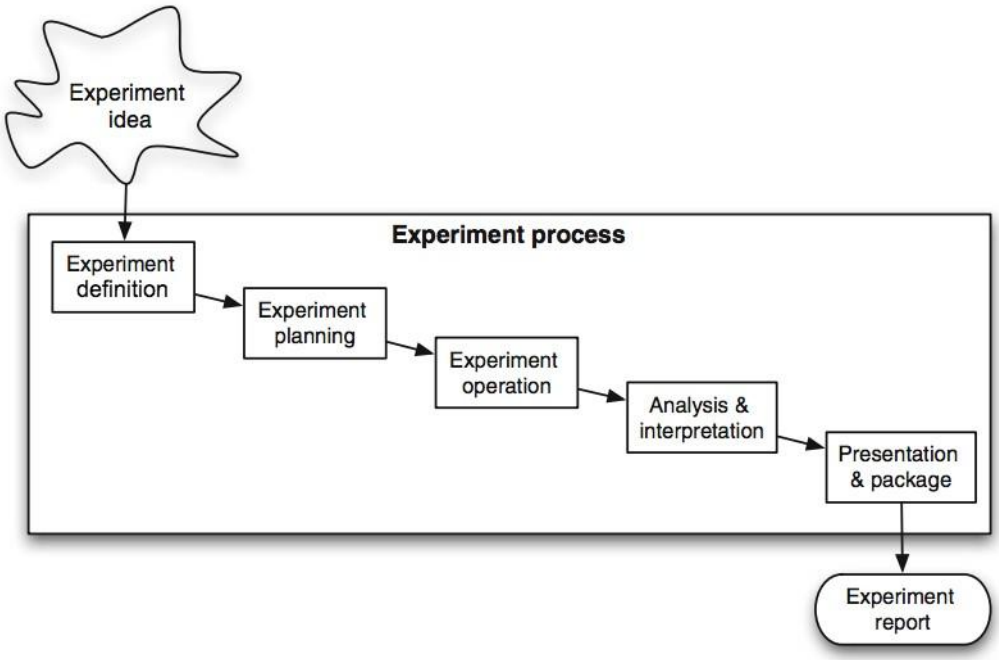



**Figure 4.9. Steps to follow in the proposed evaluation process *[141]*.**

### 4.6.1. Definition of the evaluation

This evaluation purpose is to answer the following question and hypothesis:

Can state-of-the-art movement artifact denoisers improve the performance of state-of-the-art bradycardia and tachycardia classification algorithms based on PPG/APB signals?

H0: The use of Wavelet, EMD or smoothing filter-based denoising methods don't affect the specificity, sensitivity or ROC measures of state-of-the-art Bradycardia and Tachycardia classification algorithms based on PPG/APB signals tested on a simulated noise to the PPG signals obtained from the Physionet Arrhythmia database

H1: The use of Wavelet, EMD or smoothing filter-based denoising methods don't improve the specificity, sensitivity or ROC measures of state-of-the-art Bradycardia and Tachycardia classification algorithms based on PPG/APB signals tested on a simulated noise to the PPG signals obtained from the Physionet Arrhythmia database

The following section describes the experiment in order to test the hypothesis.

### 4.6.2. Design of the evaluation
#### 4.6.2.1. Selection of the evaluation method

This evaluation is presented as two separate results, the first with the purpose of showing the distortion that the noise induces into the behavior of the current classifiers which use PPG/ABP signals. This first result is purely descriptive since the hypothesis is found later in the second result. The second with the purpose of showing which denoising method induces the best possible improvement over the results of the challenge scoring.
To accomplish the third specific objective of this undergraduate thesis work, a pre- and post evaluation is established. In the hypothesis testing experiment, the pre-evaluation refers to the classification results without the application of a denoising process, and the post-evaluation refers to the classification results with the application of the denoising process. Furthermore, every time the mechanism is run, the next variables are calculated: sensitivity, specificity, precision and F1 score. From the first two variables we can obtain also a new graphic variable called ROC curve and the AUC. A quantitative type evaluation is then presented. The evaluated variables are exposed in the next subsection.

#### 4.6.2.2. Selection of variables to evaluate

These variables will assess the effectiveness of each classification script. As the outcomes of the challenge are TP, FP, TN, and FN, it is possible to compute the sensitivity, specificity, among others, for each arrhythmia on each script. Additionally, other variables as precision and F-measure are obtained to make better classifier discrimination for obtaining a better performance [143,144,145]. More specifically, as the primary arrhythmias for this work are

bradycardia and tachycardia, the variables previously mentioned will be calculated for each one of the arrhythmias. The definition of these variables is made down below:

- **Sensitivity:** it is better known as the true positive rate (TPR), sometimes is also called a recall. The test or classifier can qualify correctly the ill patients who possess the illness. In other words, it is the proportion of the patients which results are positive for an illness test or classification, among all the patients who possess the illness condition [144][145].
  It can be calculated from the following formula:

$$Sensitivity = \frac{TP}{TP + FN}$$

  Where TP stands for True Positive (a person who has the illness and is correctly classified as having the illness) and FN stands for False Negative (a person who has the illness but was incorrectly classified as not having it).
  Then the TP+FN term refers to all the population who have the illness.

- **Specificity:** it is better known as the true negative rate (TNR) or selectivity. The test or classifier can reject correctly the patients who are not ill or do not suffer from any condition. In other words, it is the proportion of the patients who do not possess the illness condition and which were correctly classified as not having it, among the totality of the patients who are not ill [144][145]. It can be calculated from the following formula:

$$Specificity = \frac{TN}{TN+FP}$$

  Where TN stands for True Negative (a person who does not have the illness and is correctly classified as not having it) and FP stands for False Positive (a person who does not have the illness but it is incorrectly classified as having the condition).
  Then the TN+FP term refers to all the population who do not have the illness.

- **Precision:** it is also called a positive predictive value (PPV). It must not be confused with accuracy, which stands for the proportion of correct results a classifier obtained over all the results. Precision is the proportion of patients who were correctly classified as having the condition, among the totality of the patients who were classified as having the condition (correctly and incorrectly) [144]. It is calculated from the following formula:

$$Precision = \frac{TP}{TP + FP}$$

  Then this measure gives us an idea of how many times the alarm was triggered with the real purpose of arrhythmia detection.

- **F-measure:** while the precision can be seen as a measure of exactness (quality of the classifier for positive prediction), the sensitivity can be seen as a measure of completeness (how many times the classifier correctly predicted a person with the condition). These two measures can be jointly analyzed, which is called as F-measure or F1 score [143]. This measure is an indicator of the balance between these two parameters, although there are other types of F measures, such as F2 or F0.5, where the first one has the weighting of the sensitivity higher than precision, and the second one has a weighting of the precision higher than sensitivity. In this occasion, the F1 measure (equal weights for both precision and sensitivity) is shown, which can be calculated from the formula:

$$F1\ score = 2 * \frac{Precision * Sensitivity}{Precision\ +\ Sensitivity}$$

  It is essential to highlight that given the case the classifier obtains a 100% sensitivity; it is taken as a trivial case because the classifier labels everything as 'true.' Oppositely, if the previous statement accomplishes, this implies a deficient result in precision, turning the classifier useless. The same will occur if we search for an excellent precision value (sensitivity will decrease).
  The conclusion over these two measures is a balance between both, which can be easily calculated with the F1 score [146].
  However, F1 score does not take into account the true negatives in the classification, and it is for this reason that the specificity is taken as well as the variable for discarding some of the results thrown by the different approaches for denoising.

- **ROC curve:** ROC (Receiver Operating Characteristics) curve will also be used. ROC is a graphics tool which is useful to visualize a classifier's performance. This curve graphs the sensitivity or TPR in the *y*-axis, and the false positive rate (FPR) in the *x*-axis at different threshold settings. The FPR or fall-out measure is equal to 1-specificity. Then this curve graphs the tradeoff between sensitivity and specificity (because any increment on sensitivity implies a decrement on specificity). A cutoff point will always exist between positive and negative results, where positive refers to ill or with an arrhythmia and negative means healthy or without arrhythmia(s) [147]. These curves are used in many ways, especially in applications with cost-sensitive learning and allow us to [145]:
  - ➔ Determine a decision threshold which minimizes error rate or misclassification cost under determined cost functions.
  - ➔ Identify regions where a classifier works better than the other.
  - ➔ Identify regions where a classifier works worse than chance.

  The idea is to determine the best threshold or cutoff point that maximizes either the sensitivity or specificity and minimizes the mistaken results such as false positives (FP) or false negatives (FN) [146]. Nevertheless, obtaining a 100% accuracy is the ideal case

and very unlikely since FP and FN are mostly present in practice. Therefore, the only thing that can be manageable is the cutoff point. If the classifier estimates a positive value following its belief, and this value is a positive value, it will increase either TPR and FPR. If this decision threshold or cutoff point is varied from its maximal to its minimal value results in a piecewise linear curve, the ROC curve is obtained such that each segment has a non-negative slope [145].
The closer this curve is from the upper-left borders, it means that the classifier is more accurate. The closer the curve is from the diagonal line (chance line) drawn at 45 degrees from (0,0) to (1,1) coordinates, it means that the classifier is less accurate. Anything that passes under this diagonal is considered as inaccurate or wrong at classification and finally, any point on the diagonal means that the classification follows the chance criteria [146].

- **AUC:** the AUC (Area Under the Curve) is the measure of the classifier's accuracy. When this area is equal to 1, it represents a complete classification. There exists a rough guideline to classify the results of accuracy for the AUC [146], which is shown as follows:

| AUC value | Classification of the result | AUC value | Classification of the result |
|---|---|---|---|
| 0.9 - 1 | Excellent | 0.6 - 0.7 | Poor |
| 0.8 - 0.9 | Good | 0.5 - 0.6 | Fail |
| 0.7 - 0.8 | Fair | | |

**Table 4.4. Classification criteria of AUC values for a classifier.**

AUC can also be seen as a discrimination measure, in other words, is the classifier's ability for distinguishing those positive results (people who have the disease) from the negative results (people who do not have the disease) [146].

The graphing procedure for this curve can be done in many ways. Thus, three methods are retrieved from various tools found at the documentation center of MATLAB software, these are explained in more detail in *Appendix F*, adjoint with the results for each graphing method.

#### 4.6.2.3. Evaluation criteria

Evaluation criteria for this kind of evaluation are as follows:

- The impact that generates the use of the proposed denoisers over each classification script. This result means the improvement that each denoiser introduces over the classifiers compared with the performance of the classifiers under noise conditions (for bradycardia and tachycardia) using the modeled noise.

- As a secondary evaluation, we perform a comparison between the performance of the denoised classification scripts against the original results for each script. In other words, a comparison between the results obtained with the proposed denoising method and the results of the unchanged scripts from the Physionet 2015 Challenge. This is done because the original results are considered without artificial noise models or under any motion artifact conditions. Therefore, they represent the 'ceiling' results that can be used as a reference to test our mechanism performance.
- The mechanism (classifier with its corresponding denoising method) that throws the best results regarding the evaluation variables named previously for both arrhythmias detection.
- As an additional result, although not so important, the results of Bonomi will be shown with respect to those obtained with the classifiers. Because of the differences in the evaluation method from Bonomi and our case, it is not as significant to compare ourselves with him.

### 4.6.3. Operation

#### 4.6.3.1. Procedures before the operation

After the evaluation variables selection and the evaluation criteria statement, several procedures are executed before the operation. These are detailed in *Appendix D*. For each different mechanism (different classifiers joined several denoisers), the following steps are made:

- Choose a range from the Dynamic Variance Moving Average model that generates a deviation over the original results of the classifiers. These deviations are computed by removing the filters that these classifiers possess internally, because we want to simulate Bonomi's scenario. Then a classification is executed using them, without any type of preprocessing. This step is made only once for all of the scripts, because it throws the noise models for contaminating all the records from the arrhythmia database.
- Introduce the chosen noise models in each of these classification scripts and obtain the results with noise.
- Introduce then the set of denoisers (one by one) in each script and obtain the results for the impact of each one of them over the different classifiers.
- Choose the best denoising case for each classification script.
- Compare the final results of the best cases for each classifier with the corresponding original results and noise results.
- Compare the final results of the best cases for each classification script with the ones in the other scripts.
- Select then the best classification mechanism.

We expect that the results serve as a proposal of a mechanism that supports bradycardia and tachycardia detection under low intensity movement artifact conditions.

### 4.6.4. Results and analysis

#### 4.6.4.1. Results after the addition of noise: simulation of Bonomi's scenario

As it was mentioned in the previous section, the noise addition over the classification results should induce some change in the results, specially if the proper filters used by each script are removed. The filters used by each work are explained in *Appendix D.* They are reduced to simple notch low-pass filtering (for the base script and Miguel Caballero) and Butterworth passband filtering (for Chengyu and Sardar scripts). Results are taken from the original case, in other words, the scripts are run as they come from the Physionet page repository. These results should be the same obtained from the authors. Unfortunately, these cannot be revised because they only mention testing results in their work, so the original results for training are trusted.

First, a graphical analysis of how much the low-intensity noise can damage the classification results is shown. As this noise is a low intensity one, it cannot be expected that this makes the signal unrecognizable or extremely damaged, thus making the SNR analysis far inconvenient for this problem. Then, an analysis for the variables which determine the classification is made. Specifically, for bradycardia and tachycardia, two variables are calculated respectively: low heart rate for bradycardia and high heart rate for tachycardia. A signal quality index is also calculated for both of the arrhythmias to execute the validation of the alarm.

The following if-else statements (translated into a logical gate schema) are made for determining the existence of brady- or tachycardia in the base script:

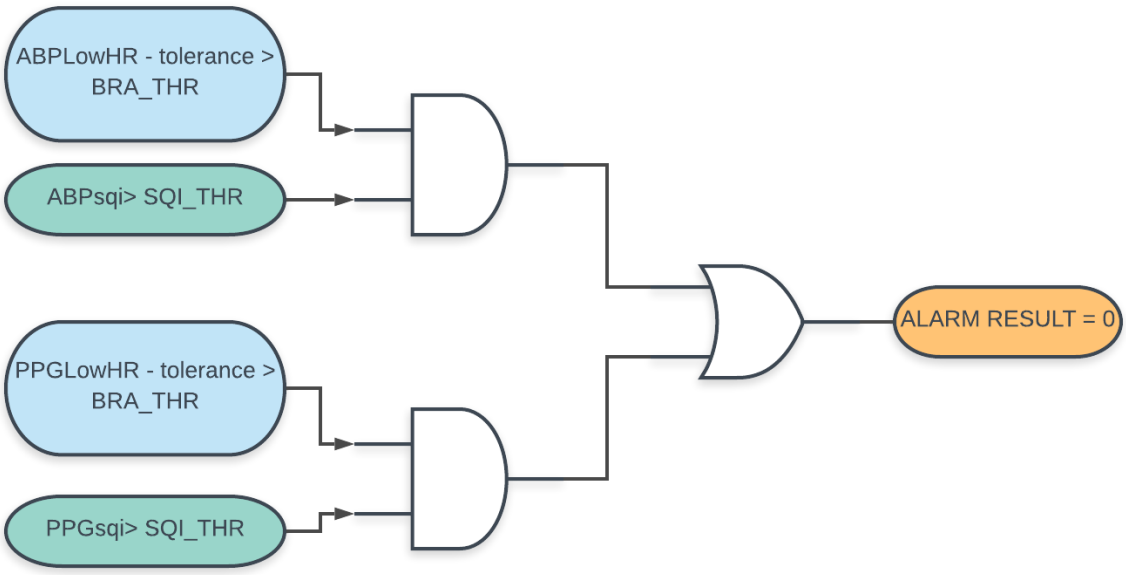



**Figure 4.10. Gate schema for the bradycardia determination block in base script.**

The two conditions are evaluated by a logical conjunction operator AND, so any of the results which turn their logical state to zero will make impossible the alarm off. In most of the cases, there exists just one of the pulsatile signals, making the schema active in just one of the sides.

Tolerance and bradycardia thresholds (BRA_THR) for the determination have already been established in the challenge statement, however code could be changed for the convenience of some competitors, as we will see later. The schema for tachycardia is similar to the previous one, but the condition is side changed, as it can be seen in the following figure.

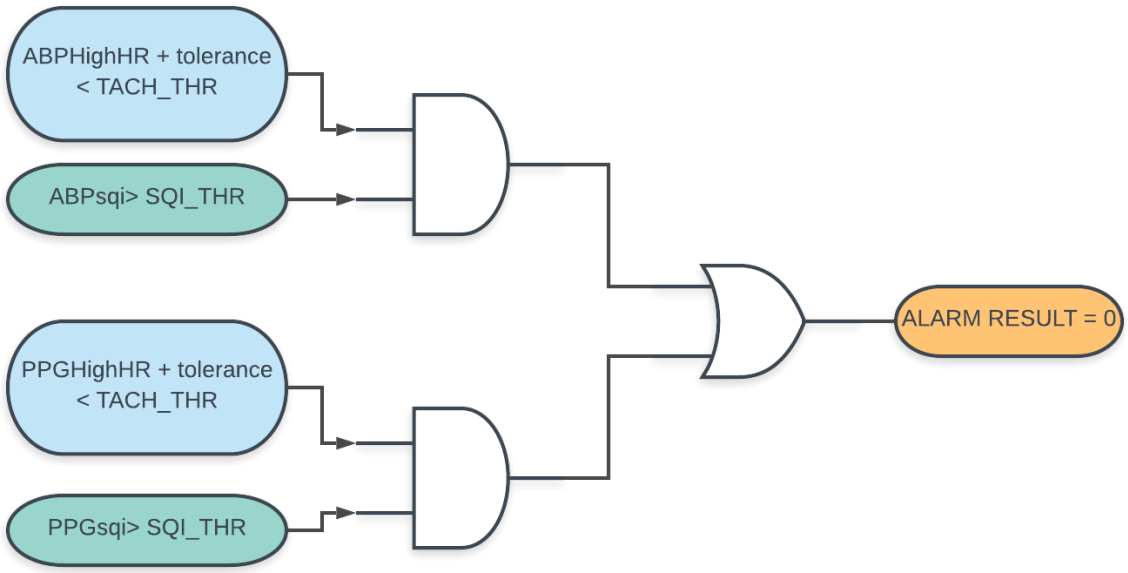



**Figure 4.11. Gate schema for the tachycardia determination block in base script.**

The tests for showing how much the noise reaches to distort the signal are then made in terms of the discrimination variables. For each one of the cases (bradycardia and tachycardia) these features were calculated for the noise conditions and the original conditions. The following figures show the variance in the heart rate results for each one of the arrhythmias: low heart rate for bradycardia and high heart rate for tachycardia.

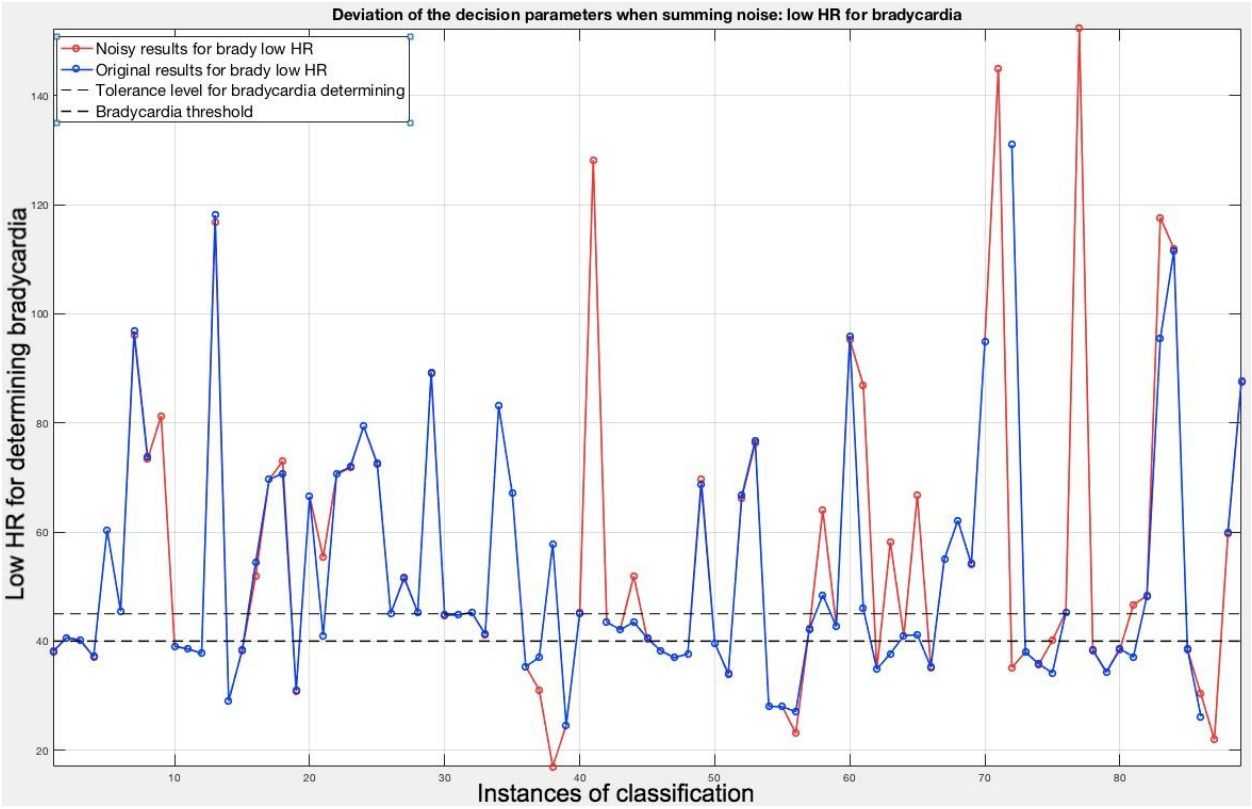



**Figure 4.12. Deviation of the low heart rate decision variable for bradycardia.**

It should be highlighted at this point that some of the results for these variables are retrieved as NaN, because the algorithm initializes them like that. In the case these variables cannot be calculated for validation reasons, these are left as NaN. This is the reason why in the continuous *figure 4.12*, some pieces of the original signal aren't plotted, but this is an internal problem of the algorithm, thus further explanation on this will not be given.

In dotted lines are shown the thresholds for the classification of the specific instance between having or not having bradycardia. The former figure allows to see that even low intensity movements provoke this low heart rate results to distort, therefore we can infer that the heart rate derived from each record is being biased at some point. As the decision for the determination of bradycardia is also affected by a tolerance level, which is a constant value for all records, the solving of the inequation would give a new determination threshold. In this case, the upper dotted line is the edge value for bradycardia determination: all the points located over it would be classified as 'non bradycardia'. However, these changes must be equally backed by the quality index variation. If both these values don't produce enough variation to go beyond the threshold, then the classification will not change in terms of sensitivity or specificity. This is the same case for tachycardia.

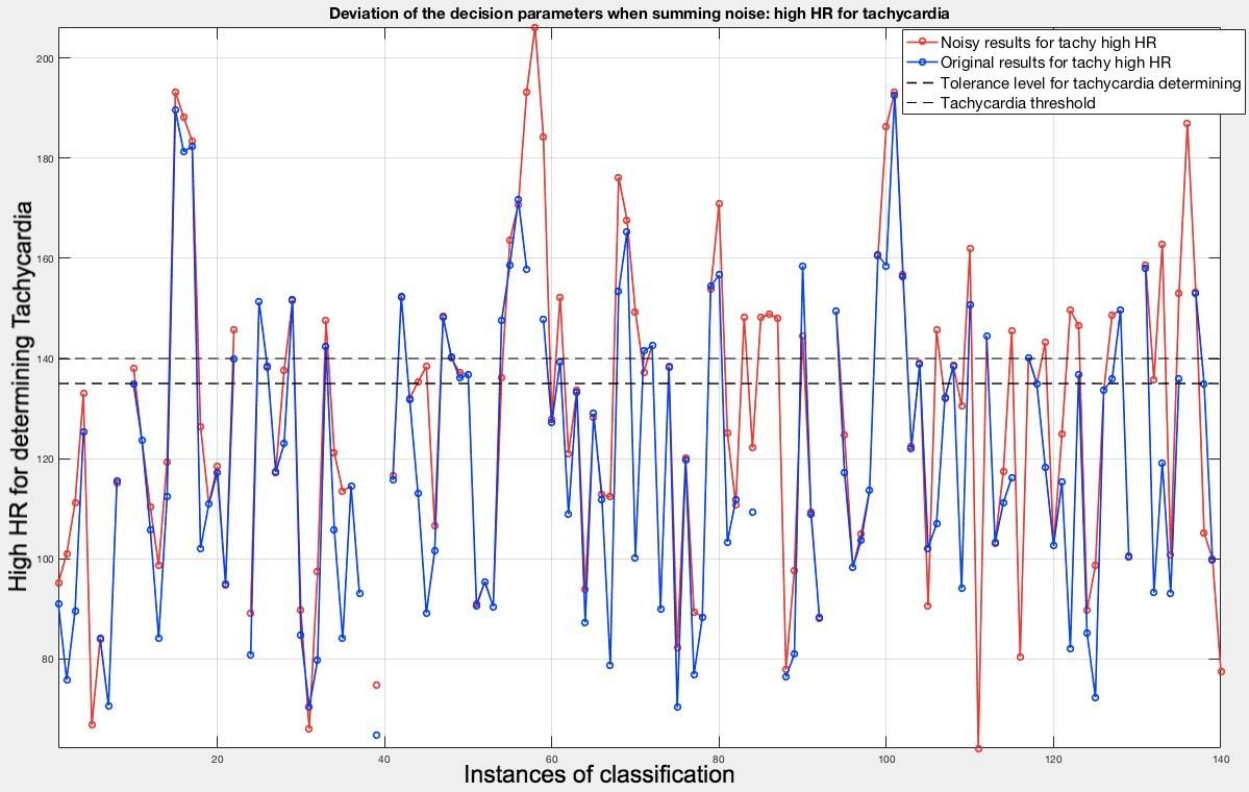



**Figure 4.13. Deviation of the high heart rate decision variable for tachycardia.**

Tachycardia results are far more affected by the simulation of the noisy scenario, as it can be evidenced by the *figure 4.13.* In this case, the dotted line in the lower part is the one that delimits the edge value for determining tachycardia. Then, any value below this line is considered as a false tachycardia alarm, of course, if the results of the signal quality index are equally distorted. A correlation analysis is also done over these heart rate variables to prove that even in low quantity, results get distorted. The scattering of the original results vs the noise results should provide some deviation evidence around a 45-degree diagonal. These figures are shown below adjoint with the correlation coefficients between the original and noisy results. This coefficient should be less than one, demonstrating that the noisy results are being biased from the originals.

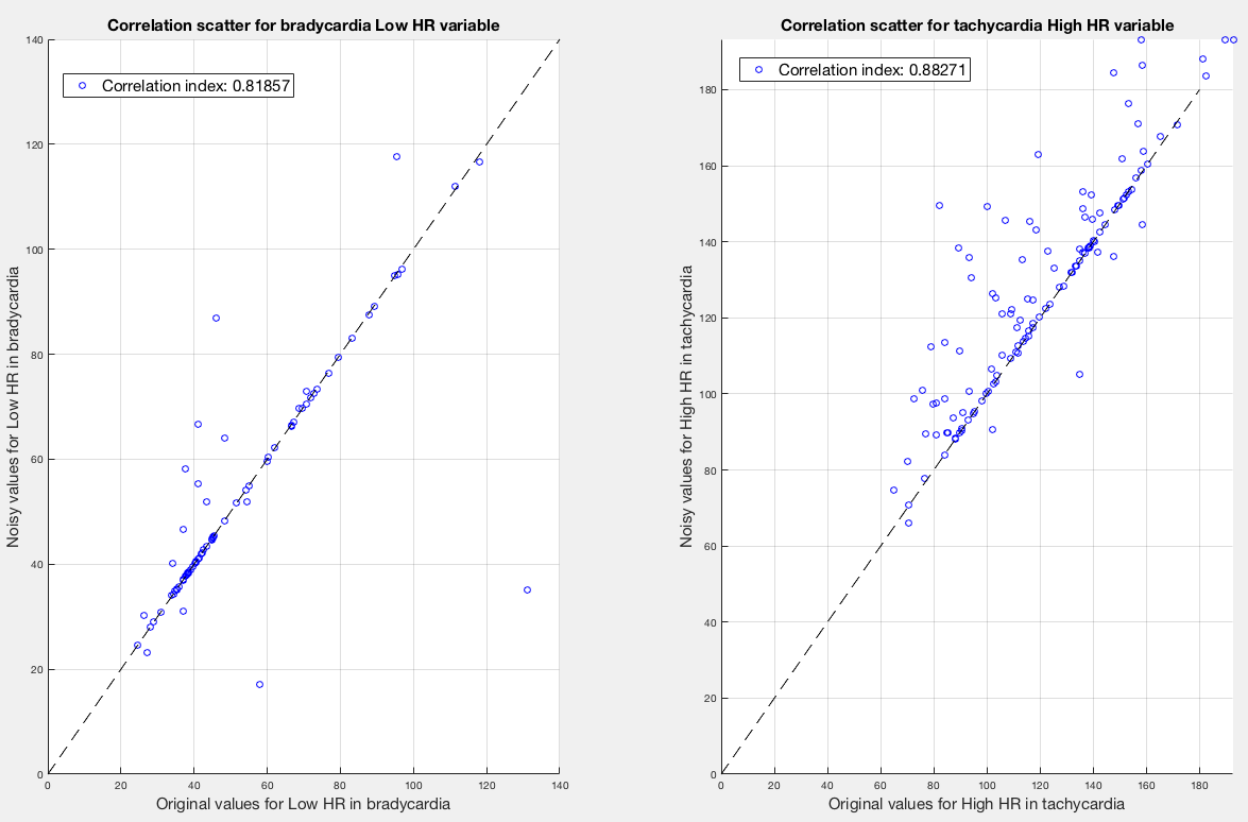



**Figure 4.14. Correlation between the decision variables after adding noise.**

The correlation index in tachycardia and bradycardia heart rate show the real situation that the classifier in base script passes through given the deviation of the noisy results from the original results. However, regarding sensitivity and specificity results, it's much harder to obtain a noticeable change, since this specific heart rate condition is not the only determining variable for turning a zero-classification result into a one classification result. However, these variations are shown to indicate that the noise is corrupting pulsatile signals in the measure that these are low intensity movement artifacts.

Further results are given in terms of sensitivity and specificity in the tables below. Here we can see that even when the heart rate estimations for each arrhythmia vary, it's actually difficult to move the threshold of the signal quality index. Therefore, the general classification results do not change in a major quantity.

A comparison between the sizes of the noise signal and the arrhythmia signal allows to explain this. The noise signal does not have such a comparable value with respect to the arrhythmia signal. *Figure 4.15* enables to see the amplification factor needed for the noise to be size comparable to the signal. However, we consider this does not indicate a poor noise characterization. Tests conducted in *chapter 3* established that this noise model is the most suitable for a low intensity movement artifact representation. Oppositely, this indicates that even when the noise is small, it damages in a certain measure the heart rate results for each arrhythmia. Bonomi's affirmation seems to be right, if we consider that his evaluation is made

just from heart rate features. In our case, the signal quality index is a differing factor from Bonomi, avoiding the results to change greatly. We can still observe the change in some of the classification results from *Table 4.5.*

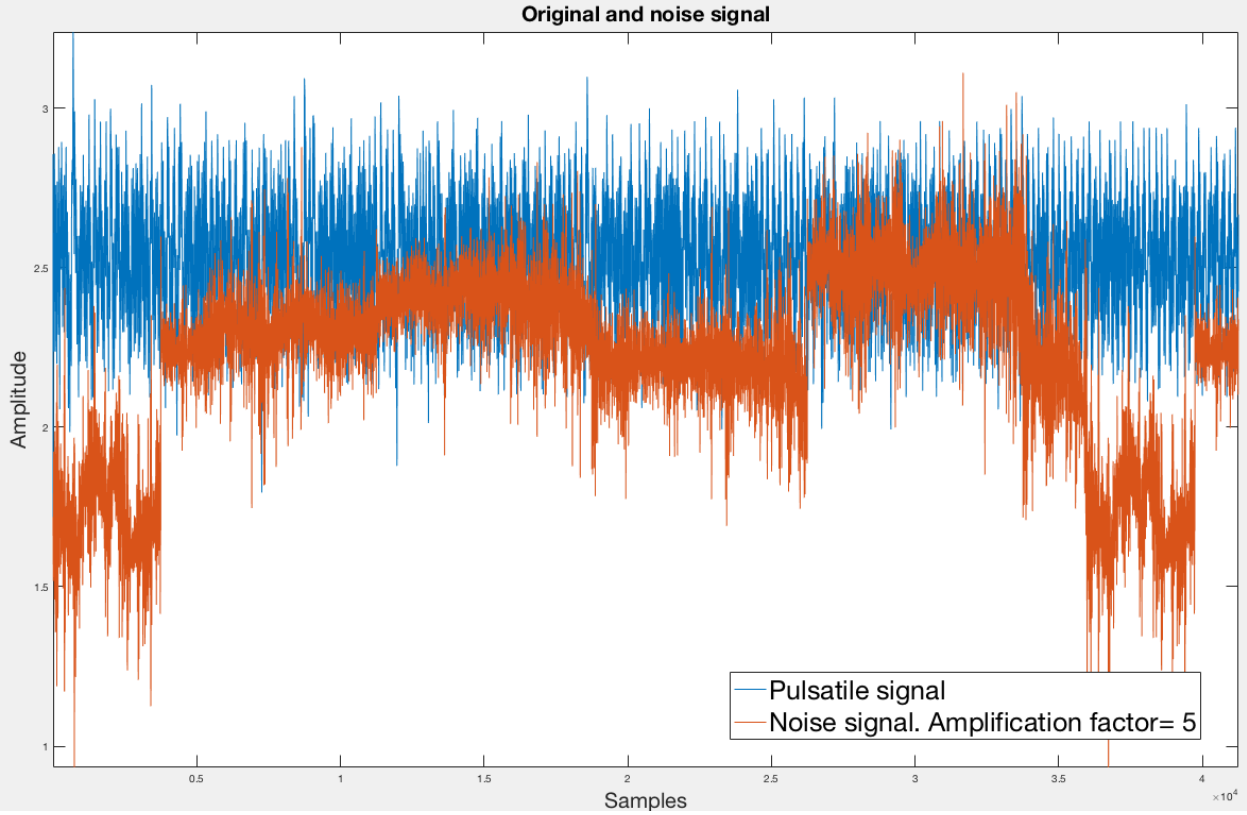



**Figure 4.15. Comparison between sizes of original and noise signal.**

The following tables have been retrieved from *Appendix C* and these contain the results in the original conditions and noisy conditions. Original conditions refer to the use of the classification scripts as they come from the Physionet repository, containing thus all of the preprocessing stages each author considered. Noisy conditions refer to the removal of these preprocessing techniques and adding the low intensity movement artifact noise, in hope of simulating the scenario from Bonomi.

| | Original results | | Noisy results | |
|---|---|---|---|---|
| Arrhythmia | Sensitivity | Specificity | Sensitivity | Specificity |
| Bradycardia | 97,826 | 58,140 | 95,652 | 58,140 |
| Tachycardia | 87,023 | 55,556 | 90,076 | 55,556 |

**Table 4.5. Results after adding noise to the base classification script.**

| | Original results | | Noisy results | |
|---|---|---|---|---|
| Arrhythmia | Sensitivity | Specificity | Sensitivity | Specificity |
| Bradycardia | 91,304 | 79,069 | 84,783 | 83,271 |
| Tachycardia | 94,656 | 55,555 | 94,656 | 55,556 |

**Table 4.6. Results after adding noise to the Miguel Caballero classification script.**

| | Original results | | Noisy results | |
|---|---|---|---|---|
| Arrhythmia | Sensitivity | Specificity | Sensitivity | Specificity |
| Bradycardia | 93,478 | 79,070 | 91,304 | 81,395 |
| Tachycardia | 100 | 77,778 | 100,000 | 33,333 |

**Table 4.7. Results after adding noise to Chengyu Liu classification script.**

| | Original results | | Noisy results | |
|---|---|---|---|---|
| Arrhythmia | Sensitivity | Specificity | Sensitivity | Specificity |
| Bradycardia | 91,304 | 83,721 | 89,130 | 79,070 |
| Tachycardia | 100 | 44,444 | 98,473 | 44,444 |

**Table 4.8. Results after adding noise to Sardar Ansari classification script.**

#### 4.6.4.2. Results after the operation: best denoisers for each classification script

This section presents the results obtained with each one of the selected denoisers for the impact evaluation over the F1 score (sensitivity and precision) and specificity measures regarding the brady- and tachycardia detection.

The first test comprehends the comparison of the two main performance measures for the classifiers, which are sensitivity and specificity. With this test we desire to obtain a denoiser which causes the greatest positive impact over sensitivity and specificity in each classifier. In this way, we assure that the denoiser tends to improve both measures as much as possible.

The second test uses the F1 score to reaffirm the first test. If the greater average F1 score coincides with the results of the first tests, then we conclude that the denoising case is making an improvement of both sensitivities (for bradycardia and tachycardia) in a balanced way. F1 average score gives us an idea of the increase in both sensitivities (bradycardia and tachycardia) at the same time.

Then, the specificity is reviewed, since F1 score does not directly relate to this metric. If the specificity results show a negative impact, then the denoising case is discarded, because it

shows a great improvement on sensitivities at the cost of diminishing specificities. Oppositely, if specificity results show a maximization or at least an improvement of this measure, then the denoising case is considered as the best denoising case classification script.

For a better comprehension of the figures shown below, each one of the denoisers was named as a ‘case'. These cases are the ones mentioned in *subsection 4.4.*

- Case 1: wavelet denoising with Daubechies db8 wavelet with a three-level reconstruction tree structure.
- Case 2: wavelet denoising with Daubechies db6 wavelet with a four-level reconstruction tree structure.
- Case 3: wavelet denoising with Daubechies db10 wavelet with a four-level reconstruction tree structure.
- Case 4: wavelet denoising with Daubechies db6 wavelet with a five-level reconstruction tree structure.
- Case 5: wavelet denoising with symlets sym4 wavelet with a three-level reconstruction tree structure.
- Case 6: wavelet denoising with symlets sym6 wavelet with a three-level reconstruction tree structure.
- Case 7: Savitzky-Golay smoothing filtering with parameters 3 (cubic fitting) and 41 (window size selected).
- Case 8: empirical mode decomposition (open source code[7]).
- Case 9: wavelet denoising with thresholding.
- Case 10: moving median filtering with parameters 11 (window size selected)

To determine which wavelet would be chosen for the thresholding case, an observation of the results for thresholding in all cases is made. This implies the calculation of the thresholding for all of the wavelet cases, followed by a comparison of the best results between the thresholding cases. Then, the wavelet thresholding case which induces the most positive increase over both parameters is taken as the case 9. All of these tests can be seen with more detail in *Appendix C.*

First, the results for the base script are presented. For the sake of space limitation, these compiled results can be found in the bar graphs from *Appendix E.* The results are presented for bradycardia from case 1 to case 5 (for space and visualization reasons), then in another figure are represented the rest of the cases, from 6 to 10. Tachycardia results will be presented using the same structure.
The bar graphs are formed by the quantities corresponding to the sensitivity presented under noise conditions, then the sensitivity presented when the denoising is done. Together with

[7] Code taken from MATLAB File exchange central, available here: https://la.mathworks.com/matlabcentral/fileexchange/52502-denoising-signals-using-empirical-mode-decomposition-and-hurst-analysis.

these two results, there is an additional bar showing the increase or decrease of these quantities. The same process will be done with the specificity of the issued cases.
Given that these results cannot be reproduced in this same document we show only the deviation outcomes. These are the ones that give more information about the impact of each denoiser, and correspond to the tables from *4.9* to *4.17.*

*Table 4.9* shows the results after denoising. Almost all cases for bradycardia show an improvement over sensitivity. In cases like 1, 5, 7 and 10, sensitivity improvement does not represent a decrement over specificity. Therefore, we can state that the performance of the classifier algorithm for this specific arrhythmia improved in the desired way. In cases 2, 6 and 9, the specificity reaches an improvement, maintaining sensitivity results invariant. This suggests in this case, the denoising method works for greater detection of TN. Case 10 oppositely increases sensitivity and maintains specificity unchanged. Cases like 3 and 8 show a significant improvement over sensitivity results, although this implies a specificity decrease. This can be the product of a tradeoff between these two parameters, showing that the denoising in these cases moves the cutoff point but not in the best way. In case 4, the decrease of the specificity is enormous, and the sensitivity reaches a higher result. We suspect a trivial classification case might be occurring, given the structure of the base script decision (the result is true unless the contrary is proofed).
Now, if we look at tachycardia results, it shows a less favorable scenario for almost every denoisers. The only one which generates an improvement over both parameters is the case 3. Cases 2,5 and 6 improve sensitivity and maintain specificity.
Lastly, a cross of the tachycardia and bradycardia cases is made. We assume case 5 as the denoiser which exposes the best results for the base script. Although it maintains as equal the specificity results for tachycardia, it also improves in a balanced way the other parameters, without demonstrating the accomplishment of a trivial case.
We can usually see that an increment in bradycardia parameters generates slight decrements in tachycardia parameters and also the opposite way. For example, case 3 goes over significant improvements of both parameters for tachycardia, but this implies a worsening over bradycardia specificity.

| DEVIATION RESULTS FOR EACH DENOISER CASE IN BASE SCRIPT | | | | |
|---|---|---|---|---|
| CASE | BRADYCARDIA | | TACHYCARDIA | |
| | SENSITIVITY (%) | SPECIFICITY (%) | SENSITIVITY (%) | SPECIFICITY (%) |
| 1 | +2,174 | +4,651 | -0,763 | 0 |
| 2 | 0 | +2,326 | +1,257 | 0 |
| 3 | +4,348 | -2,326 | +2,290 | +11,111 |
| 4 | +2,174 | -25,581 | +9,161 | -33,3338 |

| | | | | |
|---|---|---|---|---|
| 5 | +2,174 | +6,977 | +0,764 | 0 |
| 6 | 0 | +4,651 | +1,527 | 0 |
| 7 | +2,174 | +2,326 | -3,053 | +11,111 |
| 8 | +4,348 | -4,652 | +5,344 | -33,334 |
| 9 | 0 | +4,651 | -0,763 | 0 |
| 10 | +2,174 | 0 | -0,763 | 0 |

**Table 4.9. Deviation results for each denoiser case in base script.**

*Table 4.10* presents the values of F1 score in the noisy case and then the 10 denoising cases. The score presented in this table is an averaging of both measures for tachycardia and bradycardia in each case. We wish an improvement of this score in both arrhythmias, in a balanced way. With this averaging, we avoid the choosing of a case where F1 score increases significantly for one arrhythmia while the other is very low.

| Case | Average F1 Score (%) | Case | Average F1 Score (%) |
|---|---|---|---|
| Noisy | 87,381 | 6 | 87,426 |
| 1 | 88,485 | 7 | 87,851 |
| 2 | 88,181 | 8 | 88,600 |
| 3 | 88,893 | 9 | 88,697 |
| 4 | 86,008 | 10 | 87,713 |
| 5 | 89,304 | | |

**Table 4.10. Average F1 Score results for each denoiser case in base script**

Improving greatly the F1 score value means a better balance between precision and sensitivity parameters in both arrhythmias. In this script, the best case for an F1 score improvement is case 5, which reaffirms the results obtained in *Table 4.9.* The best case of denoising for this script increases in the right way the sensitivity for both arrhythmias, improves specificity for bradycardia and maintains constant the specificity for tachycardia. This approach demonstrates a positive impact on bradycardia detection and the need for a better technique over tachycardia (such as case 3), especially over specificity.

*Table 4.11* exposes Miguel Caballero script results. Bradycardia and tachycardia results fall into a less improved outlook than the base script. We notice a greater difficulty for finding a denoiser which improves specificity and sensitivity results. However, there are some cases for

bradycardia where specificity is maintained constant, and sensitivity is increased. These are the cases 1, 5, and 6.
Regarding tachycardia, cases like 1, 2, and 6 show the same behavior (sensitivity improves in very slight steps, while specificity is maintained constant). Case 10 is the only one which improves both parameters for tachycardia but causes, in turn, a decrement over bradycardia specificity. Other cases like 2, 3, 7, 9 and 10 represent a tradeoff.
This case is more challenging for us. It is needed to weigh all the elements, looking for a case which increases in a balanced way both of the parameters. However, this seems improbable. In this kind of situation, we prefer the denoiser which gives small steps over sensitivity and does not decrease specificity.
*Table 4.11* enables to see that there exist cases which present significant improvements for tachycardia specificity. However, this comes along with a diminished sensitivity. Case 6 seems to be pretty stable over specificity, suggesting that the denoising method allows detecting more TP, without having any adverse effect over TN for both arrhythmias.

| DEVIATION RESULTS FOR EACH DENOISER CASE IN CABALLERO SCRIPT | | | | |
|---|---|---|---|---|
| CASE | BRADYCARDIA | | TACHYCARDIA | |
| | SENSITIVITY (%) | SPECIFICITY (%) | SENSITIVITY (%) | SPECIFICITY (%) |
| 1 | +4,348 | 0 | +0,764 | 0 |
| 2 | +6,522 | -13,954 | +0,764 | 0 |
| 3 | +4,348 | -6,977 | +1,527 | +11,111 |
| 4 | 0 | -16,279 | +4,581 | -33,333 |
| 5 | +2,174 | 0 | 0 | 0 |
| 6 | +4,348 | 0 | +0,764 | 0 |
| 7 | +10,870 | -6,977 | -2,290 | +11,111 |
| 8 | -6,522 | -6,977 | +3,054 | -33,333 |
| 9 | +6,521 | -2,326 | 0 | 0 |
| 10 | +6,521 | -2,326 | -0,763 | 0 |

**Table 4.11. Deviation results for each denoiser case in Miguel Caballero script.**

| Case | Average F1 Score (%) | Case | Average F1 Score (%) |
|---|---|---|---|
| Noisy | 90,268 | 6 | 91,694 |
| 1 | 91,694 | 7 | 91,451 |
| 2 | 89,661 | 8 | 87,251 |
| 3 | 90,730 | 9 | 91,425 |
| 4 | 87,902 | 10 | 91,627 |
| 5 | 90,888 | | |

**Table 4.12. Average F1 Score results for each denoiser case in Miguel Caballero script**

The best results for F1 score in *Table 4.12* are the cases 1 and 6. Both demonstrate an increment over sensitivity without varying specificity. This result implies that denoisers help in better TP detection, triggering the alarm where there actually exists an arrhythmia.
In this way, we conclude that the preferent cases in this script are case 1 and 6.

In the next two scripts (Chengyu and Sardar), an additional variable is added, which is the use of the ECG signal. As we stated before, the desired simulated scenario is the one where there exists movement artifact noise due to low intensity movements and this hasn't been preprocessed. To do so, the filter from both signals ECG and ABP/PPG are removed. This is a Butterworth passband filter which deals with high and low-frequency artifacts at the same time. Then the noise is added to ABP/PPG signals. However, we don't dispose a noise characterization for ECG signal, as these are not within our scope. It is later decided to remove the preprocessing filters for ECG signals. We hope that this could cause any effect or at least don't turn the classification into a matter of only ECG signal information. For resembling the denoised scenario, these ECG filters are put again in addition to the implementation of each denoising case for pulsatile signals.
*Table 4.13* shows the results for Chengyu script, where we can see a general improvement in almost all bradycardia sensitivities. Those cases that improve bradycardia sensitivity without damaging specificity or even improving it, are cases 5, 6 and 9. In all other cases, it is shown a decrease of the specificity, in different quantities. Case 8 seems to be a trivial case because the decrement of the specificity implies a great increment over sensitivity.

| DEVIATION RESULTS FOR EACH DENOISER CASE IN CHENGYU SCRIPT | | | | |
|---|---|---|---|---|
| CASE | BRADYCARDIA | | TACHYCARDIA | |
| | SENSITIVITY (%) | SPECIFICITY (%) | SENSITIVITY (%) | SPECIFICITY (%) |
| 1 | +4,348 | -2,326 | -0,763 | +44,445 |
| 2 | +2,174 | -16,279 | -3,053 | +44,445 |
| 3 | +6,522 | -9,302 | -3,817 | +44,445 |
| 4 | 0 | -20,923 | -2,290 | +33,333 |
| 5 | +2,174 | +2,326 | -0,763 | +44,445 |
| 6 | +2,174 | 0 | -0,763 | +44,445 |
| 7 | +4,348 | -4,651 | -3,053 | +44,445 |
| 8 | +6,522 | -20,930 | -1,527 | +44,445 |
| 9 | +2,174 | 0 | -0,763 | +44,445 |
| 10 | +4,348 | -2,326 | -0,763 | +44,445 |

**Table 4.13. Deviation results for each denoiser case in Chengyu Liu script.**

Tachycardia obtains a noticeable improvement over all specificities, depending on the issued case. All sensitivities are worsened in different quantities. Although we don't consider this as a bad result, taking into account that the result for noisy conditions in tachycardia sensitivity is 100%. As it can be seen in tables from *Appendix C*, this result for 100% of sensitivity (in noise conditions) is not attributed to a proper classification. On the contrary, the results for specificity fall into a trivial case. This is the reason why the specificity changes are so abrupt after denoising. The denoising method is inducing the results into a more proper set of results, pushing the classifier out of the trivial case. Those cases which make a better work with tachycardia are cases 1, 5, 6, 9 and 10.

Although this specificity improvement cannot be entirely attributed to pulsatile signals, at least it can be stated that a combination of both signals cleaned adequately, allows a proper classification of these arrhythmias.

| Case | Average F1 Score (%) | Case | Average F1 Score (%) |
|---|---|---|---|
| Noisy | 90,268 | 6 | 93,760 |
| 1 | 93,874 | 7 | 92,846 |
| 2 | 90,192 | 8 | 90,903 |

| | | | |
|---|---|---|---|
| 3 | 92,338 | 9 | 93,760 |
| 4 | 88,855 | 10 | 93,874 |
| 5 | 94,221 | | |

**Table 4.14. Average F1 Score results for each denoiser case in Chengyu Liu script.**

*Table 4.14* displays the best value for F1 score, which falls again in case 5. With this in mind, we can state that the best denoising case for this script is case 5.
Ansari script results show a constant improvement over tachycardia sensitivities, while specificity remains unchanged in most of the cases.
Sensitivities of both tachycardia and bradycardia show an increase, but specificities decrement or remain unchanged.

| DEVIATION RESULTS FOR EACH DENOISER CASE IN SARDAR SCRIPT | | | | |
|---|---|---|---|---|
| CASE | BRADYCARDIA | | TACHYCARDIA | |
| | SENSITIVITY (%) | SPECIFICITY (%) | SENSITIVITY (%) | SPECIFICITY (%) |
| 1 | +2,174 | +2,325 | +1,527 | 0 |
| 2 | +2,174 | 0 | +1,527 | 0 |
| 3 | +4,348 | +2,325 | +1,527 | 0 |
| 4 | +4,348 | -2,326 | +0,764 | +11,112 |
| 5 | +2,174 | 0 | +1,527 | 0 |
| 6 | +4,348 | 0 | +1,527 | 0 |
| 7 | +4,348 | -2,326 | +1,527 | 0 |
| 8 | +2,174 | -2,326 | +1,527 | 0 |
| 9 | +4,348 | 0 | +1,527 | 0 |
| 10 | +4,348 | 0 | +1,527 | 0 |

**Table 4.15. Deviation results for each denoiser case in Sardar Ansari script.**

With these results, we state that the most suitable case for improving almost all parameters is case 3. This outcome is reaffirmed with the results for F1 score, where case 3 stands in the first place.

| Case | Average F1 Score (%) | Case | Average F1 Score (%) |
|---|---|---|---|
| Noisy | 91,388 | 6 | 92,941 |
| 1 | 92,814 | 7 | 92,498 |
| 2 | 92,363 | 8 | 91,920 |
| 3 | 93,364 | 9 | 92,941 |
| 4 | 92,491 | 10 | 92,941 |
| 5 | 92,362 | | |

**Table 4.16. Average F1 Score results for each denoiser case in Sardar Ansari script.**

With the results for the best denoising case in each script, an evaluation stage is conducted for the different classifiers performance. Three ROC curves for each script will be calculated, one for the corrupted results (called as noisy results from now on), other for the original results and the last one will be for the best denoiser case in that specific script. This analysis is made to assess how much improvement is induced through the denoisers in terms of a graphic representation.

The procedures for obtaining the ROC curves estimations can be seen in *Appendix F* along with a more detailed comparison of the confusion matrices information for each script. For each script result, it will be displayed the ROC curve estimation that gives a better visual properness of the empirical curve. Altogether with this, the AUC tables which compare the noisy, denoised, and original results under the different ROC curves estimation are shown. Finally, this appendix shows an improvement of the best denoising cases over the noisy results and sometimes an improvement or a equaling with the original results.

The figures presented below make a better appropriation of the empirical curve, as it was analyzed in *Appendix F*. In *Appendix C* are presented the other ROC curves estimations with the three remaining methods.

First, we will start with the ROC and AUC results for the base script. For future references and visualization issues, the x-axis will represent the FPR and the y-axis will represent the TPR.

| | BRADYCARDIA RESULTS | | | TACHYCARDIA RESULTS | | |
|---|---|---|---|---|---|---|
| | ORIGINAL | NOISY | DENOISED | ORIGINAL | NOISY | DENOISED |
| ESTIMATED AUC | 0,8473 | 0,8365 | 0,8751 | 0,8168 | 0,8282 | 0,8321 |

**Table 4.17. AUC comparison for the three conditions on both arrhythmias (base script).**

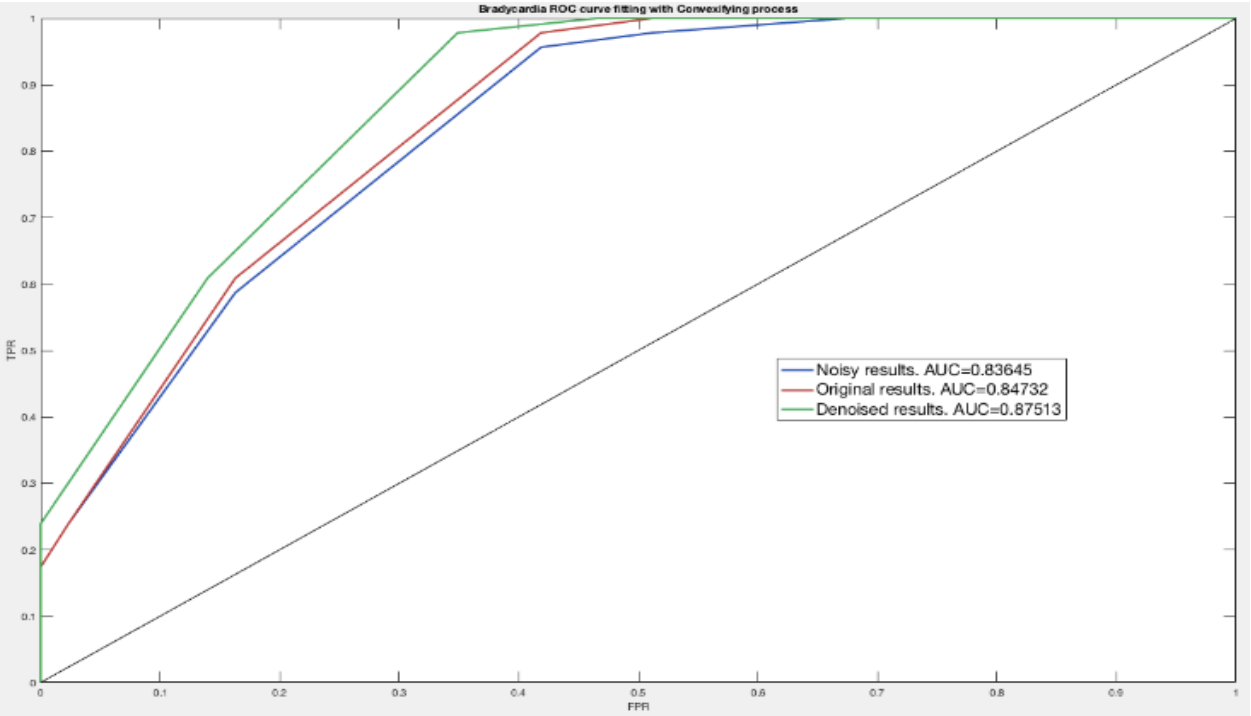



**Figure 4.16. Base script ROC curves for the noisy, denoised, and original conditions in bradycardia.**

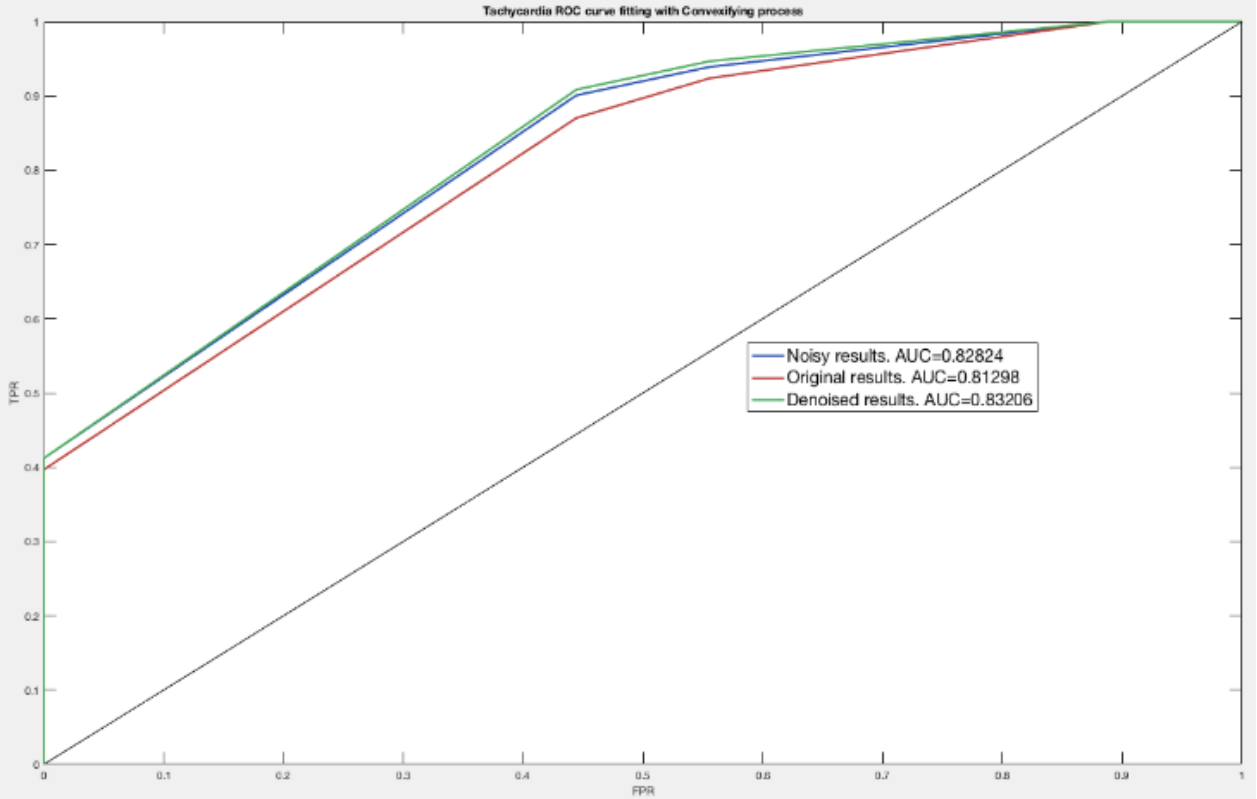



**Figure 4.17. Base script ROC curves for the noisy, denoised, and original conditions in tachycardia.**

The improvement of the denoised results over the noisy ones is seen in *figures 4.16* and *4.17.* These curves also verify situations such as the improvement of the sensitivity when the noise was added. In the case of tachycardia, the area under the curve comprised by the denoised case is slightly larger than the area under the curve of the noisy results. The denoised curve presents a tiny higher value on the y-axis, since the specificity is maintained and what increases is the sensitivity at about +0.77.
In the bradycardia ROC curves, the improvement over both the original and denoised cases is more noticeable, presenting an increase in both sensitivity and specificity.
The curves chosen for the base script representation were the convex hull estimations over the empirical curve because the binormal estimation showed slight hooks at the beginning of the curve. However, if this kind of visual improperness is little (such as the one in ROC curves for base script), the error is negligible.
Although, since we want to present a proper concave ROC curve, the convex hull estimation was chosen with the awareness that this induces an AUC overestimation. This overestimation is almost equal to all the cases (original, noisy, denoised) and for this reason, we consider it is not a mistake because the magnitudes are maintained in the same scale for the three areas under the curve. However, for a more realistic result about the classifier's performance, we suggest referring to the empirical ROC curve estimation for the AUC.

Now, in Caballero's script, we have the following results:

| | BRADYCARDIA RESULTS | | | TACHYCARDIA RESULTS | | |
|---|---|---|---|---|---|---|
| | ORIGINAL | NOISY | DENOISED | ORIGINAL | NOISY | DENOISED |
| ESTIMATED AUC | 0,9201 | 0,9057 | 0,9321 | 0,8511 | 0,8511 | 0,8550 |

**Table 4.18. AUC comparison for the three conditions on both arrhythmias (Caballero's script).**

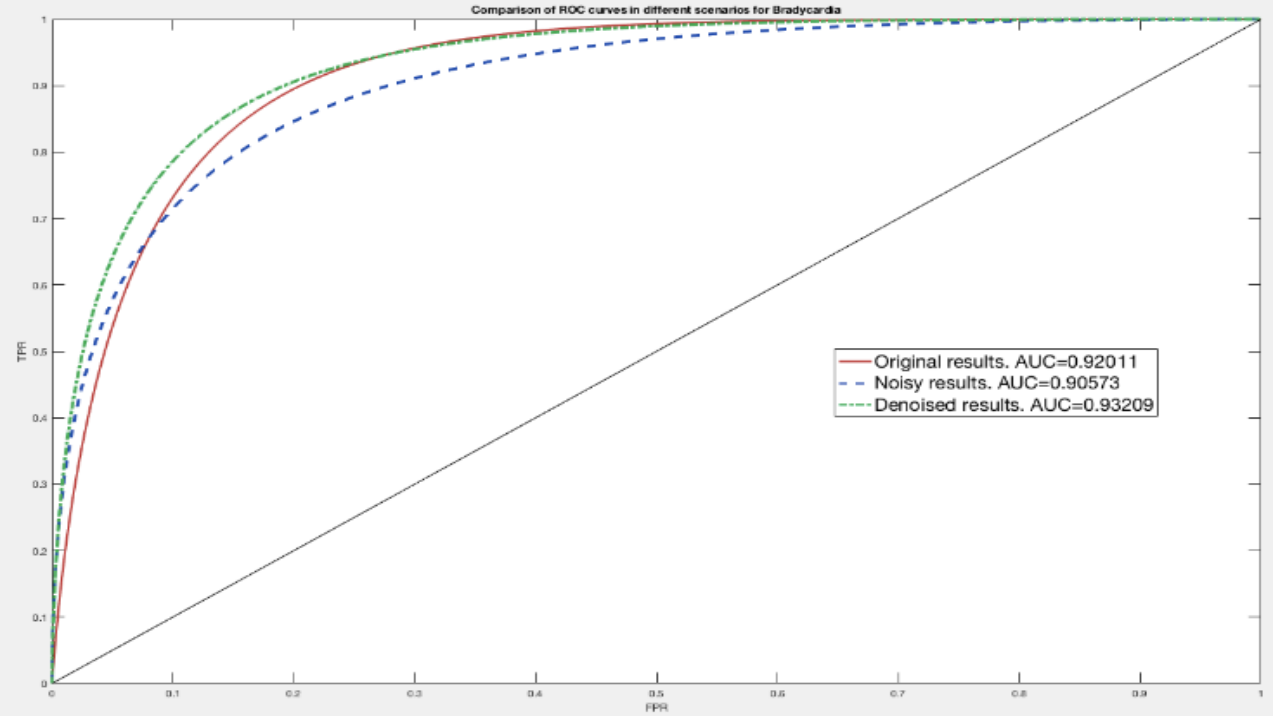



**Figure 4.18. Caballero's script ROC curves for the noisy, denoised, and original conditions in bradycardia.**

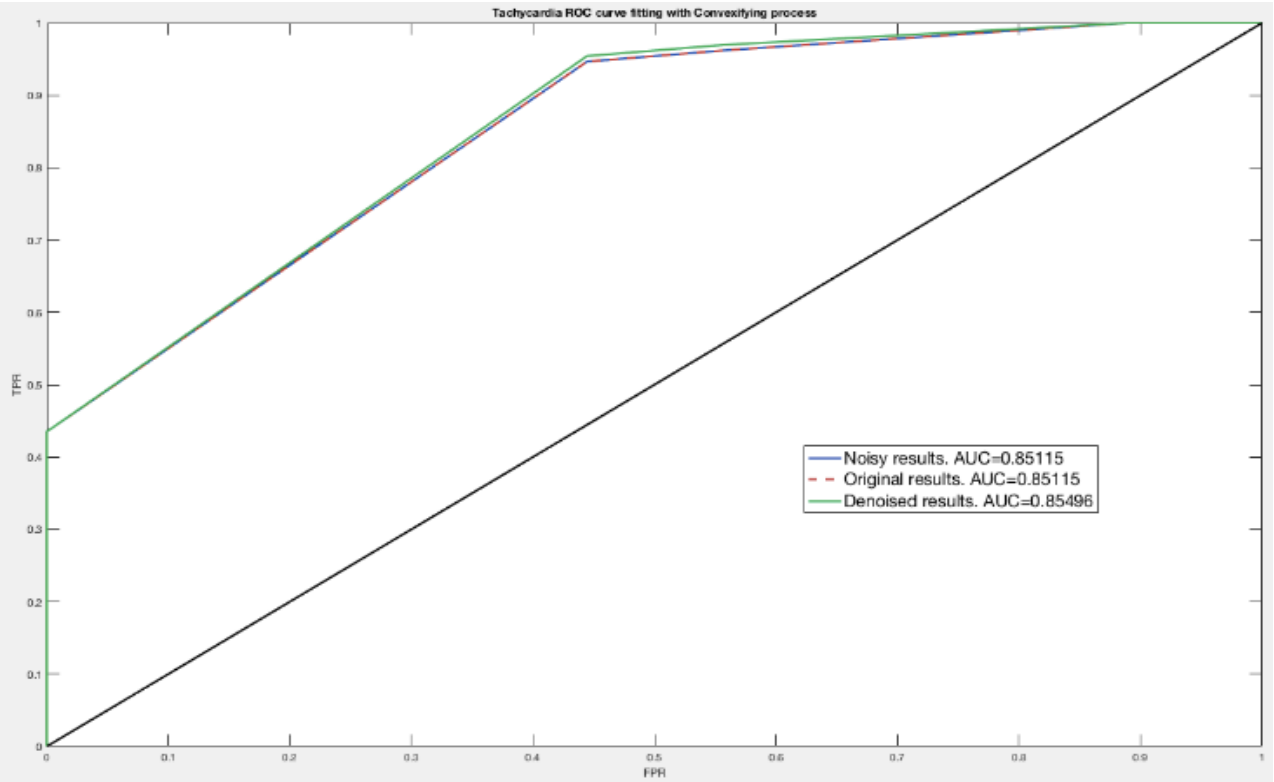



**Figure 4.19. Caballero's script ROC curves for the noisy, denoised, and original conditions in tachycardia.**

*Figure 4.18 and figure 4.19* expose the achievement of a better AUC with the denoised results over both the noisy and original results. It is also shown the tradeoff that occurs in bradycardia when the sum of the noise is done over the clean signal. The noisy ROC curve presents a greater closeness to the left side of the graph because the specificity has been improved while the closeness to the upper corner is diminished because the sensitivity has been decreased. In tachycardia results, it is possible to see the invariance of the results with and without noise (ROC curves for these two cases are overlapped). Therefore, this makes the denoised results to be compared with either of the two other results. However, the AUC improvement is minimal, but it's an increment over the original results too.
The convex hull curve was chosen for tachycardia representation because binormal presented a little hook at the beginning again.

The results for Chengyu's script are as follows.

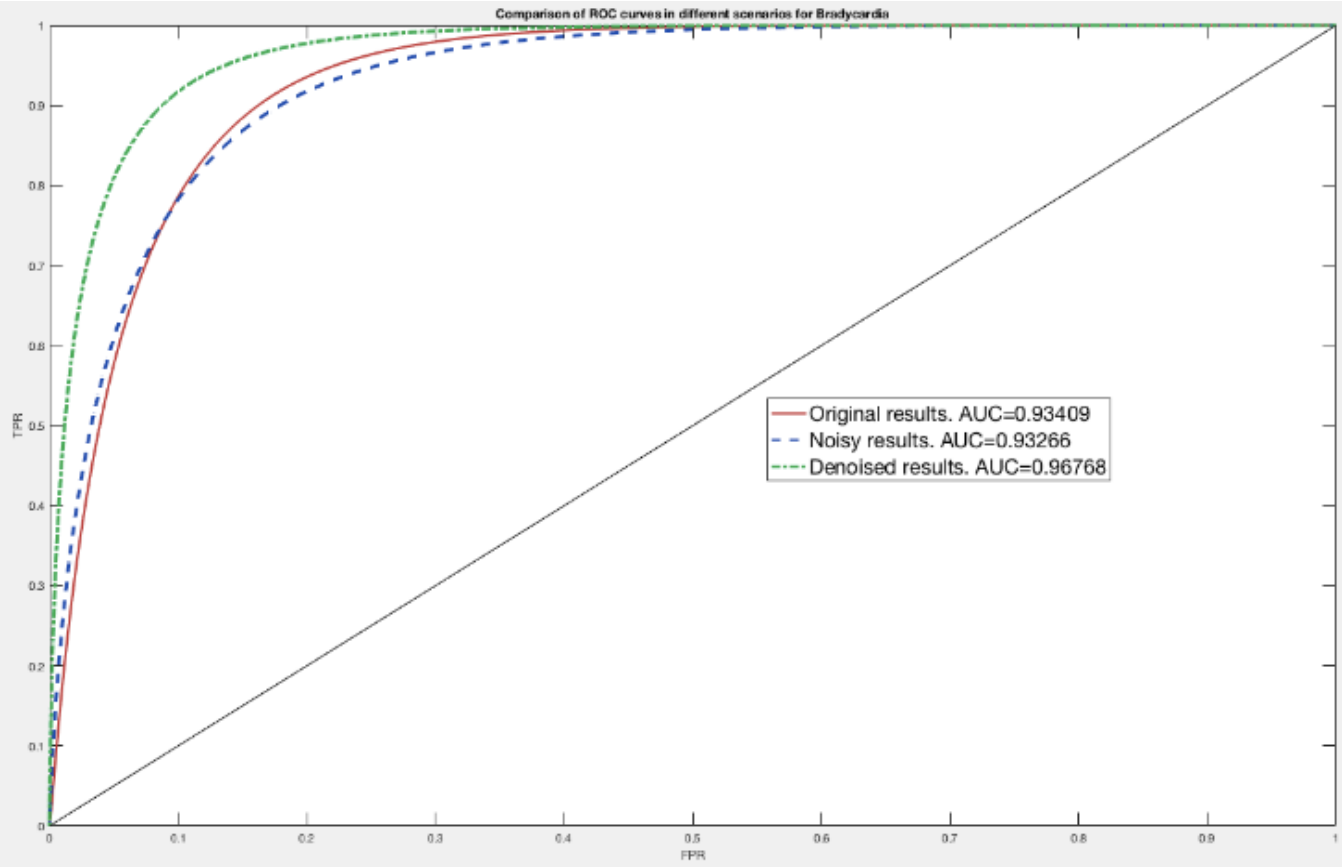



**Figure 4.20. Chengyu's script ROC curves for the noisy, denoised, and original conditions in bradycardia.**

| | BRADYCARDIA RESULTS | | | TACHYCARDIA RESULTS | | |
|---|---|---|---|---|---|---|
| | ORIGINAL | NOISY | DENOISED | ORIGINAL | NOISY | DENOISED |
| ESTIMATED AUC | 0,9341 | 0,9327 | 0,9677 | 0,9406 | 0,8312 | 0,9385 |

**Table 4.19. AUC comparison for the three conditions on both arrhythmias (Chengyu's script).**

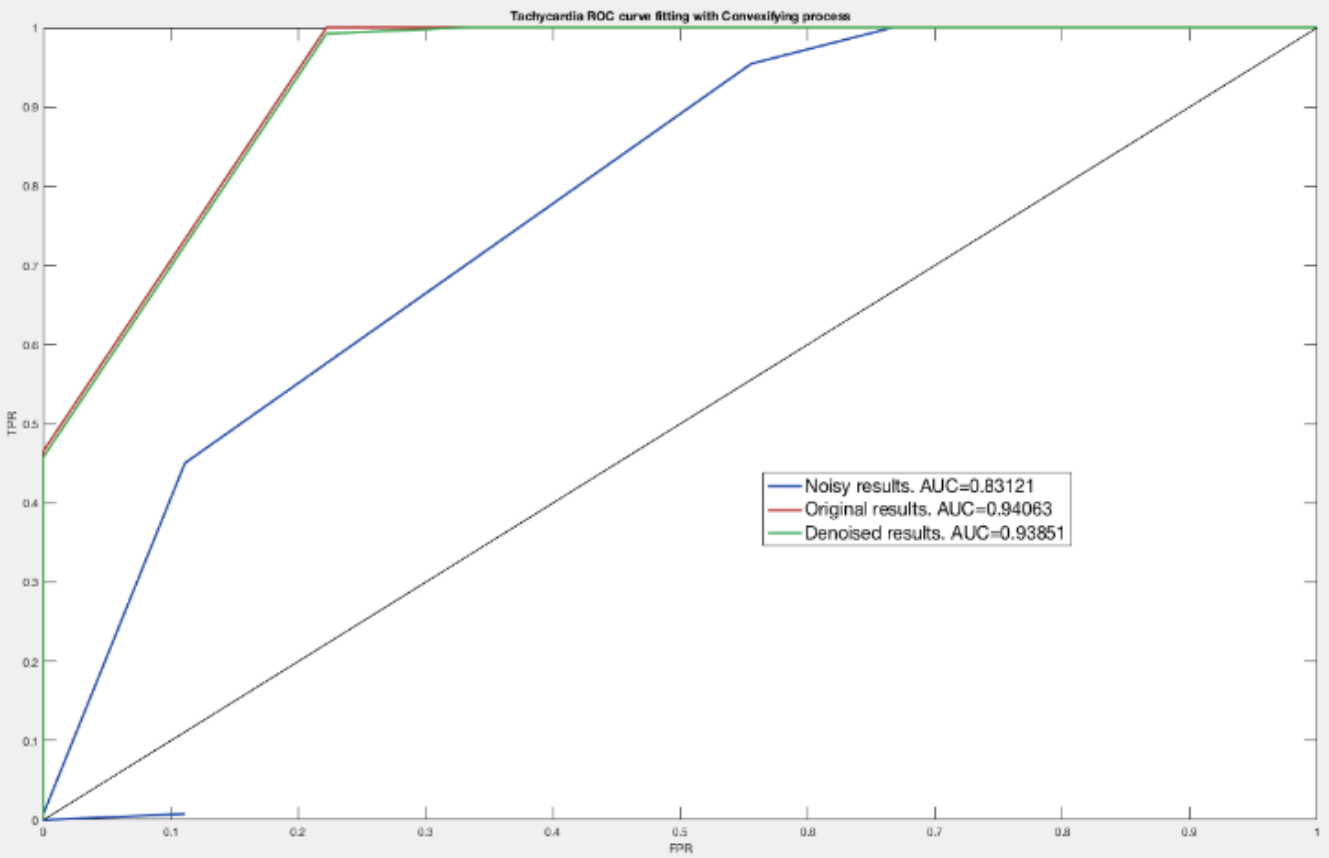



**Figure 4.21. Chengyu's script ROC curves for the noisy, denoised, and original conditions in tachycardia.**

The convex hull estimation for tachycardia is chosen because the binormal estimation could not fit at all the empirical curve form. This because of the use of cumulative distribution functions for the computation of the AUC (as it is explained in *Appendix F*). The improvement over the original results was achieved only in bradycardia while on tachycardia, original results were closely tracked by denoised results, but not equaled. Other details as the tradeoff caused by the noise sum in bradycardia are evident in *Figure 4.19*.
The original curve from *Figure 4.21* is slightly larger on the Y-axis, because the sensitivity of the original conditions could not be reached (only differentiated by -0,763).

However, the improvement from the noisy results is noticeable, caused by the change in the specificity. This makes the curve's slope slower, and therefore closer to the chance line. We could presume that the noise addition and the filtering removal for ECG signals distort the results in a great measure for this classifier.

In last place, Sardar results are exposed

| | BRADYCARDIA RESULTS | | | TACHYCARDIA RESULTS | | |
|---|---|---|---|---|---|---|
| | ORIGINAL | NOISY | DENOISED | ORIGINAL | NOISY | DENOISED |
| ESTIMATED AUC | 0,9449 | 0,9060 | 0,9459 | 0,8643 | 0,8605 | 0,8643 |

**Table 4.20. AUC comparison for the three conditions on both arrhythmias (Sardar's script).**

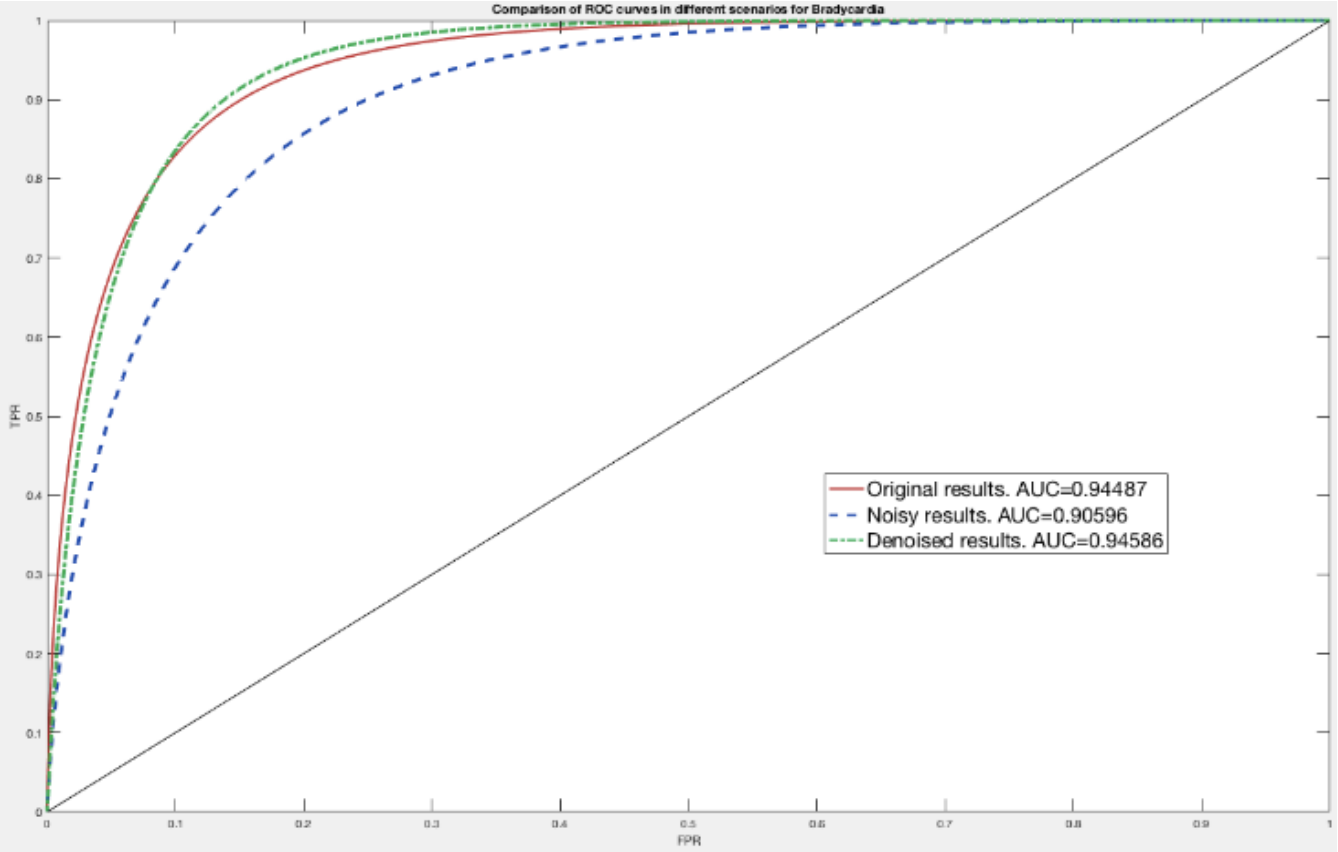



**Figure 4.22. Sardar's script ROC curves for the noisy, denoised, and original conditions in bradycardia.**

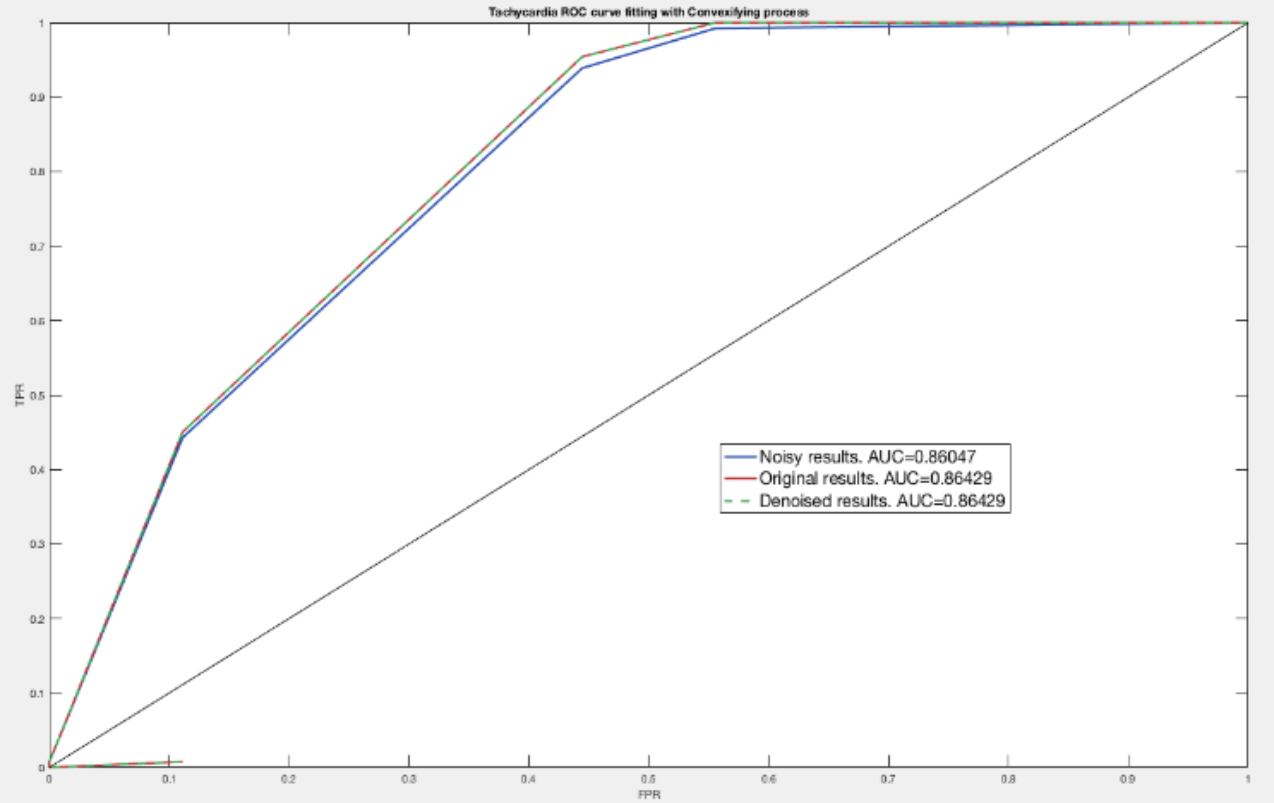



**Figure 4.23. Sardar's script ROC curves for the noisy, denoised, and original conditions in tachycardia.**

Noisy results show improvements over original results for bradycardia, using the binormal estimation and the convex hull estimation for AUC calculation. For tachycardia the original results were equaled by the denoised results. For this reason, a dotted line has been plotted. Tachycardia ROC curves suggest that the noise addition and filter removal from ECG signals is not as damaging as expected. This could be explained by the robustness of Sardar's algorithm, where he takes much more features from both signals and does not discard any of them. Chengyu approach makes a comparison of the features and then discard the ones which do not provide any significant information. Sardar creates multiple indicators that help in greater or lower measure to the correct alarm detection, without discarding any information.
According to the empirical estimation of the ROC curve, original bradycardia results could not be achieved in the denoised conditions. They are followed closely, showing that the denoiser can generate satisfactory results for a 'ceiling' that has excellent conditions.

With the results of the best denoising case for each script, we can finally assemble four different mechanisms (denoising stage plus an arrhythmia detection stage). A comparison between them is made in order to declare which one has the best performance. We know that the algorithms that only use ABP/PPG signals are not going to obtain as good results as the ones who use both ECG and ABP/PPG. However, this will allow us to analyze which classifier is the best depending on the signals used.

The given score by organizers is used to qualify each one's performance, just as it was made at the competition.
The equation that's used in the final scoring for the competition can be seen in [81], where a higher weight is placed upon the FN than the FP. This because the purpose is to avoid the triggering of the ICU alarm when it is not the case of arrhythmia detection, along with the correct identification of an alarm when the patient is experiencing an arrhythmia. FN is an alarm which should have been noticed, but it did not occur, thus diminishing the patients proper and opportune care.

$$Final\ Score = \frac{TP + TN}{TP + TN + FP + 5 * FN}$$

The scoring for each mechanism is computed, and then these two scores (one for tachycardia and one for bradycardia) are averaged, to present a global score for both arrhythmias in each script. These scores are presented in *Table 4.21.*

| Script | Partial scores | | Global score |
|---|---|---|---|
| Base | Bradycardia | 78,4946 | 72,2260 |
| | Tachycardia | 65,9574 | |
| Miguel Caballero | Bradycardia | 70,6422 | 74,9552 |
| | Tachycardia | 79,2683 | |
| Chengyu Liu | Bradycardia | 78,2178 | 86,6784 |
| | Tachycardia | 95,1389 | |
| Sardar Ansari | Bradycardia | 77,2277 | 86,8281 |
| | Tachycardia | 96,4286 | |

**Table 4.21. Final scoring table for tachycardia and bradycardia in each script.**

It is expected that these results are not as high as a normal accuracy, because the scoring equation is taking five times the mistakes made in the FN case. However, the results for the normal accuracy expression can be seen in *Appendix C.*

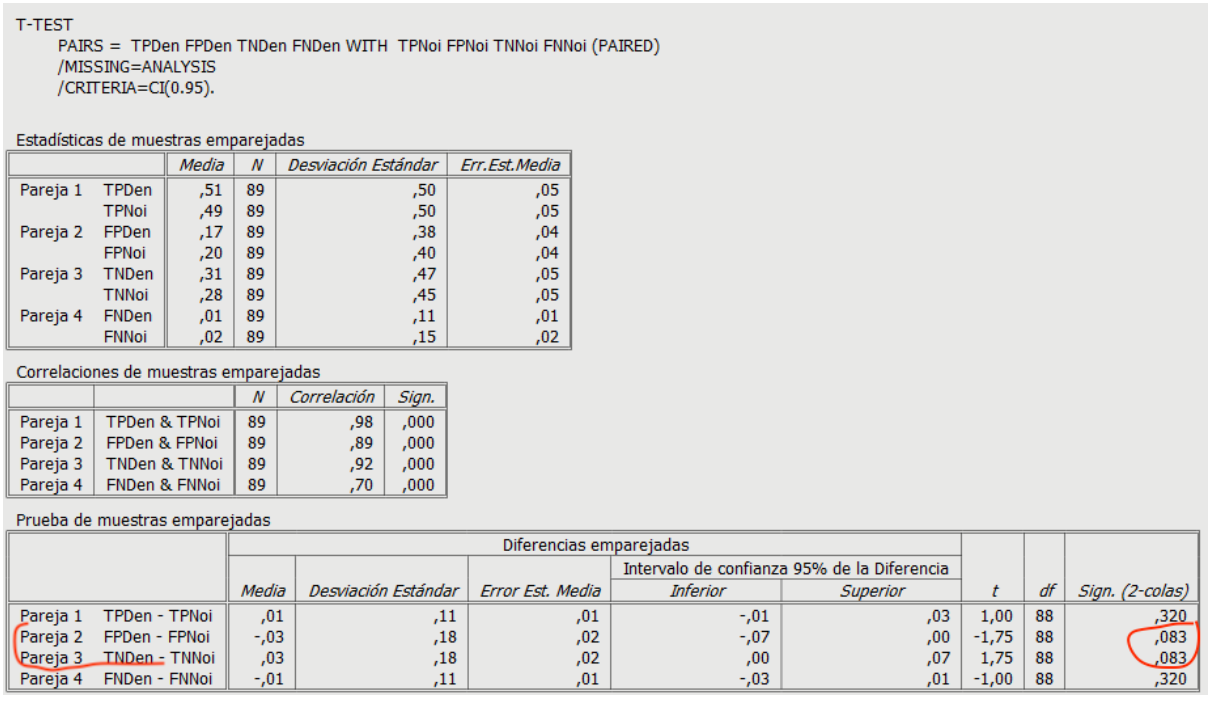
As an extra result, the comparison with Bonomi's results is made. It has already been considered that these two results are not comparable because of the differences on the classification process. Since the main problem for Bonomi was the ‘low’ sensitivity results, these same values are compared with the ones from Bonomi, summarized in the next table.

| Arrhythmia | Sensitivity results (%) | | | | |
|---|---|---|---|---|---|
| | Bonomi | Base | Caballero | Chengyu | Sardar |
| Tachycardia | 85,000 | 97,826 | 89,130 | 93,478 | 93,478 |
| Bradycardia | 89,000 | 90,840 | 95,420 | 99,237 | 100,000 |

**Table 4.22. Comparison of final results with Bonomi's.**

#### 4.6.4.3. Statistical tests

In order to reject or accept the alternative hypothesis stated in the first part of this evaluation, and given the quantitative character of the evaluation variables, a statistical test is proposed.
The tool used for this statistical test in SPSS, which is a simple and intuitive tool for making the t-paired test analysis, since we want to see the differences in the means before and after making the denoising process in the arrhythmia detection task.
The procedure for calculating this test was inputting the confusion matrices results (TP, FP, TN, and FN) before and after the denoising process, in other words, a comparison between the results in the noisy conditions and the results in the denoised conditions is made.
This test is run for both tachycardia and bradycardia results in each one of the detection scripts declared all along the former sections.
Results in *Figure 4.24.* show an acceptable significance level for the base script results in terms of FP and TN.



```
T-TEST
    PAIRS =  TPDen FPDen TNDen FNDen WITH  TPNoi FPNoi TNNoi FNNoi (PAIRED)
    /MISSING=ANALYSIS
    /CRITERIA=CI(0.95).
```

Estadísticas de muestras emparejadas

| | | Media | N | Desviación Estándar | Err.Est.Media |
|---|---|---|---|---|---|
| Pareja 1 | TPDen | ,51 | 89 | ,50 | ,05 |
| | TPNoi | ,49 | 89 | ,50 | ,05 |
| Pareja 2 | FPDen | ,17 | 89 | ,38 | ,04 |
| | FPNoi | ,20 | 89 | ,40 | ,04 |
| Pareja 3 | TNDen | ,31 | 89 | ,47 | ,05 |
| | TNNoi | ,28 | 89 | ,45 | ,05 |
| Pareja 4 | FNDen | ,01 | 89 | ,11 | ,01 |
| | FNNoi | ,02 | 89 | ,15 | ,02 |

Correlaciones de muestras emparejadas

| | | N | Correlación | Sign. |
|---|---|---|---|---|
| Pareja 1 | TPDen & TPNoi | 89 | ,98 | ,000 |
| Pareja 2 | FPDen & FPNoi | 89 | ,89 | ,000 |
| Pareja 3 | TNDen & TNNoi | 89 | ,92 | ,000 |
| Pareja 4 | FNDen & FNNoi | 89 | ,70 | ,000 |

Prueba de muestras emparejadas

| | | Diferencias emparejadas | | | | | t | df | Sign. (2-colas) |
|---|---|---|---|---|---|---|---|---|---|
| | | Media | Desviación Estándar | Error Est. Media | Intervalo de confianza 95% de la Diferencia Inferior | Superior | | | |
| Pareja 1 | TPDen - TPNoi | ,01 | ,11 | ,01 | -,01 | ,03 | 1,00 | 88 | ,320 |
| Pareja 2 | FPDen - FPNoi | -,03 | ,18 | ,02 | -,07 | ,00 | -1,75 | 88 | ,083 |
| Pareja 3 | TNDen - TNNoi | ,03 | ,18 | ,02 | ,00 | ,07 | 1,75 | 88 | ,083 |
| Pareja 4 | FNDen - FNNoi | -,01 | ,11 | ,01 | -,03 | ,01 | -1,00 | 88 | ,320 |

**Figure 4.24. Results for T-paired test for bradycardia detection in the Base script.**

T-TEST

T-TEST
PAIRS = TPDen FPDen TNDen FNDen WITH TPNoi FPNoi TNNoi FNNoi (PAIRED)
/MISSING=ANALYSIS
/CRITERIA=CI(0.95).

Estadísticas de muestras emparejadas

| | | Media | N | Desviación Estándar | Err.Est.Media |
|---|---|---|---|---|---|
| Pareja 1 | TPDen | ,93 | 140 | ,26 | ,02 |
| | TPNoi | ,94 | 140 | ,25 | ,02 |
| Pareja 2 | FPDen | ,01 | 140 | ,12 | ,01 |
| | FPNoi | ,04 | 140 | ,20 | ,02 |
| Pareja 3 | TNDen | ,05 | 140 | ,22 | ,02 |
| | TNNoi | ,02 | 140 | ,15 | ,01 |
| Pareja 4 | FNDen | ,01 | 140 | ,08 | ,01 |
| | FNNoi | ,00 | 140 | ,00 | ,00 |

Correlaciones de muestras emparejadas

| | | N | Correlación | Sign. |
|---|---|---|---|---|
| Pareja 1 | TPDen & TPNoi | 140 | ,95 | ,000 |
| Pareja 2 | FPDen & FPNoi | 140 | ,57 | ,000 |
| Pareja 3 | TNDen & TNNoi | 140 | ,65 | ,000 |
| Pareja 4 | FNDen & FNNoi | 140 | NaN | NaN |

Prueba de muestras emparejadas

| | | Diferencias emparejadas | | | | | t | df | Sign. (2-colas) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | Intervalo de confianza 95% de la Diferencia | | | | |
| | | Media | Desviación Estándar | Error Est. Media | Inferior | Superior | | | |
| Pareja 1 | TPDen - TPNoi | -,01 | ,08 | ,01 | -,02 | ,01 | -1,00 | 139 | ,319 |
| Pareja 2 | FPDen - FPNoi | -,03 | ,17 | ,01 | -,06 | ,00 | -2,02 | 139 | ,045 |
| Pareja 3 | TNDen - TNNoi | ,03 | ,17 | ,01 | ,00 | ,06 | 2,02 | 139 | ,045 |
| Pareja 4 | FNDen - FNNoi | ,01 | ,08 | ,01 | -,01 | ,02 | 1,00 | 139 | ,319 |

**Figure 4.24. Results for T-paired test for tachycardia detection in Chengyu script.**

*Figure 4.23* evinces that the bradycardia detection reaches a significance level under 0.1 for the t-paired test. In other words, a 90% of certainty that these results are reproducible for a larger population is obtained.

*Figure 4.24* indicates that tachycardia detection reaches a higher significance level, with a value lower than 0.05.

Therefore, in the specific bradycardia detection case within the base script, we can reject the null hypothesis and accept the alternative one, as well as in Chengyu script regarding tachycardia.

Both these results are obtained for the variables that allow calculating specificity, FP, and TN. Then we conclude that the mechanism allows the improvement on tachycardia and bradycardia estimation for this specific two conditions: the base mechanism allows to obtain significant improvements over bradycardia detection, and the Chengyu mechanism allows to obtain significant improvements over tachycardia detection.

With this in mind, we can infer from this results that these mechanisms would serve as a suitable separator of the population that does not have the arrhythmia. A level of certainty of 90% in the estimation of bradycardia (with the base mechanism) and a level of certainty of 95% in the estimation of the tachycardia (with the mechanism of Chengyu).

### 4.6.5. Discussion

#### 4.6.5.1. Discussion of the simulation of low-intensity noise over tachycardia and bradycardia records

The simulation of a noisy scenario was the first objective of this evaluation. From *figures 4.12* and *4.13* it is possible to evidence the real variation that the noise addition and preprocessing removal induce to the decision variables (in those cases, the heart rate decision variables for each arrhythmia). At some points, the noise makes the heart rate variables to change and surpass the threshold established by the competition organizers. This would infer a change in the classification outcome, as we would expect. However, *figures 4.10* and *4.11* suggest that the general outcome deviation for the classification is not solely affected by heart rate variables. The heart rate may vary, but if the quality index turns to be fine for that record, the general classification remains unchanged.

Notice that the change of heart rate variability calculated for each arrhythmia does not directly indicates a decrement of the signal quality index. The heart rates for each arrhythmia just take into account the number of detected onset pulses (regarding script base and Caballero). The quality index makes a comparison of the correlation with a template clean signal and thus, calculates a general estimate of how much the signal differs from the templates. In other words, peaks might be induced or removed, but in general terms, the entire signal taken into consideration for the classification doesn't show great corruption. Additional results made from simulation show a correlation of 0,8186 and 0,8827 for the calculated heart rates for bradycardia and tachycardia, respectively. Whereas for the signal quality index of both these arrhythmias, correlation results show a lower bias with 0.9530 for bradycardia and 0.9119 for tachycardia results. This is an expected result since the damaging of the signals by low-intensity movement artifacts is slight. Indeed, *figure 4.15* evidences the amplitude differences and thus, the small contribution of the noise over for damaging the signal quality index can be inferred

Results in *tables 4.5* and *4.6* expose the difficulty that was thought of beforehand: it is easy to generate a damaging of the heart rates. However, for general classification results, it is difficult for the noise to damage the sensitivity and specificity. This is the case for both specificities in base script or sensitivity and specificity for tachycardia in Miguel Caballero script.

An interesting outcome can be seen in the base script case, in *Table 4.5* where the noisy results obtain a greater sensitivity than the original outcomes. This can be explained by the fact that the retrieved population for tachycardia classification is greatly imbalanced (this can be referred to in *Table 1* from *Appendix F*). The classification outcomes of the confusion matrix can be seen as a conditional probability. Then, these are strongly affected by a priori probability of the real population or the gold standard population. The classifier may change the results of the confusion matrix, but this should not always be a negative impact. At least, in this particular case, the actual negative population is considerably smaller than the actual positive population (only 9 out of 140 registers). Once the noise reaches to make greater high heart rate estimates than originals, the change in the classification outcomes is more prone to fall into a truly positive register rather than in a truly negative register (as we only have nine of them). The great a priori

probability for the actual tachycardia registers (approximately 0.9357) makes the conditional probability turn in favor of the true tachycardia registers.
For bradycardia in base script, it can be seen that this measure diminishes its value due to the decrement of one TP and an increment of one FN.
Miguel Caballero results expose that noise cannot vary results in tachycardia, but it does it well in bradycardia, where occurs a tradeoff between sensitivity and specificity. The sensitivity decreases, caused by an increase of three FN while TP decreased by three as well; the specificity increases caused by a decrease of two FP and an increase of the same magnitude in TN.
In the case of tachycardia, none of the results vary, which is a concerning result given that the same dataset and the same registers as base script are used. However, a deeper insight into Caballero script allows to see that the author of this algorithm changes the threshold for tachycardia (TACH_THR in *figure 4.11*) final determination. The reasons for this change are uncertain, nevertheless, the author claims that the logical decision tree that he obtained allowed to make a better and more robust estimation of the arrhythmia.
Chengyu results suggest a tradeoff between bradycardia sensitivity and specificity, while tachycardia sensitivity remains unchanged. However, and taking into account the sharp decrease of tachycardia specificity, it makes us inquire that the removal of both filters for both of the signals is forcing the classifier into a trivial case for tachycardia.
Lastly, Ansari results show a decrease over almost all parameters, except for tachycardia specificity. Results in bradycardia suggest that the removal of both filters and addition of the noise to pulsatile signals can decrement both sensitivity and specificity results, in other words, a general worsening of the classifier for this arrhythmia.
At the end of this first analysis, it is possible to appreciate that the low intensity noise induces changes in the decision variables. Each one depending on the conditions established by the authors of the classification scripts. Those which only use pulsatile signals are difficult to contaminate or distort in terms of tachycardia records. The high heart rates calculated for tachycardia were quite variable, however, the signal quality index didn't vary as much. For this reason, the general result variation is not so remarkable. Those which used ECG signals as well are also distorted because of the filter removal for these other signals as well.
The simulation of the same scenario found by Bonomi led us to prove that the heart rates estimations were sometimes distorted. These first results back up the assertion from his study: the fact that even in a rest environment, low intensity movements reach to distort the heart rate values. Even if this noise doesn't demonstrate a great power level. Because of this, we suggest to take into consideration these kind of artifacts in ambulatory scenarios. In our work's case and due to the classification logic used in the scripts, results were not so easily distorted. But in the case that the only variable for deciding the existence of an arrhythmia was the heart rate, it should be expected that the classifier commits more mistakes.

#### 4.6.5.2. Discussion of the results for the best denoiser case after adding low intensity noise

The purpose of this second test was to prove that the low intensity noise induced in a measurement environment for tachycardia and bradycardia can be removed using different techniques. The improvement of this noisy condition results was proposed by assembling a mechanism which consisted of two parts: a denoising part and a detection part.
Our data proved that this improvement can be reached, depending on many factors, which were reviewed along the evaluation procedure. Firstly, it should be noticed that the characterized and used noise for these tests is a low intensity noise. With this in mind, significant changes in sensitivities and specificities after summing it, were not expected. Different kind of noises (noises with different variances) were summed, causing that some signals were distorted in a greater quantity than others. We do not know with certainty what noise is being added to which signal, we only know that a set of noises are being added randomly to different signals.
Second, we can see that several of the denoising methods did not show positive results. Then we can state that there doesn't exist a particular denoising method for preprocessing a signal before a posterior procedure. Literature from the review in *section 4.3* exposed various articles which compared different denoising methods for a specific signal: the PPG signal. This also affirms that there is no specific method for all applications, as it depends on both the data used for the test and the purpose of the posterior test.
In this case, a trial and error approach was made, which sought to find a determined denoiser for each script and thus conform the mechanism. Each author extracted different features from the signals. Then, we expect that the best denoiser will be the one that delivers the PPG signal with the characteristics desired by each arrhythmia detector to make a later decision.
The algorithms which use only pulsatile signals are those who have a lower scoring. The use of only one signal for diagnosis throw less accurate results than the use of two, as expected.
However, their sensitivity and specificity results have improved with respect to those obtained with noise. Some have even increased these results over the ones obtained originally. This seems to indicate that the preprocessing method used for the pulsatile signals in these algorithms were not considering a proper bandwidth for the sake of the PPG signal. Both base and Caballero's scripts used a simple notch low pass filter, which was more intended for the isolation of the powerline interference at 50 Hz and the frequencies beyond. Using a greater granularity over the bandwidth, it was possible to see that wavelets denoising can be considered as a convenient method for this task, either with a fixed reconstruction or using soft thresholding.
The best denoising case for bradycardia detection in base script generates an improvement on both parameters. A specificity result 6,97% better than the original was obtained. It is also the case of sensitivity in tachycardia, where original results are outperformed as well in 3,81%.
About Miguel Caballero's case, the improvement of sensitivity over noisy signals is accomplished for both arrhythmias. Tachycardia sensitivity reaches a 0,76% improvement over the original results.

The improvement over original results suggest that the best denoising method for this two first mechanisms reaches to deliver more valuable information or desired features of the PPG signal than the filtering method used originally by the author. It is possible to reconstruct a signal accurately after being contaminated with this low intensity noise, thus obtaining a better quality of the signal received by the arrhythmia detector.

Better results were obtained using both ECG and ABP/PPG signals, because of the greater information provided by the extra signals.

In Chengyu results, there was an accomplishment of the sensitivity improvement just for bradycardia, concerning the noisy values. In tachycardias, denoisers generated a good performance, differing from original results in only -0,76%. Reaching a sensitivity result of 99,24% after denoising can be seen as a suitable output for a classification result. An improvement of 4,65% in the specificity of bradycardias from original results was obtained as well.

Sardar case also accomplished the sensitivity improvement of both arrhythmias over noisy results. Regarding specificity results, one increased (tachycardia), and the other is held constant (bradycardia). With the best denoiser case, it is possible to obtain a 100% sensitivity in tachycardias, equaling original results. Also, in bradycardia sensitivity, a little improvement of 2,174% over the original results was reached.

A common factor between all these mechanisms is the use of wavelets as the best denoising method for the pulsatile signals. The improved performance parameters were obtained in two scripts with symlet 4, while symlet 6, daubechies 8 and daubechies 10 wavelets obtained the best results for one script each.

Both the base and Chengyu scripts obtained the best results by using sym4 wavelet, and Caballero obtained the best results using symlet 6 and daubechies 8. One thing that is common to the three first scripts is that regardless of the wavelet family used and the wavelet level, all of the three used a decomposition-reconstruction level of 3. A frequency analysis over the clean signals of the new dataset allow to evidence that the pulsatile signal information is more concentrated within the (0-7) Hz bandwidth, which is approximately the same interval reconstructed by Mallat's algorithm when the reconstruction level is chosen is 3. Looking into the *Figure 4.4*., if we divide the sampling frequency used for the signals between $2^{N+1}$, where *N* relates to the decomposition level, we obtain a frequency interval of approximately (0 - 7,81) Hz.

In the classifier case for Ansari, the denoising reconstruction level for the best results is 4, which indicates that the frequency band reconstructed would be approximately (0-3,9) Hz. This could be explained by the fact that this author cuts the signal to a really little time window before making the preprocessing. This cut of the signal generates a weaker spectrum, where the information comprises even more in the lowest frequencies.

Better results were obtained in general for all the mechanisms, demonstrating that the noise conditions can be improved if the right choice of an appropriate denoiser for the detectors is made. However, this can be a cumbersome and exhausting job, suggesting that future work towards wavelets denoising techniques over PPG signals should be conducted. In this case, results can be primarily focused towards the wavelets symlets, as a suggestion to clean and collect wise PPG signal characteristics. The levels of deconstruction and reconstruction 3 and

4 obtain a good performance for PPG signals in our case, depending on the subsequent processing that we want to do on them. However, these levels of deconstruction are not generalizable, since the Mallat algorithm takes into account the sampling frequency used, and therefore, it can change from study to study.

The thresholding method included in the analysis obtained good results, sometimes better than some of the fixed reconstruction cases. However, the best of the wavelet thresholding case could not surpass the best of the fixed reconstruction case. The deconstruction levels for thresholding were intentionally taken as nine, so a deeper differentiation of the detail's frequency sub bands could be executed by the soft thresholding command. Nevertheless, results indicate that maybe in the process of rescuing only the valuable information from the noise, some of the important frequency components were erased since the noise overlaps with the signal spectra. It is taken as future work to investigate more in this technique and its extents. The new research path could be either using hard thresholding or executing thresholding into the approximation levels as well. Wavelet packets are also seen as a future investigation for these kinds of signals. From *table 4.15*, it can be seen that the thresholding was close to obtain the best results. In a similar way it happens with the other scripts, generating a great improvement in some results of the sensitivity and specificity, but in turn generating a tiny decrease in some other result, reason why it could not be chosen. A greater rigor in the variables used for this thresholding could have obtained better results, even the best.

Statistical tests allowed us to confirm the assumption that was stated previously in this same discussion. The use of different and complementary signals along with the pulsatile ones result in a greater improvement after being denoised, as it can be seen from Chengyu significance results. If the condition of noised ECG signals was fulfilled, then even greater changes (worse performance) in the noisy results should be expected. The differences between noisy and denoised conditions would be more extensive, expecting the obtention of a higher significance level.

Besides this considerations, significant results of the improvements could be seen in the specificity, although the sensitivity also obtained improvements in all the scripts for both arrhythmias (except for Chengyu in the tachycardia). Significance results were mostly obtained for the variables that handle specificity results, thus showing that the detection allows to be more selective about the patients that do not have the arrhythmia condition.

As an additional test, we would like to mention the application of hybrid wavelets denoising over the signals. The best denoising case was partitioned into two different classification issues: one for bradycardia and one for tachycardia. This approach was only proved for base and Caballero scripts, since we have the certainty that the classification is made only with PPG signals. Bradycardia records maintained the same wavelet denoising case, while tachycardia changed to another case. *Appendix C* contains the results for these approaches, were ROC curves are also displayed. This enables us to conclude that even greater results could be reached with a finest granularity regarding the arrhythmia case. Since time constraints are exceeded for this type of approach, we leave this conclusion as the base for future works conducted in the same direction.

### 4.6.6. Conclusions

The application of ten denoising techniques over different brady- and tachycardia classification algorithms allows to conclude that there is not a defined denoiser for all of them. This decision even varies depending on the procedure that the classification scripts make internally.

We sought for a denoiser-classifier combination which produced an improvement over sensitivities in the detection of two specific arrhythmias. Results thrown a coincidence over the MRA using the DWT, applying symlet family in the majority of the cases. In two cases, the best results were obtained using daubechies family wavelets.

As an answer to the research question made in the evaluation definition, we state that the wavelet denoising contribute to achieve an improvement over the results obtained with noise. Significance tests point to better results obtained for specificity. However, sensitivity was also improved as it was evidenced by the AUC results, which allow comprising the information of the quality of each classifier.

Bonomi's results were compared to ours. Although his sensitivity results were acceptable, we highlight his excellent specificity results. This might be also a reason for his sensitivity values, because it's highly improbable that a classifier reaches such great results for both measures.

Another important fact to take into account is that the sole use of pulsatile signals can perform well into a classification problem, but the results are reinforced when other kinds of complementary signals such as ECG is used. The results are much improved, showing that a combination of both signals and the correct denoising process over them would help significantly arrhythmia detection. Nevertheless, as the future context for this work is the use of the non-invasive signals obtention methods for an ambulatory environment, ECG signals are inconvenient for this kind of scenarios and again, all the detection responsibility would lay onto the pulsatile signals.

Further analysis is needed on why some results for other denoisers threw unexpected results, but the causes might be the signal points that every classification script has into account, however, this is marked as future work because of time constraints.

# CHAPTER 5. CONCLUSIONS AND OUTLOOK.

## 5.1. Conclusions

A first state of the art review was conducted and summarized in *Table 2.1.* This table presents a wide set of arrhythmia detection techniques. However, some of these do not consider the effect of the movement artifacts over PPG signals. The majority of the studies which worked under noise conditions are oriented towards a heart rate estimation using ECG signal as the gold standard. Few works were encountered to detect arrhythmias, yet these are not within the ones focused by us: tachycardia and bradycardia. This sets a current gap in the analyzed studies: there is a lack of studies that perform an estimation of this specific arrhythmias under noise conditions.

Only one relevant study was found on this topic, the author is: A. Bonomi. He carries a study around the bradycardia and tachycardia detection under free-living conditions. His work deals with motion artifacts and makes an estimation of these two arrhythmias at the same time. However, Bonomi's results show a limitation: the sensitivity of the brady- and tachycardias estimation was affected by low intensity movements induced by patients.

We conclude in this first part of the job that there is a shortage of studies conducted towards the brady- and tachycardia diagnosis under low intensity movement artifacts. Bonomi's dataset, which allowed the study of movement artifacts, is confidential (as most of the other literature datasets). Thus, a major limitation for this study is the lack of experimental (field work) data.

As far as we know, there are very few datasets of PPG signals under noise conditions which could help us in this task. This is backed up by other results in this very same state of the art review. The majority of the conducted works carried out their own measurements. This restraint is understandable because of the many factors that must be taken into account to obtain proper results. As a result of this limitation, we propose to generate a similar dataset as the one used by Bonomi. This analysis entails the second part of this thesis work.

The generation of a dataset similar to the one from Bonomi was needed. If we obtain this, then we can analyze the possibility of proving that these low intensity noise condition can be improved.

To do so, the obtention of a noise model was planned, for the posterior contamination of the signals containing the specific arrhythmias. A second state of the art review exhibited the lack of pertinent works in this area. Solosenko names in his study the need of PPG signals datasets that contain arrhythmia events as well. Even if brady- and tachycardias are not present in his results, this author demonstrates that it is possible to model PPG signals with various cardiopathies features, using ECG signals as the base. The only author who developed a study towards noise modeling is Wartzek, with his artifact modeling for capacitive ECG signals.

Given the small set of data we disposed to make a suitable characterization for the PPG movement artifacts noise, it is decided to take some suggestions from his studies in order to create our own noise model.

Later on, the characterization processes for the noise were carried out. These processes were experimentally tested, because there is not a reference study which allowed to establish a

defined process for the characterization of the noise at determined activities. Most of the studies which generated a purposely distortion over biomedical signals, considered white noise as a good estimation. But we think this white noise approach might be resembling all kinds of artifacts affecting these signals. The procedures or approximations towards noise characterization were empirical. Several hints were taken from the limited results from the bibliographic review about noise characterization processes. A noise model that fits the desired results is obtained, having as the gold standard the ECG signal. Here, we state the major conclusion over this second part of the work, settled in the generation of a dataset containing arrhythmias and low intensity movements. A new dataset with the desired simulated data can be generated through a noise characterization process.

We faced a difficulty with this noise modeling process, as the topic has not been widely explored and the information volume is reduced. However, our approach can be seen as a contribution for future study fields, were noisy datasets are needed for a denoiser validation procedure. Others could even think of an accuracy validation on the arrhythmia detection methods under noise conditions. Precisely, this last action is pursued in the third part of this work, but the focus is put into the brady- and tachycardia signals.

In the third and last part of this work, the objective is set in the improvement of the brady- and tachycardias detection under these low intensity movement artifacts. An important factor is highlighted here: these arrhythmias are detected in rest conditions. For example, a tachycardia condition can be easily confused in an exercise environment. This is the reason why Bonomi cuts parts of the signal by looking at the accelerometer signal threshold. As the contamination of our available signals is done with these low intensity artifacts, we don't need any clipping in this case.

A third state of the art review was conducted and summarized in *Table 4.1*. The most remarkable technique within this review is wavelet denoising. Other methods like adaptive filtering or smoothing filtering are also noticed. This last review along with the review from the challenge outcomes take us to the next step of this work.

After the denoisers application over the arrhythmia detectors, we get quite variable results. Further steps into the task of finding a determined and specific method for denoising these signals should be taken. Our work proved that wavelets are desirable techniques due to their versatility and wide range of applications. An advantage is seen as they don't define a specific wavelet for the application, instead, the application is first defined and then the wavelet is chosen.

Wavelets show better performance as compared to other denoisers. The wavelet family which exhibited the majority of best results for this specific application were symlets. We found the best decomposition levels at a 125 Hz sampling frequency for PPG signals are 3 and 4. This configuration is variable taking into account the length of the signal and the sampling frequency used for the signal acquirement. These two aspects affect the spectra distribution. For this reason, the wavelet parameters for this specific work could not be generalizable to all of the studies. Each study must have its own considerations and margins.

Thresholding methods are recommended too, since very good results were obtained compared with the best results of each script. Also, the reconstruction for this technique is desired as it deals with the noise in a more suitable and automatic way. However, since the noise

frequencies overlap the signal's frequencies, thresholding could have distinguished a main portion of the signal spectra.
The wavelet choosing took parts of the literature review and our suggestions. Thus, this comparative evaluation can be recognized as one of the first steps for identifying possible wavelet families that accomplish a good performance regarding PPG signal denoising.
Final results were visible thanks to ROC curves, confusion matrices and AUC final estimation. We can close up this last section with the conclusion of this part. It is possible to improve the performance of the brady- and tachycardia classification when PPG signals are affected by low intensity movements. The denoising of this seemingly unimportant kind of artifacts produces more accurate results, without the needing of cutting all of the signal parts which contain movement.
Wavelets are exposed as an adaptable and useful method, that demonstrates a wide range of possibilities for obtaining good results. Not only over these signals, but we presume that these have already been used over other type of biomedical signals. Because of the broad range of functions that it possesses, it is also deduced that wavelets present multiple solutions for the PPG signal analysis in more applications than just denoising.
An improvement over the noisy results was achieved and some of them demonstrated a better performance than original results.
This means that it is possible to return desired features of the signal, taking into account the subsequent process carried with the PPG signal. Also, we infer that the denoisers used in the original scenario could have shown better results if a more precise approach over the frequency band of the pulsatile signals were made. However, we cannot penalize the original results from the participants of the challenge, because they supposed these were clean signals already. The detection of these two arrhythmias under acceptable noise conditions in an ambulatory scenario is seen as an interesting field for future works.
Significance results demonstrated that the correct selectivity of patient groups with these cardiopathies could be obtained. Two specific mechanisms acted as a good estimator of the population that does not possess the illness. Of course, the sensitivity was also a desired indicator since it is the one that denotes the population that actually has the illness. However, this particular work couldn't meet the certainty percentage needed for generalizing the sensitivity improvements to a broader population. Although we declare that selectivity is also important, because the patient will be sure of their condition and then won't have to spend on any unnecessary treatments. It can be seen as a facilitator for the medical specialist to discriminate both cases of the populations.
Nevertheless, we propose a future work with a real case assessment of these mechanisms, because of the simulated environment that the tests were run on.
Finally, we answer the research question stated in *Chapter 1* of this document. The wavelet based denoising is the best option to improve sensitivity of bradycardia and tachycardia detection under low-intensity movements. We are aware of the existence of more methods for this purpose, but the granularity that wavelets provide and the multiple possibilities of choice, transforms them into the best choice for this occasion.

## 5.2. Contributions

With the development of this present work, the following points are marked as contributions:

- Article titled “Modeling of motion artifacts on PPG signals for heart-monitoring using wearable devices” presented to SIPAIM event: 15th International Symposium on Medical Information Processing and Analysis 2019. This event will be carried out in Medellín city, Colombia, from 6 to 8 of November of 2019.
- Article titled “Tachycardia and Bradycardia Detection using Photoplethysmography under low-intensity Motion Artifacts” sent to Sensors MDPI journal under the special issue of “Wireless body sensors”.
- An approach to the use of an alternative electro-optical method for sensing cardiac signals and the further diagnosis of cardiopathies on the basis of this signals, promoting less-invasive measurement methods than other more invasive as ECG signals.
- The development of a noise model, extracted for characterizing low intensity activities, which is completely variable, allowing to generate a different noise model using the variance factor[8].
- An approach to a more general noise model, which is completely random and ruled by random variables. In this same way, five possible bases for the creation of noise models with a greater sophistication.
- A database containing arrhythmia events and movement artifacts due to low-intensity movements.
- A mechanism for the removal of induced artifacts due to low-intensity movements in the measurement and detection of brady- and tachycardias (in a simulated environment).
- An experimental evaluation process, collecting data from 4 different arrhythmia detectors under the influence of 10 denoising cases[9][10][11][12]. This evaluation serves as the base for future works in the area, which are focused towards the utilization of preprocessing techniques over PPG signals.
- Relationship bonding with the Valle de Pubenza clinic, especially with Dr. Nelson Muñoz, who guided the process of establishing the requirements for the correct planning of the scope of the present work.

[8] The repository with the codes for the noise models generation can be found in: https://github.com/sebasmos/PPGpeakDetection1

[9] The repository with the codes for the base script mechanisms can be found in: https://github.com/malelang/ScriptBase

[10] The repository with the codes for the Miguel Caballero script mechanisms can be found in: https://github.com/sebasmos/Script_Miguel_Caballero

[11] The repository with the codes for the base script mechanisms can be found in: https://github.com/sebasmos/Multi_Feature-Fusion-Method--Chenyu

[12] The repository with the codes for the base script mechanisms can be found in: https://github.com/malelang/Script_Ansari

### 5.3. Outlook

The main contribution of this work is a mechanism that allows the reduction of noise due to low intensity movements and posterior detection of brady- and tachycardias over PPG signals. Having in mind the current limitations and observations along the process of this thesis, the following future works are contemplated:

- Implementing pattern recognition techniques over noise signals produced by quasi-periodic movements, for example: the arm movement at the moment of walking or jogging. In this way a better characterization and posterior modeling of the artifacts produced by these movements is made, along with the collaboration of an accelerometer or gyroscope signal, if possible.
- Inquire into the alternative of generating a gaussian noise model which can resemble in certain measure the noise produced by determined activities or movements in PPG signals. Use a greater quantity of realizations of noise and thus, a greater quantity of people who execute these movements under the proper conditions for the measurement. Use also the information provided by accelerometer, maybe
- Inquire into the possibility of generating models of accelerometry signals when executing certain quasi-periodic movements or else, determined movements of short lasting, such as gesticulation, rise and lower the hand, moving horizontally or vertically the forearm, etc. By generating this accelerometry models, it would exist also a way of recognize patterns in future studies that need to make a movement discrimination and thus, recognize the possible activities that the patient has been making in free-living conditions or an ambulatory measurement scenario, where they are not being monitored by hospital staff.
- Explore different techniques for the PPG signal denoising which could not be executed in this work, and use other kind of signals as a helper for denoising, as an example, the design of adaptive filters.
- Integrate another kind of signals to the tachycardia and bradycardia detection process, by having then a more valuable information quantity. Some of these can be non-invasive ECG and respiratory.
- Explore and indague the denoising possibilities through other wavelet families. Also, dig deeper into the thresholding method, the penalization methods within this thresholding or using other wavelet modalities of frequencies decomposition, for example: wavelet packets.
- Integrate the mechanisms algorithm codes which obtained the best results to and application that allows to generate an anomalous cardiac rhythms evaluation, for people who suffer from these conditions. Send this information through wireless technologies from the wearable to a mobile device, thus the detection of tachycardias and bradycardias can be made into determined time intervals or windows. This could serve as a support for the assessment and diagnosis of cardiopathies carried by a health professional

# REFERENCES


[1] "About Arrhythmia", Heart.org, 2018. [Online]. Available: http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/About-Arrhythmia_UCM_002010_Article.jsp#.Wx1_WVMvxQI.

[2] E. Benjamin, et al, "Heart Disease and Stroke Statistics—2017 Update: A Report From the American Heart Association", Circulation, vol. 135, no. 10, pp. e146-e603, 2017.

[3] K. Kearley, M. Selwood, A. Van den Bruel, M. Thompson, D. Mant, F. Hobbs, D. Fitzmaurice and C. Heneghan, "Triage tests for identifying atrial fibrillation in primary care: a diagnostic accuracy study comparing single-lead ECG and modified BP monitors", BMJ Open, vol. 4, no. 5, p. e004565, 2014.

[4] L. Pava-Molano and P. Perafán-Bautista, "Generalidades de la fibrilación auricular", Revista Colombiana de Cardiología, vol. 23, pp. 5-8, 2016. Available: 10.1016/j.rccar.2016.10.003.

[5] "What is Atrial Fibrillation (AFib or AF)?", Heart.org, 2018. [Online]. Available: http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/What-is-Atrial-Fibrillation-AFib-or-AF_UCM_423748_Article.jsp#.WkxCqlQ-eu4.

[6] J. Díaz-Martínez, M. Duque-Ramírez, J. Marín-Velásquez, J. Aristizábal-Aristizábal, J. Velásquez-Vélez and W. Uribe-Arango, "Costos asociados a la fibrilación auricular", Revista Colombiana de Cardiología, vol. 23, pp. 192-197, 2016.

[7] H. Leutheuser, S. Gradl, P. Kugler, L. Anneken, M. Arnold, S. Achenbach and B. Eskofier, "Comparison of real-time classification systems for arrhythmia detection on Android-based mobile devices", 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014.

[8] S. Shan, S. Tang, P. Huang, Y. Lin, W. Huang, D. Lai and A. Wu, "Reliable PPG-based algorithm in atrial fibrillation detection", 2016 IEEE Biomedical Circuits and Systems Conference (BioCAS), 2016.

[9] A. KAVSAOĞLU, K. POLAT and M. BOZKURT, "An innovative peak detection algorithm for photoplethysmography signals: an adaptive segmentation method", TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES, vol. 24, pp. 1782-1796, 2016.

[10] D. McDuff, S. Gontarek and R. Picard, "Remote Detection of Photoplethysmographic Systolic and Diastolic Peaks Using a Digital Camera", IEEE Transactions on Biomedical Engineering, vol. 61, no. 12, pp. 2948-2954, 2014.

[11] S. Kuntamalla and L. Ram Gopal Reddy, "An Efficient and Automatic Systolic Peak Detection Algorithm for Photoplethysmographic Signals", International Journal of Computer Applications, vol. 97, no. 19, pp. 18-23, 2014.

[12] A. Sološenko, A. Petrėnas, V. Marozas and L. Sörnmo, "Modeling of the photoplethysmogram during atrial fibrillation", Computers in Biology and Medicine, vol. 81, pp. 130-138, 2017.

[13] A. Bonomi, F. Schipper, L. Eerikainen, J. Margarito, R. Aarts, S. Babaeizadeh, H. de Morree and L. Dekker, "Atrial Fibrillation Detection Using Photoplethysmography and Acceleration Data at the Wrist", 2016 Computing in Cardiology Conference (CinC), 2016.

[14] H. Han and J. Kim, "Artifacts in wearable photoplethysmographs during daily life motions and their reduction with least mean square based active noise cancellation method", Computers in Biology and Medicine, vol. 42, no. 4, pp. 387-393, 2012.
[15] K. Sweeney, T. Ward and S. McLoone, "Artifact Removal in Physiological Signals—Practices and Possibilities", IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 3, pp. 488-500, 2012.
[16] S. Shashikumar, A. Shah, Q. Li, G. Clifford and S. Nemati, "A deep learning approach to monitoring and detecting atrial fibrillation using wearable technology", 2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), 2017.
[17] J. Song, D. Li, X. Ma, G. Teng and J. Wei, "A Robust Dynamic Heart-Rate Detection Algorithm Framework During Intense Physical Activities Using Photoplethysmographic Signals", Sensors, vol. 17, no. 11, p. 2450, 2017.
[18] N. Pradhan, S. Rajan, A. Adler and C. Redpath, "Classification of the quality of wristband-based photoplethysmography signals", 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2017.
[19] A. Bonomi, L. Eerikäinen, F. Schipper, R. Aarts, H. de Morree and L. Dekker, "Detecting Episodes of Brady- and Tachycardia Using Photoplethysmography at the Wrist in Free-living Conditions", 2017 Computing in Cardiology Conference (CinC), 2017.
[20] C. Veiga, D. Rivera, A. Paz, D. Castineira, J. Farina, J. Rodriguez-Andina, E. Garcia and A. Iniguez, "Optimized PPG-based wearable acquisition unit for massive analysis of heart rhythms", 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), 2017.
[21] A. Solosenko, A. Petrenas and V. Marozas, "Photoplethysmography-Based Method for Automatic Detection of Premature Ventricular Contractions", IEEE Transactions on Biomedical Circuits and Systems, vol. 9, no. 5, pp. 662-669, 2015.
[27] P. Chapman, CRISP-DM 1.0. Chicago: SPSS, 2000.
[29] Z. Lin, "Heart rate estimation using wrist-acquired photoplethysmography under different types of daily life motion artifact." 2015 IEEE ICC SAC- Communications for E- health, 2015.
[30] D. Jarchi and A. Casson, "Description of a Database Containing Wrist PPG Signals Recorded during Physical Exercise with Both Accelerometer and Gyroscope Measures of Motion," Data, vol. 2, no. 1, p.1,2016.
[33] "Bradycardia: Slow Heart Rate", www.heart.org, 2018. [Online]. Available: https://www.heart.org/en/health-topics/arrhythmia/about-arrhythmia/bradycardia--slow-heart-rate
[35] "Taquicardia - Síntomas y causas - Mayo Clinic"; Mayoclinic.org, 2018. [Online]. Available: https://www.mayoclinic.org/es-es/diseases-conditions/tachycardia/symptoms-causes/syc-20355127.
[22] Bumgarner JM, Lambert CT, Hussein AA, Cantillon DJ, Baranowski B, Wolski K, Lindsay BD, Wazni OM, Tarakji KG, Automated Atrial Fibrillation Detection Algorithm Using Smartwatch Technology, Journal of the American College of Cardiology (2018), doi: 10.1016/j.jacc.2018.03.003.
[23] Pradhan, N., Rajan, S. and Adler, A. (2017). "Classification of the Quality of Wristband-based

Photoplethysmography Signals". 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2017.
[24] L. M. Eerikanen, L. Dekker, A. G. Bonomi, R. Vullings, F. Schipper, J. Margarito, H. M. D. Morree, and R. M. Aarts, "Validating Features for Atrial Fibrillation Detection from Photoplethysmogram under Hospital and Free-living Conditions," 2017 Computing in Cardiology Conference (CinC), 2017.
[25] A. J. Casson, A. V. Galvez, and D. Jarchi, "Gyroscope vs. accelerometer measurements of motion from wrist PPG during physical exercise," ICT Express, vol. 2, no. 4, pp. 175–179, 2016.
[26] N. S. Paradkar and S. R. Chowdhury, "Fuzzy Entropy based Detection of Tachycardia and Estimation of Pulse Rate through Fingertip Photoplethysmography," Journal of Medical and Bioengineering, vol. 4, no. 1, pp. 19–23, 2015.
[28] A. Solosenko and V. Marozas, "Automatic Premature Ventricular Contraction detection in photoplethysmographic signals," 2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings, 2014.
[31] A. Seyed. "Motion and Noise Artifact Detection and Vital Signal Reconstruction in ECG/PPG based Wearable Devices". Doctoral Dissertations, 2015.
[32] D. Kasper, A. Fauci, S. Hauser, D. Longo, J. Jameson, J. Loscalzo and T. Harrison, Harrison's principles of internal medicine, 18th ed. McGraw Hill, 2013, pp. 864-878.
[36]"El sistema circulatorio", calameo.com, 2019. [Online]. Available: https://es.calameo.com/read/0017664031a8daace9912.
[37]J. Sepúlveda Saavedra and R. Medina Hernández, Histología, 6th ed. Madrid: Mc Graw Hill Interamericana, 2014.
[38] T. Fagan, Lo esencial en sistema cardiovascular, 2nd ed. Madrid: Elsevier, 2007, pp. 3-4.
[39]"Sistema cardiovascular: Anatomía", Infermeravirtual.com, 2019. [Online]. Available: https://www.infermeravirtual.com/files/media/file/100/Sistema%20cardiovascular.pdf?1358605522.
[40]"¿Cuáles son los movimientos del corazón?", aboutespanol, 2019. [Online]. Available: https://www.aboutespanol.com/los-movimientos-del-corazon-sistole-y-diastole-1184993.
[41]M. Gaea Marelle Miranda, "Estructura y función del corazón", News-Medical.net, 2019. [Online]. Available: https://www.news-medical.net/health/Structure-and-Function-of-the-Heart-(Spanish).aspx.
[42] "Las válvulas cardíacas | Texas Heart Institute", Texas Heart Institute, 2019. [Online]. Available: https://www.texasheart.org/heart-health/heart-information-center/topics/las-valvulas-cardiacas/.
[43] Mefanet, «Cardiac cycle» [Online]. Available: https://mefanet.lfp.cuni.cz/download.php?fid=31.
[44] D. Ahmad and D. Ahmad, "Cardiac Cycle,Phases of Cardiac Cycle,Cardiac Cycle & ECG", Medicosite.com, 2019. [Online]. Available: https://www.medicosite.com/cardiac-cycle/.
[45] A. Hertzman and C. Spealman, "Observations on the finger volume pulse recorded photoelectrically", Am.J.Physiol, no. 119, pp. 334-335, 1937.

[46] J. Allen, "Photoplethysmography and its application in clinical physiological measurement", Physiological Measurement, vol. 28, no. 3, pp. R1-R39, 2007. Available: 10.1088/0967-3334/28/3/r01.
[47] A. Alian and K. Shelley, "Photoplethysmography", 2014. [Online]. Available: http://www.elsevier.com/locate/bean.
[48] M. Susha Cheriyedath, "Photoplethysmography (PPG)", News-Medical.net, 2019. [Online]. Available: https://www.news-medical.net/health/Photoplethysmography-(PPG)-(Spanish).aspx. [49] "How Does PPG Technology Works? - SoulFitBlog", Soulfit Blog, 2019. [Online]. Available: https://soulfit.io/blog/how-does-ppg-technology-works/.
[50] M. O'Rourke, "Wave reflections and the arterial pulse", Archives of Internal Medicine, vol. 144, no. 2, pp. 366-371, 1984. Available: 10.1001/archinte.144.2.366.
[51] K. Shelley, M. Dickstein and S. Shulman, "The detection of peripheral venous pulsation using the pulse oximeter as a plethysmograph", Journal of Clinical Monitoring, vol. 9, no. 4, pp. 283-287, 1993. Available: 10.1007/bf02886699.
[52] G. Celi, M. Rocha and M. Yapur, "Mediciones Fotopletismográficas", Escuela Superior Politécnica del Litoral, Ecuador, 2011.
[53] D. Sanchez, "DISEÑO DE UN DISPOSITIVO PARA LA DETECCIÓN DEL ESTRÉS A PARTIR DE LA SEÑAL DE FOTOPLETISMOGRAFÍA", Undergraduate, ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA.
[54] C. Salud, Q. cardiovascular and Q. cardiovascular, "Qué es el riesgo cardiovascular", Riojasalud.es, 2019. [Online]. Available: https://www.riojasalud.es/ciudadanos/catalogo-multimedia/nefrologia/que-es-el-riesgo-cardiovascular.
[55]"Cardiovascular disease", nhs.uk, 2019. [Online]. Available: https://www.nhs.uk/conditions/cardiovascular-disease/.
[56] "Centro de Información Cardiovascular | Texas Heart Institute", Texas Heart Institute, 2019. [Online]. Available: https://www.texasheart.org/salud-cardiovascular/centro-de-informacion-cardiovascular/.
[57] S. Landinez and C. Villamil, "SISTEMA ECG MÓVIL PARA SOPORTAR LA EVALUACIÓN DEL RIESGO CARDIOVASCULAR", Undergraduate, Universidad del Cauca, 2016.
[58]"Factores de riesgo - Fundación Española del Corazón", Fundaciondelcorazon.com, 2019. [Online]. Available: https://fundaciondelcorazon.com/prevencion/riesgo-cardiovascular.html.
[59]Prevention of cardiovascular disease. Geneva: World Health Organization, 2007, pp. 8-22.
[60] "Factores de riesgo cardiovascular | Texas Heart Institute", Texas Heart Institute, 2019. [Online]. Available: https://www.texasheart.org/heart-health/heart-information-center/topics/factores-de-riesgo-cardiovascular/.
[61] "Tachycardia: Fast Heart Rate", www.heart.org, 2019. [Online]. Available: https://www.heart.org/en/health-topics/arrhythmia/about-arrhythmia/tachycardia--fast-heart-rate. [62]H. Lohninger, "Signal and Noise", Statistics4u.com, 2019. [Online]. Available: http://www.statistics4u.com/fundstat_eng/cc_signal_noise.html.
[63] R. Banerjee, A. Ghose, A. Dutta Choudhury, A. Sinha and A. Pal, "Noise cleaning and Gaussian modeling of smart phone photoplethysmogram to improve blood pressure estimation", 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015. Available: 10.1109/icassp.2015.7178113.

[64] T. Wartzek, "Modeling of Motion Artifacts in Contactless Heart Rate Measurements", Computing in Cardiology, vol. 40, pp. 931-934, 2019. Available: https://ieeexplore.ieee.org/abstract/document/6713531.
[65] Lee J. "Motion artifacts reduction from PPG using cyclic moving average filter." Technology and Health Care. pp:409-17, 2014.
[66] M. Hassani and M. Karami, "Improved EEG Segmentation Using Non-linear Volterra Model in Bayesian Method", IETE Journal of Research, vol. 64, no. 6, pp. 832-842, 2017. Available: 10.1080/03772063.2017.1379889.
[67] M. Taherisadr, O. Dehzangi and H. Parsaei, "Single Channel EEG Artifact Identification Using Two-Dimensional Multi-Resolution Analysis", Sensors, vol. 17, no. 12, p. 2895, 2017. Available: 10.3390/s17122895.
[68] K. Suresh and P. Manimegalai, "Detection and separation of EEG artifacts using wavelet transform", Journal of Engineering and Applied Sciences, vol. 13, no. 11, pp. 4165-4172, 2018. Available: 10.3923/jeasci.2018.4165.4172.
[69] C. Lin, C. Huang, W. Yang, A. Singh, C. Chuang and Y. Wang, "Real-Time EEG Signal Enhancement Using Canonical Correlation Analysis and Gaussian Mixture Clustering", Journal of Healthcare Engineering, vol. 2018, pp. 1-11, 2018. Available: 10.1155/2018/5081258.
[70] B. Singh and H. Wagatsuma, "Two-stage wavelet shrinkage and EEG-EOG signal contamination model to realize quantitative validations for the artifact removal from multiresource biosignals", Biomedical Signal Processing and Control, vol. 47, pp. 96-114, 2019. Available: 10.1016/j.bspc.2018.08.014.
[71] M. Hassani and M. Karami, "Noise estimation in electroencephalogram signal by using volterra series coefficients", Journal of Medical Signals & Sensors, vol. 5, no. 3, p. 192, 2015. Available: 10.4103/2228-7477.161495.
[72] L. Han-Wook, L. Ju-Won, J. Won-Geun and L. Gun-Ki, "The Periodic Moving Average Filter for Removing Motion Artifacts from PPG Signals", International Journal of Control, Automation, and Systems, vol. 5, no. 6, pp. 701-706, 2007.
[73] Kejariwal, A. and Daoudi, M. (2019). Filtering Out The Noise | Signal Processing. [online] Catchpoint's Blog - Web Performance Monitoring. Available at: http://blog.catchpoint.com/2016/06/29/signal-vs-noise/.
[74] Zhilin Zhang, Zhouyue Pi, Benyuan Liu, ‘‘TROIKA: A General Framework for Heart Rate Monitoring Using Wrist-Type Photoplethysmographic Signals During Intensive Physical Exercise,''IEEE Trans. on Biomedical Engineering, vol. 62, no. 2, pp. 522-531, February 2015
[75]. A. Savitzky and M. Golay, "Smoothing and Differentiation of Data by Simplified Least Squares Procedures.", Analytical Chemistry, vol. 36, no. 8, pp. 1627-1639, 1964. Available: https://pubs.acs.org/doi/abs/10.1021/ac60214a047.
[76] S. Chang, Bin Yu and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression", IEEE Transactions on Image Processing, vol. 9, no. 9, pp. 1532-1546, 2000. Available: 10.1109/83.862633.
[77] J. Makhoul, "Linear prediction: A tutorial review", Proceedings of the IEEE, vol. 63, no. 4, pp. 561-580, 1975. Available: 10.1109/proc.1975.9792.
[78] D. R. Cox and H. D. Miller, The theory of stochastic processes, 1st ed. London: Chapman & Hall, 1996.

[79] "Chapter 9. Noise", Ocw.mit.edu, 2019. [Online]. Available: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-02-introduction-to-eecs-ii-digital-communication-systems-fall-2012/readings/MIT6_02F12_chap09.pdf.
[80] "¿Qué es la distribución normal? - Minitab", Support.minitab.com, 2019. [Online]. Available: https://support.minitab.com/es-mx/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/normality/what-is-the-normal-distribution/.
[81]"Reducing False Arrhythmia Alarms in the ICU: the PhysioNet/Computing in Cardiology Challenge 2015", Physionet.org, 2019. [Online]. Available: https://physionet.org/challenge/2015/.
[82] Chambrin MC. Review: Alarms in the intensive care unit: how can the number of false alarms be reduced? Critical Care. 2001 Aug; 5(4):184-8. Epub 2001 May 23.
[83] Meyer TJ, Eveloff SE, Bauer MS, Schwartz WA, Hill NS, Millman RP. Adverse environmental conditions in the respiratory and medical ICU settings. Chest. 1994 Apr; 105(4), 1211-16.
[84] Parthasarathy S, Tobin MJ. Sleep in the intensive care unit. Intensive Care Med. 2004 Feb; 30(2), 197-206.
[85] Johnson AN. Neonatal response to control of noise inside the incubator. Pediatr Nurs. 2001 Nov-Dec; 27(6), 600-5.
[86] Slevin M, Farrington N, Duffy G, Daly L, Murphy JF. Altering the NICU and measuring infants' responses. Acta Paediatr. 2000 May; 89(5), 577-81.
[87] A.J. Cropp, L.A. Woods, D. Raney, D.L. Bredle, Name that tone. The proliferation of alarms in the intensive care unit. Chest, 105 (4) (1994), pp. 1217-1220 Apr
[88] M.A. Novaes, A. Aronovich, M.B. Ferraz, E. Knobel, Stressors in ICU: patients' evaluation, Intens Care Med, 23 (12) (1997), pp. 1282-1285 Dec
[89] M. Topf, S. Thompson, Interactive relationships between hospital patients' noise induced stress and other stress with sleep, Heart Lung, 30 (4) (2001), pp. 237-243 Jul-Aug
[90] W.E. Morrison, E.C. Haas, D.H. Shaffner, E.S. Garrett, J.C. Fackler, Noise, stress, and annoyance in a pediatric intensive care unit, Crit Care Med, 31 (1) (2003), pp. 113-119 Jan
[91] S. Berg, Impact of reduced reverberation time on sound-induced arousals during sleep, Sleep, 24 (3) (2001), pp. 289-292 May 1
[92] Donchin Y, Seagull FJ. The hostile environment of the intensive care unit. Curr Opin Crit Care. 2002 Aug;8(4):316-20.
[93] S.T. Lawless, Crying wolf: false alarms in a pediatric intensive care unit, Crit Care Med, 22 (6) (1994), pp. 981-985 Jun
[94] C.L. Tsien, J.C. Fackler, Poor prognosis for existing monitors in the intensive care unit, Crit Care Med, 25 (4) (1997), pp. 614-619 Apr
[95] Li Q, Mark RG, Clifford GD. Artificial arterial blood pressure artifact models and an evaluation of a robust blood pressure and heart rate estimator. Biomed Eng. Online 2009; 8:13
[96] A. Biswas, M. Singha Roy and R. Gupta, "Motion Artifact Reduction from Finger Photoplethysmogram Using Discrete Wavelet Transform", Recent Trends in Signal and Image Processing, pp. 89-98, 2018. Available: 10.1007/978-981-10-8863-6_10.
[97] A. K R, R. Srivathsa and B. M, "Improved Heart Rate Estimation from Photoplethysmography During Physical Exercise Using Combination of NLMS and RLS

Adaptive Filters", TENCON 2018 - 2018 IEEE Region 10 Conference, 2018. Available: 10.1109/tencon.2018.8650495.
[98] W. He, Y. Ye, L. Lu, Y. Cheng, Y. Li and Z. Wang, "Robust Heart Rate Monitoring for Quasi-periodic Motions by Wrist-Type PPG Signals", IEEE Journal of Biomedical and Health Informatics, pp. 1-1, 2019. Available: 10.1109/jbhi.2019.2912708.
[99] B. Keerthiveena and S. Esakkirajan, "Denoising of PPG signal by wavelet packet transform", 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2017. Available: 10.1109/icicict1.2017.8342632.
[100] P. Kasambe and S. Rathod, "VLSI Wavelet Based Denoising of PPG Signal", Procedia Computer Science, vol. 49, pp. 282-288, 2015. Available: 10.1016/j.procs.2015.04.254.
[101] M. Raghuram, K. Venu Madhav, E. Hari Krishna and K. Ashoka Reddy, "Evaluation of wavelets for reduction of motion artifacts in photoplethysmographic signals", 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), 2010. Available: 10.1109/isspa.2010.5605443.
[102] M. Islam, M. Shifat-E-Rabbi, A. Dobaie and M. Hasan, "PREHEAT: Precision heart rate monitoring from intense motion artifact corrupted PPG signals using constrained RLS and wavelets", Biomedical Signal Processing and Control, vol. 38, pp. 212-223, 2017. Available: 10.1016/j.bspc.2017.05.010.
[103] J. Xiong, L. Cai, F. Wang and X. He, "SVM-Based Spectral Analysis for Heart Rate from Multi-Channel WPPG Sensor Signals", Sensors, vol. 17, no. 3, p. 506, 2017. Available: 10.3390/s17030506.
[104] P. Mullan et al., "Unobtrusive heart rate estimation during physical exercise using photoplethysmographic and acceleration data", 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015. Available: 10.1109/embc.2015.7319787.
[105] G. Joseph, A. Joseph, G. Titus, R. Thomas and D. Jose, "Photoplethysmogram (PPG) signal analysis and wavelet de-noising", 2014 Annual International Conference on Emerging Research Areas: Magnetics, Machines and Drives (AICERA/iCMMD), 2014. Available: 10.1109/aicera.2014.6908199.
[106] M. Raghuram, K. Madhav, E. Krishna, N. Komalla, K. Sivani and K. Reddy, "Dual-tree complex wavelet transform for motion artifact reduction of PPG signals", 2012 IEEE International Symposium on Medical Measurements and Applications Proceedings, 2012. Available: 10.1109/memea.2012.6226643.
[107] P. Thamarai and D. K.Adalarasu, "An Effective Method to Denoise EEG, ECG and PPG Signals based on Meyer Wavelet Transform", International Journal of Engineering & Technology, vol. 7, no. 334, p. 678, 2018. Available: 10.14419/ijet.v7i3.34.19415.
[108] "Denoising of EEG, ECG and PPG signals using wavelet transform", Journal of Pharmaceutical Sciences and Research, vol. 10, no. 1, pp. 156-161, 2019. Available: https://www.researchgate.net/publication/322939862_Denoising_of_EEG_ECG_and_PPG_signals_using_wavelet_transform.
[109] L. Eerikainen, J. Vanschoren, M. Rooijakkers, R. Vullings and R. Aarts, "Decreasing the false alarm rate of arrhythmias in intensive care using a machine learning approach", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7408644.

[110] V. Kalidas and L. Tamil, "Enhancing accuracy of arrhythmia classification by combining logical and machine learning techniques", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411015.
[111] C. Hoog Antink and S. Leonhardt, "Reducing false arrhythmia alarms using robust interval estimation and machine learning", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7408642.
[112] M. Caballero and G. Mirsky, "Reduction of false cardiac arrhythmia alarms through the use of machine learning techniques", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411124.
[113] C. Tsimenidis and A. Murray, "Reliability of clinical alarm detection in intensive care units", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411128.
[114] N. Sadr, J. Huvanandana, D. Nguyen, C. Kalra, A. McEwan and P. de Chazal, "Reducing false arrhythmia alarms in the ICU by Hilbert QRS detection", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411125.
[115] S. Fallet, S. Yazdani and J. Vesin, "A multimodal approach to reduce false arrhythmia alarms in the intensive care unit", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7408640.
[116] A. Serackis, V. Abromavicius and A. Gudiskis, "Identification of ECG signal pattern changes to reduce the incidence of Ventricular Tachycardia false alarms", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411130.
[117] S. Ansari, A. Belle and K. Najarian, "Multi-modal integrated approach towards reducing false arrhythmia alarms during continuous patient monitoring: The Physionet Challenge 2015", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411127.
[118] M. Xu, J. Shen and H. Yu, "Multimodal data classification using signal quality indices and empirical similarity-based reasoning", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411131.
[119] R. He et al., "Reducing false arrhythmia alarms in the ICU using novel signal quality indices assessment method", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411129.
[120] S. Teo et al., "Reducing false arrhythmia alarms in the ICU", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411126.
[121] C. Liu, L. Zhao and H. Tang, "Reduction of False Alarms in Intensive Care Unit using Multi-feature Fusion Method", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411017.
[122] J. Gieraltowski, I. Grzegorczyk, K. Ciuchcinski, K. Kosna, M. Solinski and P. Podziemski, "Algorithm for life-threatening arrhythmias detection with reduced false alarms", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411132.
[123] F. Plesinger, P. Klimes, J. Halamek and P. Jurak, "False alarms in intensive care unit monitors: Detection of life-threatening arrhythmias using elementary algebra, descriptive statistics and fuzzy logic", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7408641.

[124] C. Daluwatte, L. Johannesen, J. Vicente, C. Scully, L. Galeotti and D. Strauss, "Heartbeat fusion algorithm to reduce false alarms for arrhythmias", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7411018.
[125] C. Burrus, R. Gopinath and H. Guo, Introduction to wavelets and wavelet transforms. Upper Saddle River, N.J.: Prentice Hall, 1998.
[126] I. Daubechies, Ten lectures on wavelets. Philadelphia, Pa.: Society for Industrial and Applied Mathematics, 1992.
[127]"Comparación de Wavelet y Transformada de Fourier", Um.edu.ar, 2019. [Online]. Available: http://www.um.edu.ar/math/wavelets/compa.htm.
[128] A. Graps, "An Introduction to Wavelets", Eecis.udel.edu. [Online]. Available: https://www.eecis.udel.edu/~amer/CISC651/IEEEwavelet.pdf.
[129] I. Daubechies, "Orthonormal bases of compactly supported wavelets", Communications on Pure and Applied Mathematics, vol. 41, no. 7, pp. 909-996, 1988. Available: 10.1002/cpa.3160410705.
[130] Introducción a la Transformada Wavelet. Buenos Aires: Universidad Nacional del Centro de la provincia de Buenos Aires, 2006, pp. 1-13.
[131] "Introduction to Wavelet Families- MATLAB & Simulink", Mathworks.com, 2019. [Online]. Available: https://www.mathworks.com/help/wavelet/gs/introduction-to-the-wavelet-families.html#f3-1008627.
[132] G. Clifford et al., "The PhysioNet/Computing in Cardiology Challenge 2015: Reducing false arrhythmia alarms in the ICU", 2015 Computing in Cardiology Conference (CinC), 2015. Available: 10.1109/cic.2015.7408639.
[133] T. O'Haver, A Pragmatic Introduction to Signal Processing. Maryland: The University of Maryland, 2008.
[134] W. Press, Numerical recipes, 3rd ed. Cambridge, UK: Cambridge University Press, 2007, pp. 766-772.
[135]"Moving median - MATLAB movmedian", Mathworks.com, 2019. [Online]. Available: https://www.mathworks.com/help/matlab/ref/movmedian.html?s_tid=doc_ta#bu6j0h1.
[136] S. Tang, Y. Goh, M. Wong and Y. Lew, "PPG signal reconstruction using a combination of discrete wavelet transform and empirical mode decomposition", 2016 6th International Conference on Intelligent and Advanced Systems (ICIAS), 2016. Available: 10.1109/icias.2016.7824118.
[137] Y. Kopsinis and S. McLaughlin, "Development of EMD-Based Denoising Methods Inspired by Wavelet Thresholding", IEEE Transactions on Signal Processing, vol. 57, no. 4, pp. 1351-1362, 2009. Available: 10.1109/tsp.2009.2013885.
[138] A. Boudraa, J. Cexus and Z. Saidi, "EMD-Based Signal Noise Reduction", International Journal of Electronics and Communication Engineering, vol. 1, no. 2, pp. 321-324, 2007.
[139] N. Huang and Z. Wu, "A review on Hilbert-Huang transform: Method and its applications to geophysical studies", Reviews of Geophysics, vol. 46, no. 2, 2008. Available: 10.1029/2007rg000228.
[140] W. Denoiser, "Choose a Wavelet- MATLAB & Simulink", Mathworks.com, 2019. [Online]. Available: https://www.mathworks.com/help/wavelet/gs/choose-a-wavelet.html

[141] Apa, C., Robaina, R., de León, S. and Vallespir, D. (2019). Reporte Técnico RT 10-02 Conceptos de Ingeniería de Software Empírica. [online] Fing.edu.uy.
[142] C. Wohlin, P. Runeson, M. Höst, M. Ohlsson, B. Regnell and A. Wesslén, Experimentation in Software Engineering. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.
[143] Brownlee, J. (2019). Classification Accuracy is Not Enough: More Performance Measures You Can Use. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/classification-accuracy-is-not-enough-more-performance-measures-you-can-use/.
[144] Powers, D. (2019). Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. [online] Adelaide, Australia: School of Informatics and Engineering Flinders-University of South Australia, pp.1-3.
[145] Sammut, C. (2011). Encyclopedia of machine learning. New York, NY: Springer, pp.901-902.
[146] A. Descoins, "Why accuracy alone is a bad measure for classification tasks, and what we can do about it", TyroLabs, 2013.
[147] T. Tape, "Interpreting Diagnostic Tests", Gim.unmc.edu. [Online]. Available: http://gim.unmc.edu/dxtests/Default.htm.
[148] "Assay Validation Methods", Fws.gov, 2019. [Online]. Available: https://www.fws.gov/aah/PDF/SandS.pdf.
[149] R. Polikar, "Part IV. MULTIRESOLUTION ANALYSIS: THE DISCRETE WAVELET TRANSFORM.", The wavelet tutorial, 2018.
[150] T. Tamura, Y. Maeda, M. Sekine and M. Yoshida, "Wearable Photoplethysmographic Sensors—Past and Present", Electronics, vol. 3, no. 2, pp. 282-302, 2014. Available: 10.3390/electronics3020282.
[151] C. Fischer, B. Domer, T. Wibmer and T. Penzel, "An Algorithm for Real-Time Pulse Waveform Segmentation and Artifact Detection in Photoplethysmograms", IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 2, pp. 372-381, 2017. Available: 10.1109/jbhi.2016.2518202.
[152]"Wind Energy Conversion Systems Fault Diagnosis Using Wavelet Analysis", International Review of Electrical Engineering, vol. 3, no. 4, pp. 646-652, 2019.
[153] R. GANESAN, T. DAS and V. VENKATARAMAN, "Wavelet-based multiscale statistical process monitoring: A literature review", IIE Transactions, vol. 36, no. 9, pp. 787-806, 2004. Available: 10.1080/07408170490473060.
[154] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, R. Morales-Caporal and I. Cruz-Vega, "Vibration Analysis of Partially Damaged Rotor Bar in Induction Motor under Different Load Condition Using DWT", Shock and Vibration, vol. 2016, pp. 1-11, 2016. Available: 10.1155/2016/3530464.