Pervasive Electrocardiography and Health Monitoring

Ana Rute Sá Oliveira Banha

Thesis to obtain the Master of Science Degree in

Biomedical Engineering

Supervisor(s): Prof. Ana Luísa Nobre Fred
Dr. Rui Cruz Ferreira

Examination Committee
Chairperson: Prof. Doutor Mário Jorge Costa Gaspar da Silva
Supervisor: Prof. Ana Luísa Nobre Fred
Member of the Committee: Prof. Maria Margarida Campos da Silveira

July 2018
Acknowledgments

A few people were fundamental to the successful elaboration of this thesis and I use this section to acknowledge their contribution.

First, I’d like to thank my colleagues from CGI for their support and patience, especially to my boss Horácio Gonçalves for always helping me and giving me the time and words I needed to work and finish this work.

I would also like to thank Diana Baptista, Joana Santos, Miguel Martinho and André Manso for all their help and guidance during this process.

I would also like to thank the “Tuna Feminina do Instituto Superior Técnico” for being my home for the last 8 years and for making me grow to be the person I am today. A special thanks to my “Tuna sisters” Ana Baptista and Tatiana Sirgado for all the help and encouragement they gave me to go through with this thesis.

To Hospital de Santa Marta, namely to Dr. Rui Ferreira and Eng. Rui César das Neves, for providing access to the anonymised data.

I also thank my family for always supporting me and giving me everything I needed to get an education and all the support for finishing this last task in the university phase.

To my best friends Sofia Alves and Mariana Lopes, thank you for helping me make the most difficult choices in the last few months and for giving me the strength and motivation to finish this thesis. You two helped me more than you’ll ever know.

To the most important person in my life, Dulce Costa, thank you so much for everything. For staying with me until late hours on work days to help me work with LateX, for dealing with all my stress crisis, for helping me in every way that you could everyday and never giving up on me.

Last but not least I thank my supervisor, Ana Fred, for everything she did for me. For all the knowledge transmitted, the ideas and advice she gave me and for always being patient and understanding of my lack of time due to my job.
Resumo

As doenças cardíacas são das mais perigosas e presentes condições em todo o mundo. O método mais utilizado para detectar e diagnosticar este tipo de doença é o Eletrocardiograma (ECG), que é uma técnica simples, não invasiva e de baixo custo para efeitos de diagnóstico e monitorização.

Este trabalho foca-se em duas áreas principais: Eliminação de ruído e troços de sinal corrompido de batimentos cardíacos de um ECG e classificação de arritmias cardíacas. Para a primeira tarefa foram testados vários métodos, nomeadamente dois filtros de mediana com um filtro passa baixo de 40 Hz e métodos combinando diferentes “wavelets” e filtros de média móvel. No presente trabalho, os melhores resultados foram obtidos para a técnica que consistia em utilizar a decomposição com a wavelet quadrática de nível 6, que alcançou 0.835 de semelhança de cosenos. Após um passo de remoção de outliers, o resultado passou para 0.930. A temática de classificação consistiu em classificar os dados de ECG em 5 classes, usando Support Vector Machines (SVM) e K-nearest neighbors (KNN). Duas bases de dados foram utilizadas nesta tarefa, a base de dados MIT-BIH, usada como referência, cuja classificação atingiu uma exactidão de 88.70% e a base de dados de Santa Marta (StM), cuja tarefa de classificação atingiu uma exactidão de 81.06%. Foram ainda realizados dois testes extra usando esta base de dados, um para testar o efeito do melhor método de remoção de ruído e troços de sinal corrompido na tarefa de classificação, o que melhorou os resultados em termos de exactidão.

Abstract

Heart diseases are one of the most dangerous and present conditions worldwide. The common method for the detection and diagnosis of this type of disease is electrocardiography (ECG), which is a simple, non-invasive and cost-effective technique for diagnosing and monitoring purposes.

This work focused on two main areas: Denoising and outlier removal of ECG heartbeats and Classification of arrhythmic heartbeats. For the first task several method were tested, namely two median filters plus a 40 Hz Low pass filter, wavelet based denoising methods, using several mother wavelets and a moving average filter. In the present work, the best results were obtained with the technique consisting of using a decomposition level 6 quadratic spline wavelet based denoising which achieved a 0.835 cosine similarity. After an outlier removal step, the result was improved to 0.930.

The classification task consisted in separating a dataset into 5 classes and using the Support Vector Machines (SVM) and K-nearest neighbors (KNN) classifiers. Two databases were used for this task, the MIT-BIH database that was used as a benchmark, whose classification achieved an accuracy of 88.70% and the Santa Marta (StM) database, whose classification task achieved an 81.06% accuracy. Also, two extra tests were done using this database, one for testing the effect of the best denoising method on the classification task, which improved the results in terms of accuracy.

Keywords: Arrhythmic heartbeats, ECG, Heart Rhythm, Support Vector Machine, K-Nearest Neighbor, Wavelet Decomposition, Outlier Removal, Denoising, Classification.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>iii</td>
</tr>
<tr>
<td>Resumo</td>
<td>v</td>
</tr>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td>Nomenclature</td>
<td>xv</td>
</tr>
<tr>
<td>Glossary</td>
<td>1</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Goals and Proposed Approach</td>
<td>2</td>
</tr>
<tr>
<td>1.3 Contributions</td>
<td>2</td>
</tr>
<tr>
<td>1.4 Structure</td>
<td>3</td>
</tr>
<tr>
<td>2 Electrophysiology</td>
<td>4</td>
</tr>
<tr>
<td>2.1 Electrical Conduction System</td>
<td>4</td>
</tr>
<tr>
<td>2.2 Electrocardiogram</td>
<td>6</td>
</tr>
<tr>
<td>2.3 Arrhythmia and ECG heartbeat types</td>
<td>7</td>
</tr>
<tr>
<td>2.3.1 Sinus node originated Heartbeats</td>
<td>8</td>
</tr>
<tr>
<td>2.3.2 Supraventricular Ectopic Beats</td>
<td>10</td>
</tr>
<tr>
<td>2.3.3 Ventricular Ectopic Beats</td>
<td>11</td>
</tr>
<tr>
<td>2.3.4 Fusion Beats</td>
<td>13</td>
</tr>
<tr>
<td>2.4 Standard 12-Lead ECG</td>
<td>13</td>
</tr>
<tr>
<td>3 State-of-the-Art</td>
<td>15</td>
</tr>
<tr>
<td>3.1 ECG Denoising Techniques</td>
<td>15</td>
</tr>
<tr>
<td>3.2 Heartbeat Classification</td>
<td>17</td>
</tr>
<tr>
<td>3.2.1 Feature Extraction and Classification</td>
<td>17</td>
</tr>
<tr>
<td>3.3 Wavelet Transform</td>
<td>22</td>
</tr>
<tr>
<td>3.4 R peak Detection</td>
<td>23</td>
</tr>
<tr>
<td>3.5 Classifiers</td>
<td>24</td>
</tr>
<tr>
<td>3.5.1 Support Vector Machine (SVM)</td>
<td>24</td>
</tr>
</tbody>
</table>
3.5.2 K-Nearest Neighbours (KNN) ................................................. 26

4 Proposed Methodology ............................................................... 27
4.1 Types of Arrhythmia ............................................................... 27
4.2 Denoising ............................................................................. 29
  4.2.1 40 Hz Low-Pass Filter ....................................................... 30
  4.2.2 Denoising using Wavelet Thresholding Techniques .................. 31
  4.2.3 Moving Average ............................................................. 32
4.3 Outlier removal ................................................................. 33
4.4 Feature Extraction ............................................................... 33
  4.4.1 Temporal Features .......................................................... 33
  4.4.2 Morphological Features .................................................... 34
  4.4.3 Evaluation and performance measures ................................... 34

5 Experimental Setup and Results ................................................... 36
5.1 Experimental Setup ............................................................. 36
  5.1.1 BITalino Bicycle (BB) Database ......................................... 36
  5.1.2 MIT-BIH Arrhythmia (MIT-BIH) Database ......................... 37
  5.1.3 Santa Marta (StM) Database .............................................. 38
5.2 Results for Denoising Techniques ............................................. 38
  5.2.1 40 Hz Low-Pass Filter Results ......................................... 39
  5.2.2 Wavelet Based Denoising Techniques .................................. 39
  5.2.3 Moving Average ............................................................ 42
  5.2.4 Outlier Removal ............................................................. 42
  5.2.5 Denoising Task Discussion .............................................. 44
5.3 Classification Results .......................................................... 50
  5.3.1 MIT-BIH - Benchmark Data Experimental Results .................. 50
  5.3.2 StM - Validation Setup Experimental Results ...................... 53

6 Conclusions and Future Work ..................................................... 58

Bibliography ............................................................................. 60

A Outlier removal in MIT-BIH Database ...................................... A.1
List of Tables

3.1 Summary of the most relevant methods for features extraction and classification presented in this section. The unavailable information is marked with an asterisk (*). ............... 18

4.1 Association for the Advancement of Medical Instrumentation (AAMI) recommended practice based classes. [54] ................................................................. 28

4.2 Classes considered for this thesis. ........................................................................ 28

4.3 Number of samples for each label in the MIT-BIH and STIM databases. ............... 29

4.4 Experiments specification for Classification Task .................................................. 35

5.1 SNR and CosSim values for the four subjects with the best results (highest values) after only doing a baseline wander removal using 2 median filters. ......................... 39

5.2 SNR and CosSim values for the four subjects with the best results (highest values) before and after 40 Hz LP filter. ................................................................. 39

5.3 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Sym12, is applied to the heartbeats without applying a 40 Hz low pass filtering to the noisy signal first. ..................................... 40

5.4 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Sym12, is applied to the heartbeats after first applying a 40 Hz low pass filtering to the noisy signal. ..................................... 40

5.5 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Db4, is applied to the heartbeats without 40 Hz low pass filtering to the noisy signal. .......................................................... 41

5.6 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Db4, is applied to the heartbeats after the 40 Hz low pass filtering to the noisy signal. .......................................................... 41

5.7 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Quadratic Spline, is applied to the heartbeats without 40 Hz low pass filtering to the noisy signal. .......................................................... 42

5.8 SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Quadratic Spline, is applied to the heartbeats after the 40 Hz low pass filtering to the noisy signal. .......................................................... 42
5.9 SNR and CosSim values for the four subjects with the best results (highest values) before and after moving average is applied to the signals. ................................................................. 43
5.10 Results for validation of outlier removal method using the MIT-BIH database. This shows the number of subjects with a difference in the initial and final number of heartbeats higher than 100 and also the maximum and minimum values of the difference from the initial and final number of heartbeats, i.e., before and after outlier removal. More detailed results are present in Appendix A. ........................................................................................................... 50
5.11 SNR and CosSim values for the four subjects with the best results (highest values) before and after Outlier Removal + Quadratic Spline, level 6, is applied to the signals. ............. 52
5.12 Results for the AAMI recommended practices based classes, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results. ................................................................. 52
5.13 Results for proposed classes for this thesis, namely N, L, S, V and F, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results. ........................................................................................................... 52
5.14 Results for classification of L and V heartbeats, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results. 52
5.15 Results for classification of the 5 labels using the Db2 wavelet with a decomposition level 4. 54
5.16 Results for classification of the 5 labels using the Quadratic Spline wavelet with a decomposition level 6. ........................................................................................................... 55
5.17 Number of samples in each class after outlier removal step. ................................. 55
5.18 Results for classification of the 5 labels after an outlier removal and denoising steps. 56
5.19 Results for classification of the 5 labels using the Db2 wavelet with a decomposition level 4 and the KNN classifier. ........................................................................................................... 57
List of Figures

2.1 The cardiac conduction system [6]. ......................................................... 5
2.2 The components of the ECG signal [8]. .................................................... 6
2.3 Left bundle branch block heartbeat [11]. .................................................. 8
2.4 Right bundle branch block (RBBB) in lead V1 [10]. .................................. 9
2.5 Atrial escape beat [10]. ............................................................................ 9
2.6 Nodal (junctional) escape beat [13]. ....................................................... 10
2.7 Atrial premature beat three possible faiths, where the dots represent normal sinus beats and the stars represent the atrial premature beats. The first APB is conducted normally, the second one follows the previous sinus beat earlier and is conducted with RBBB aberration and the third one is still a little bit earlier after the previous sinus beat and is blocked in the AV node and is, therefore, not conducted and results in a present P-wave but no QRS complex is present, followed by a non-compensatory pause [14]. .......................... 11
2.8 Nodal (junctional) premature beat [13]. .................................................. 12
2.9 Premature ventricular contraction [17]. ................................................... 12
2.10 Ventricular escape beat [13]. .................................................................. 12
2.11 Fusion Beat [13]. .................................................................................... 13
2.12 Lead placement for a 12-lead ECG [18]. ................................................ 14

3.1 Example of classification using SVM algorithm. ....................................... 25
3.2 Example of classification using KNN algorithm [53]. .............................. 26

4.1 Denoising methodology. ......................................................................... 27
4.2 Steps for the classification task. ............................................................... 27
4.3 Top: Example of the baseline signal used for measuring the results of the denoising technique (Clean Signal); Bottom: Example of signal corrupted by noise (Noisy Signal). ...... 30
4.4 Mother wavelet the Db4 wavelet .............................................................. 31
4.5 Mother wavelet of the sym12 wavelet ..................................................... 31
4.6 Mother wavelet of the quadratic spline wavelet ..................................... 32
4.7 Mother wavelet (left) and scaling function (right) of the Db2 wavelet ...... 34

5.1 Chest Band used for collecting the Clean Signal. .................................... 37
5.2 Stationary Bicycle used for collecting the Noisy Signal. ...................... 37
5.3 ECG recordings from subjects that were removed from the denoising experiment.

5.4 Set 1 (of two) of Final Prototypes for Outlier Removal. Green heartbeats represents the centroid and the red ones represent the heartbeat with the least euclidean distance to the centroid.

5.5 Set 2 (of two) of Final Prototypes for Outlier Removal. Green heartbeats represents the centroid and the red ones represent the heartbeat with the least euclidean distance to the centroid.

5.6 CosSim mean, mode and median values for subjects whose CosSim mean was higher then 0.5.

5.7 SNR mean, mode and median values for subjects whose CosSim mean was higher then 0.5.

5.8 CosSim mean, mode and median values of quadratic spline wavelet based technique with and without outlier removal for subjects whose CosSim mean was higher then 0.5.

5.9 SNR mean, mode and median values of quadratic spline wavelet based technique with and without outlier removal for subjects whose CosSim mean was higher then 0.5.

5.10 Comparison between the results for the classification of the 5 different classes(green) and taking the "F" class beats from the dataset (blue).

5.11 Comparison between the results for the classification with 681 samples per class balancing (and 31 for the "F" class) (green) and for classification after outlier removal plus quadratic spline wavelet based denoising step.

5.12 Comparison between the results for the classification of the 5 different classes(green) and taking the "F" class beats from the dataset (blue), using the KNN classifier.

A.1 Results of Outlier removal in MIT-BIH Database.
Nomenclature

Symbols

\(\gamma\)  
Gamma Parameter.

\(F_1\)  
F-score.

BB  
Blitalino Bicycle.

CosSim  
Cosine Similarity.

Db2  
Daubechies wavelet of order 2.

ECG  
Electrocardiogram

EMG  
Electromyogram

FT  
Fourier Transform

KNN  
K-Nearest Neighbour.

LBBB  
Left bundle branch block beats.

MIT  
Massachusetts Institute of Technology.

RBBB  
Right bundle branch block beats.

SNR  
Signal-to-noise Ratio

StM  
Santa Marta Hospital dataset.

SVM  
Support Vector Machines.

Sym12  
Symlets 12 wavelet.

WT  
Wavelet Transform
Chapter 1

Introduction

This chapter presents the motivation, goals and the structure of this thesis.

1.1 Motivation

Heart diseases are one of the most dangerous and present conditions worldwide. The common approach for the detection and diagnosis of this type of disease is based on electrocardiography (ECG), which is a simple, non-invasive and cost-effective technique for diagnosing and monitoring purposes. Besides the standard 12-lead ECG, which is widely used, other types of monitoring tools were developed in the last few decades, such as Holter monitors, patch monitors, among others, which shows the growing need to keep developing the heart monitoring area, in order to better diagnosing these diseases and, ultimately, saving many of lives. From this type of devices, a huge amount of data can be collected and processed, an essential feature for developing algorithms that help in their analysis.

In this work the main focus is on arrhythmic heartbeats, a disturbance in rate, rhythm or conduction of the electric signal through the heart. While some arrhythmias are harmless, others can compromise the functioning of the heart, being potentially heart threatening, a fact that increases the importance of the development of automated ways to classify the heartbeats and get a fast and accurate diagnosis to the patient. While research on heart rhythm and heartbeat automatic classification areas have gained an increased interest, with many new approaches and techniques being proposed every day, still much work has to be done towards increased diagnosis accuracy. In particular, in the area of IoT, with sensors being integrated into everyday objects (such as smart-phones or garments), analysis of the ECG data collected with these sensors poses additional difficulties, mainly due to the low quality of the acquired signals. These present artifacts and a high level of noise contamination, due to the intermitence and/or poor contact of electrodes with the skin (more notably associated with movements) and contamination with other signals (such as muscle activity).

In a step towards pervasive signal acquisition systems, most of the works are based on the placement of electrodes at the hands, wrist or chest level, corresponding to lead I placement according to the standard [1]. Hence the main focus is on diagnosis based on a single lead, mainly lead I.
1.2 Goals and Proposed Approach

This thesis focuses on two main objectives: ECG denoising and artifact removal, automatic heartbeat classification from single-lead ECG records.

For the first task of ECG denoising, several methods are tested and compared, namely two median filters plus a 40 Hz low pass filter, wavelet based denoising techniques and moving average filters. Besides these methods, an outlier removal technique is proposed and developed, aiming to remove the heartbeats that are not relevant for the classification task at hand, and the artifacts present in acquired signals. The denoising methods were tested on a database developed at IT - Instituto de Telecomunicações [2], using a BiTalino device [3], that comprises both signals with a small amount of noise, used as a baseline, acquired by placing electrodes on the chest, and signals with a big amount of noise, acquired using a static bicycle where the handles coupled to a BiTalino device were used to extract the ECG signal in the hands. The bicycle signals went through several denoising methods separately and compared to the first ones. The objective was to calculate the difference between the baseline signals and the denoised ones in terms of Signal-to-Noise ratio and cosine similarity for each method and decide which method is better for this type of task.

For the classification task, which was considered the final purpose of the thesis, three different types of features are explored: temporal features, morphological wavelet-based features and a combination of both. The aim of this task was to develop an algorithm that could identify 5 types of heartbeats: Beats originating in the sinoatrial (SA) node (N), left bundle branch block beats (L), supraventricular ectopic beats (S), ventricular ectopic beats (V) and fusion beats (F). Firstly, the best set of features for this task was chosen, based on tests performed using the MIT-BIH Arrhythmia Database, which is widely used for heartbeat and rhythm automatic classification experiments. The feature set thus selected was then used for data representation of the classifiers applied to the validation dataset, gathered at the Santa Marta Hospital, in Lisbon, the Santa Marta (StM) Database. Also for this database, a feature set based on a different mother wavelet was tested, as well as the effect of denoising and outlier removal on the classification task. Two types of classifiers were evaluated, namely: Support vector Machines (SVM); and K-nearest neighbors (KNN) algorithms.

1.3 Contributions

The contributions of this thesis are:

- Evaluates the performance of several denoising algorithms, in order to improve the quality of ECG signals;
- Presents a method for the development of prototypes for normal and pathological heartbeats that can be used both for classification and for outlier removal purposes in ECG signals;
- Analyzes the performance of time domain and morphological features, considered both individually and combined, on the classification of normal and pathological heartbeats;
• Compares the performance of two supervised learning algorithms, nearest neighbor and support vector machine, on the classification of specific heartbeat types.

1.4 Structure

This thesis is organized in six chapters:

• On the current chapter, Chapter 1, the motivation and proposed approach was outlined;

• Chapter 2 gathers important basic concepts concerning the electrical conduction system of the heart and the different types of heartbeats that are important for this study;

• In Chapter 3, a review of state-of-the-art algorithms for both ECG denoising and heartbeat and rhythm classification;

• Chapter 4 includes the materials used for this thesis, namely the used databases, the arrhythmia classification convection that was used and the methodology that was applied for both tasks;

• Experiments and results are detailed and discussed in Chapter 5;

• Chapter 6 contains the main conclusions of the work and suggestions for future developments.
Chapter 2

Electrophysiology

2.1 Electrical Conduction System of the Heart

The heart is a muscle whose function is to pump blood into the body and that is at the center of the circulatory system. This system is also constituted by a network of blood vessels, such as arteries, veins and capillaries. This muscle is controlled by an electrical system that uses electrical signals in order to contract the heart's wall. When a contraction occurs, blood is pumped into the circulatory system and the direction of its flow is controlled by the valves that exist in the heart chambers [4]. The heart has four chambers, two atria and two ventricles. The de-oxygenated blood returns to the heart, namely to the right atrium, via venous circulation. This part of the circulatory system is known as the systemic circuit. Afterwards, the blood is pumped into the right ventricle and then to the lungs, where carbon dioxide is released and oxygen is absorbed. The now oxygenated blood travels back to the heart, into the left atrium. This part of the system, called pulmonary circuit, is much smaller and its main responsibility is to oxygenate the blood.

The blood is then pumped into the left ventricle and then into the aorta, starting the circulatory system again and being supplied into the rest of the body [5]. In order to function, the heart needs a power source and uses electricity, being able to create its own electrical impulses and their route via a specialized conduction pathway. This pathway has five elements (Figure 2.1):

1. The sino-atrial (SA) node;
2. The atrio-ventricular (AV) node;
3. The bundle of His;
4. The left and right bundle branches;
5. The Purkinje fibres.

The SA node works as a natural pacemaker of the heart. It releases electrical stimuli at a regular rate (60 to 100 times per minute), which is controlled by the needs of the body. Each stimulus goes through the myocardial cells of the atria, creating a wave of contraction that spreads quickly through
both of them. The speed of this atrial contraction is so high, that it appears instantaneous. Eventually, the stimulus reaches the AV node and is slightly delayed (AV node delays impulses by 0.04 seconds), in order to allow for the contracting atria to have enough time to pump all the blood into the ventricles, followed by the closing of the valves between the atria and ventricles, once the atria are empty. This delay is one of the main functions of the AV node and some impulses may even be blocked if the atrial rate is dangerously high. Additionally, the cells in the AV node are capable of generating impulses at a rate of 40 to 60 beats per minute. At this point, the atria begin to refill and the electrical signal passes through the AV node and Bundle of His. The latest divides into the right and left bundle branches, which assures the conduction, respectively, to the right and left ventricles.

The Purkinje fibers extend from the bundle branches and spread across the ventricles. These are going to receive the impulses and, as a response, contract. These fibers can also serve as pacemakers, being able to produce impulses at a rate of 20 to 40 times per minute, sometimes less [7]. As the ventricles contract, the right one pumps blood that goes to the lungs while the left pumps blood into the aorta, beginning the systemic circuit. At this point of the process, the ventricles are empty and the atria are full, with all the valves closed. Thus beginning the process again, with the release of the impulse by the SA node. This node recharges as the atria is refilling, whilst the AV node recharges when the ventricles are being refilled. In all these processes, the terms used for the release and the recharging of the electrical are "depolarization" and "repolarization", respectively. Therefore, the 3 stages of a single heartbeat are:

1. Atrial depolarization;

2. Ventricular depolarization;

3. Atrial and ventricular repolarization.
2.2 Electrocardiogram

The heart's electrical activity creates currents that spread through the surrounding tissue to the skin so, when electrodes are attached to the skin, they can sense those currents and transmit them to an Electrocardiogram (ECG) monitor. These currents are represented in waveform, which represents the depolarization-repolarization cycle of the heart. An ECG shows the sequence of the electrical events that happen in the cardiac cell throughout that cycle and allows the monitoring of the phases of myocardial contraction and the identification of rhythm and conduction disturbances. The ECG complex is, therefore, a representation of the events that take place during one cardiac cycle and its tracings represent the conduction of the impulses from the atria to the ventricles. It consists of five different waveforms, labeled with the letters "P", "Q", "R", "S" and "T". "Q", "R" and "S" are referred to as an unit, known as the QRS complex (Figure 2.2).

![Figure 2.2: The components of the ECG signal [8].](image)

The P wave is the first component that can be seen on a normal ECG waveform. It corresponds to the atrial depolarization, i.e., the conduction of the electrical stimuli from the SA node through both atria. The duration of this type of wave can vary between 0.06 to 0.12 seconds.

The PR interval tracks the impulse from the atria and through the AV node, bundle of His and right and left branches. It is the time between the onsets of atrial and ventricular depolarization. This interval goes from the beginning of the P wave until the start of the QRS complex. It usually has a duration ranging from 0.12 to 0.2 seconds.

Following the P wave, the QRS complex represents the depolarization of the ventricles. It is often composed of three different waves, as mentioned earlier: The Q wave, which is the first negative deflection after the P wave, the first positive deflection after the P and Q waves known as R wave and the S wave, which consists of the first negative deflection after the R wave. The QRS complex, in normal situations, has a duration of 0.06 to 0.10 second, duration which should be measured from the beginning of the Q wave to the end of the S wave or from the beginning of the R wave until the end of the S wave,
if the Q wave is absent.

The ST segment is the representation of the end of the ventricular conduction or depolarization and the beginning of ventricular recovery or repolarization. It is usually an isoelectric line, extends from the S wave to the beginning of the T wave and and elevation or depression in this segment can indicate myocardial damage.

The representation of the ventricular repolarization on an ECG is the T wave. This wave should be in the same direction of the QRS complex and is asymmetrical, with the first part sloping to the peak and returning more abruptly to the baseline.

The QT interval represents the time needed for the ventricular depolarization-repolarization cycle, being measured from the beginning of the QRS complex until the end of the T wave. Depending on factors such as gender, age and heart rate, it can last from 0.36 to 0.44 seconds and shouldn’t be greater than half the distance between consecutive R waves, when a regular rhythm is observed.

Lastly, the U wave represents the period that corresponds to the recovery of the Purkinje fibers. It follows the T wave and is not always present on an ECG. This absence doesn’t represent an abnormality. These waves are usually more prominent when the heart rate is lower.

### 2.3 Arrhythmia and ECG heartbeat types

Arrhythmia, or dysrhythmia, is a very general term that can refer to all rhythms of the heart other than the normal one. It can represent any disturbance in the rate, regularity, site of origin or conduction of the cardiac impulse. As described before, the cardiac impulse starts in the SA node but, under some circumstances, cardiac cells in any other part of the heart can serve as a pacemaker of the heart. When this happens, we are in the presence of an ectopic pacemaker, that is, a pacemaker that is any other than the sinus node. The result may be ectopic beats or rhythms. Depending on the location of the pacemaker (natural - SA node - or ectopic), the heart rhythms can be identified as:

- **Sinus Node Rhythms**, which are due to impulse conduction from the sinus node. This include not only arrhythmia (Sinus arrhythmia) but also the normal sinus rhythm, i.e, the usual cardiac rhythm of the heart;
- **Atrial Arrhythmia** that originate from ectopic sites in the atria;
- **Junctional Arrhythmia** and Atrioventricular Blocks originate in the area around the AV node and the bundle of His, called AV junction;
- **Ventricular Arrhythmia** and Bundle-Branch Blocks that create the impulse in the ventricles, below the bundle of His.

An important step towards identifying an arrhythmia is the classifications of heartbeats, which helps classifying the rhythm of an ECG signal, given by the classification of several consecutive heartbeats in that signal [9]. In this work, four different main type of heartbeats were considered and are described below.
2.3.1 Sinus node originated Heartbeats

This class comprises all the heartbeats that are originated in the sinus node, which includes the normal heartbeats or the bundle branch block beats. The normal heartbeat type was already described above. The remaining types included in this class are described below.

Left bundle branch block beats (LBBB)

Left bundle branch block beat [10], where the conduction in the left bundle is slow and results in a delayed depolarization of the left ventricle. The criteria used for this type of heartbeat is as following:

- QRS complex $>0.12$ seconds;
- Broad monomorphic R waves in I and V6 with no Q waves;
- Broad monomorphic S waves in V1, may have a small R wave.

An example of this type of beat is represented in Figure 2.3.

![Figure 2.3: Left bundle branch block heartbeat [11].](image)

Right bundle branch block beats (RBBB)

In the right bundle branch block beat (displayed in Figure 2.4), the conduction in the bundle to the right ventricle is slow. While the right ventricle depolarizes, the left ventricle is, most of the times, halfway finished. The criteria used for RBBB is:

- QRS complex $>0.12$ seconds;
- Slurred S wave in lead I and V6;
- RSR' pattern in V1 where R'>R.
Atrial Escape Beats

An atrial escape beat happens after a long sinus pause due to a sinus node exit block or sinus node arrest, that is, when the depolarization that occurs in the sinus node is unable to leave the sinus node towards the atria. This type of beat can become a sustained atrial rhythm when three or more escape beats happen in a row at a rate above 60 beats-per-minute. An atrial escape beat example can be seen in Figure 2.5 and the ECG characteristics for this type of heartbeat are the following [12]:

- Every P wave is followed by a QRS complex but the shape of the P wave is different than that of the normal beat;
- QRS complex is narrow and the heart rate is generally above 60/minute.
**Nodal (junctional) escape beats**

This type of heartbeat is a delayed one that originates from an ectopic focus somewhere in the AV junction. It happens when the rate of depolarization of the SA node falls below the rate of the AV node or when it fails to reach the AV node because of a SA or AV block. The nodal, or junctional, escape beat serves as a protection for the heart’s mechanism, in order to compensate for the SA node inability of acting as a pacemaker [7]. The ECG characteristics for this heartbeat are:

- Rate of 35 to 60 bpm;
- Irregular rhythm in single escape junctional escape complex;
- The P-wave depends on the ectopic focus site. They may be inverted or appear either before or after the QRS complex or even be absent, hidden by this complex;
- The QRS complex is usually normal in duration and morphology, of less than 0.12 seconds.

The junctional escape beat is represented in Figure 2.6.

![Figure 2.6: Nodal (junctional) escape beat](image)

**2.3.2 Supraventricular Ectopic Beats**

The supraventricular ectopic beats consist in beats that result from electrical impulses with supraventricular focus (usually in the atria) [7], that can happen prematurely, being also called, supraventricular premature beats. This class includes all the types that are described below.

**Atrial premature beats (APB)**

The atrial premature beat originates from an ectopic pacing region in the atria and results in a p-wave with a different morphology from the preceding ones. If this beat follows early after a sinus beat, some of the conducting tissues may not conduct. In this way, a premature atrial complex may have different rates [14], as shown in Figure 2.7:

1. Normally conducted;
2. Conducted with aberrancy, i.e. a conduction of the supraventricular impulse to the ventricles in a markedly different way than the usual conduction [7]. Mostly right bundle branch block aberrancy occurs, since this has a longer refractory period;
3. Not-conducted. If the premature beat is very early, the AV node is incapable of conducting and the beat is not followed by a QRS complex.

ECG characteristics for this type of beat are as follows:

- Abnormal P-wave, which can be hard to see due to the fact that this wave is rather shapeless;
- Occurs earlier than expected;
- Is usually followed by a non-compensatory pause due to the fact that atrial depolarization enters the SA node and resets the sinus rhythm.

![Figure 2.7: Atrial premature beat three possible faiths, where the dots represent normal sinus beats and the stars represent the atrial premature beats. The first APB is conducted normally, the second one follows the previous sinus beat earlier and is conducted with RBBB aberration and the third one is still a little bit earlier after the previous sinus beat and is blocked in the AV node and is, therefore, not conducted and results in a present P-wave but no QRS complex is present, followed by a non-compensatory pause [14].](image)

**Nodal (junctional) premature beats**

Premature junctional beats are premature cardiac impulses that originate from the AV junction [15]. They may arise either in the AV node or in the bundle of His. Their main ECG features are:

- The P-wave may appear after complex, within the ST-segment or the T-wave;
- P-wave may not appear at all, which may be due to burial of the wave within the QRS complex.
- P-wave may appear before the QRS complex with a PR interval that is usually short, less than 0.12 s.

An example of a premature junctional beat can be seen in Figure 2.8.

### 2.3.3 Ventricular Ectopic Beats

A ventricular ectopic beat is an extra-heartbeat that originates in the ventricles, the lower chambers of the heart. They can also be called premature ventricular contractions because they occur before the beat that is triggered by the SA node, the normal site for the heart’s activation [16] (Figure 2.9). The ECG characteristics for premature ventricular contractions are described below:
Figure 2.8: Nodal (junctional) premature beat [13].

- Abnormal and wide QRS complex that occurs earlier than expected in the cardiac cycle;
- Negative ST-T segment, directed oppositely to the QRS complex;
- Preceding P-wave is absent;
- Followed by a compensatory pause.

Figure 2.9: Premature ventricular contraction [17].

This class also comprises the type below:

**Ventricular escape beats**

A ventricular escape beat is an aberrant impulse that follows a sinus pause, occurs late and is initiated in the ventricles (Figure 2.10) [18].

Figure 2.10: Ventricular escape beat [13].

It has the following ECG features:

- Absent P-wave;
- Abnormal wide QRS complex of over 10 seconds;
- They are not preceded by a pause, being often followed by a compensatory pause.
2.3.4 Fusion Beats

A fusion beat is a beat that’s triggered by more than a single electrical impulse. Ventricular fusion beats occur when two separate pacemakers compete for the control of the ventricles. The usual combination of pacemakers is the SA node and ectopic ventricular focus combination and it’s the one that is mentioned and used in this work [19]. The criteria for the recognition of this type of heartbeats has not been precisely formulated yet. Many criteria have emerged for classifying a beat as a fusion heartbeat, some of them are as following:

- The contour and duration of the QRS complex, that must be in the middle of the typical durations for the competing pacemakers;
- Width of the complex is not more than 0.06 second wider than the normal sinus one;
- Beat must occur at a moment when impulses from both sites might be expected;
- P-S interval cannot be shorter than the P-R of the SA beat;

This type of beat can be seen in Figure 2.11.

![Figure 2.11: Fusion Beat [13]](image)

2.4 Standard 12-Lead ECG

The 12-lead ECG, also referred to as the standard ECG, is a non-invasive and valuable diagnostic tool that records the hearts electrical activity as waveforms. When it’s interpreted correctly, it can detect and monitor a variety of heart conditions, from arrhythmia to coronary heart disease to electrolyte imbalance. This type of ECG makes a complete picture of the heart’s electrical activity by doing the recordings from 12 different perspectives. This ECG displays, as its name implies, 12 leads, which are derived from 10 electrodes, which placement is shown in Figure 2.12. The difference between an ECG electrode and an ECG lead is that the electrode is a conductive pad that is attached to the skin and enables the recording of the hearts activity and the leads is a graphical representation of the electrical activity of the heart and is created by analyzing several electrodes. The ECG has two sets of ECG leads: limb and chest (precordial) leads. Limb leads include three leads that are the result from comparing the electrical potentials of 2 electrodes, where one serves as the reference and the other is the one doing the exploring (standard bipolar leads I, II and III) and three augmented leads, that have two or three augmented leads (aV R, aV L and aV F). These leads record the information in the hearts frontal plane.
A transverse view is given by the precordial leads ($V_1$, $V_2$, $V_3$, $V_4$, $V_5$ and $V_6$). For this work, only lead I signals were used for the task at hand.

Figure 2.12: Lead placement for a 12-lead ECG [18].
Chapter 3

State-of-the-Art

This presents a literature review of ECG denoising, feature extraction and classification techniques.

3.1 ECG Denoising Techniques

As mentioned previously, ECG has been widely used for the diagnosis of heart diseases. In order to facilitate the interpretation of the signals and detection of heart pathological phenomena by the physicians, it is necessary to have ECG signals with good quality and with the minimum number of artifacts as possible. In ECG signals artifacts and noise are common, being the most common one the high frequency noise caused by the electromyogram (EMG) induced noise, the power line interference or the mechanical forces acting on the electrodes, which severely limits the utility of the recorded ECG. For this reason, it is necessary to remove these elements from the signal, in order to allow for a better readability and, consequently, a better clinical evaluation. Several methods of ECG denoising techniques have been and are being investigated, some of them being described in this section.

Weng et al. [20] propose an ECG denoising technique based on Empirical mode decomposition (EMD), which is able to remove high frequency noise with a minimum signal distortion. They validated their results using the MIT-BIH database, described below in section 5.1.2. The method was compared with the widely used Butterworth lowpass filtering method and demonstrated better results in terms of visual quality of the signal. In addition, the signal-to-noise ratio (SNR) is calculated for the studied method and equals 18.3 dB, which is greater than 17.45 dB achieved by the low-pass filter.

SadAbadia et al. [21] conducted a mathematical method based on the dominant scale of QRS complexes, derived from the equation 3.1, where $D_{QRS}$ is the QRS complex duration and $T_s$ is rescaled wavelet coefficients, and the locations of R-waves.

\[
a_{QRS} = 1.00 \left( \frac{D_{QRS}}{T_s} \right)
\]  

(3.1)

The task is done by using a varying length window that moves over the whole signals. Both noise (high frequencies) and base-line wandering removal tasks are evaluated for manually corrupted ECG signals and are validated for real ones. They concluded that even though it is a simple method, of fast
implementation and that preserves ECG waves, which makes it a suitable algorithm, they encountered
difficulties due to the pre-stage detection of QRS complexes and the specification of the algorithms’
parameters for varying morphology cases.

Üstündag et al. [22] presented a denoising method for weak ECG signals based on fuzzy threshold-
ing and wavelet packet analysis. Firstly, they decomposed the weak ECG into several levels using the
wavelet packet transform and then, the threshold value was determined by using the fuzzy s-function and
lastly they reconstructed the signal using inverse wavelet packet transform. Several experiments were
performed to show the effectiveness of the proposed method and were compared with the traditional
wavelet packet soft and hard thresholding methods for this type of ECG signals. The results showed that
the proposed technique achieved notable results when the signal-to-noise ratio (SNR) is about -20 dB.
As for the comparison with the usual thresholding methods, they calculated the correlation coefficient
to compare their quality. According to these results, fuzzy thresholding performs notably well compared
to the usual thresholding, showing superior SNR rates, i.e., superior performance when compared to
existing methods.

AlMahamdya et al. [23] studied different denoising techniques for ECG signals. They compared five
common and important methods, applied on real ECG signals with different levels of noise: discrete
wavelet transform, applying both universal and local thresholds; adaptive filters (least mean-squares
LMS - and recursive least-squares - RLS) and Savitzky-Golay filtering. They studied signals a noise
levels from 5 dB SNR to 45 dB SNR and the studies showed that the NeighBlock wavelet algorithm
performs better then the others but that RLS and Savitzky-Golay filters perform better in some mid-
range SNRs.

V. Supraja and S. Safiya [24] proposed a new denoising method, which they call improved thresh-
olding denoising method, that can be viewed as a compromise between soft and hard thresholding
methods. The method selects the best suitable wavelet function based on the discrete wavelet trans-
form (DWT) at a decomposition level of 5, using the mean squared error (MSE) and output SNR. This
method had the advantages of retaining both the geometrical characteristics of the original signal and
the variation of the ECG waveforms effectively. The resulting method showed to outperform the tradi-
tional wavelet thresholding denoising methods in the aspect of the remaining geometrical characteristics
as well as the in the improvement of the SNR.

More recently, Alyasseri et al. proposed a hybrid method of $\beta$-hill climbing combined with the wavelet
transform for denoising ECG signals. Selecting the mother wavelet is always a challenging task and is
usually performed based on empirical evidence. For this reason, $\beta$-hill climbing has to find the optimal
wavelet parameters for the ECG signal denoising, aiming to obtain the minimum mean square error
between the original and the denoised ECG signals. The tests were done using the MIT-BIH database
and the method was evaluated using, as criteria, the percentage root mean square difference (PRD) and
SNR. This method obtained outstanding results in terms of noise removal in ECG signals, particularly in
those with high noise. It achieved a high SNR and a low PRD.
3.2 Heartbeat Classification

Automatic ECG beat classification is essential to diagnose and monitor dangerous heart conditions. It is a very time consuming job for doctors to analyze long ECG records, therefore, many computer based methods have been proposed for automatically diagnosis of the ECG beat abnormalities. A large number of algorithms was developed both to differentiate between different types of beats and to detect and classify arrhythmia. These vary widely in terms of features extracted from the ECGs and classification methods.

Duration of waves and intervals, regularity/irregularity of the rhythm, correspondence between the P waves and the QRS complexes, amplitudes and polarities, are examples of features that have been explored in automatic ECG analysis. Researchers have therefore been able to deal with statistical parameters. In classification methods, some authors opted for heuristic knowledge based methods, such as decision rules, whilst others chose to explore more recently developed classifiers such as artificial neural networks or support vector machines.

In this section studies that focused on feature extraction, and automatic beat or rhythm classification are reviewed.

3.2.1 Feature Extraction and Classification

The classification of cardiac arrhythmia can be achieved after extracting the features of each heart beat in the ECG signal. Several methods have been proposed for extracting features of cardiac beats. The features of one cardiac cycle may be divided in time domain features and frequency domain features. Also, for cardiac arrhythmia classification several techniques have been proposed.

In this section an overview of several techniques is made and a summarized table with the most relevant methods is presented (Table 3.1).

Over the published work, the most frequent database used for benchmarking of arrhythmia classification systems is the MIT-BIH arrhythmia dataset. However, several reported work evaluates their methods over private datasets or without datasets references. In the following, unless specifically stated, this is the underlying validation setup.

Inan et al. [25] show that morphological information and timing information can provide high classification accuracy for larger datasets. When combining wavelet domain feature with RR interval features it's possible to achieve high classification accuracy as reported in [26]. Chazal [9] shows that high classification accuracy is achieved when the morphological feature is combined with the temporal feature of each patient specific data. The authors Dutta et al. [27] had proposed cross-correlation based feature for classifying PVC beats from non-PVC beats. They have used cross-correlation between each ECG heart beat signal with the normal heart beat signal which is chosen as reference signal. Khazaee et al. [28] extracted power spectral density (PSD) features of each heart beat with three timing interval features classifying cardiac abnormalities in MIT-BIH database.

Inan et al. [25] presented an approach for classifying beats of a large dataset by training a neural network (NN) classifier using wavelet and timing features. Inan et al. found that the fourth scale of a
Table 3.1: Summary of the most relevant methods for features extraction and classification presented in this section. The unavailable information is marked with an asterisk (*).

<table>
<thead>
<tr>
<th>Ref</th>
<th>Features</th>
<th>Classification Method</th>
<th>Database</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Kernel-independent component analysis Discrete wavelet transform (morphological and temporal features)</td>
<td>SVM</td>
<td>MIT</td>
<td>98.8%</td>
</tr>
<tr>
<td>24</td>
<td>Morphological features Two timing interval features.</td>
<td>SVM</td>
<td>MIT</td>
<td>99.8%</td>
</tr>
<tr>
<td>26</td>
<td>Filters Interpolation RR intervals</td>
<td>Artificial Neural Network</td>
<td>MIT</td>
<td>92.34%</td>
</tr>
<tr>
<td>27</td>
<td>RR intervals</td>
<td>Deterministic Automaton</td>
<td>MIT</td>
<td>98%</td>
</tr>
<tr>
<td>28</td>
<td>RR intervals</td>
<td>Decision Tree</td>
<td>*</td>
<td>99.1%</td>
</tr>
<tr>
<td>29</td>
<td>Adaptive filters (several features)</td>
<td>Autocorrelation Functions</td>
<td>MIT</td>
<td>*</td>
</tr>
<tr>
<td>32</td>
<td>RR interval features Heartbeat interval features Morphology features</td>
<td>Linear Discriminant</td>
<td>*</td>
<td>81.9%</td>
</tr>
<tr>
<td>34</td>
<td>Morphological and temporal QRS complex duration RR interval and average RR interval</td>
<td>SVM</td>
<td>*</td>
<td>89.72%</td>
</tr>
<tr>
<td>39</td>
<td>RR intervals 23 selected wavelet transform coefficients</td>
<td>Artificial Neural Network</td>
<td>*</td>
<td>96.77%</td>
</tr>
<tr>
<td>40</td>
<td>Wavelet coefficients Statistical properties of the coefficients Electrophysiological measures RR interval information Amplitude of waves Ratio between intervals</td>
<td>Clustering and SVM</td>
<td>*</td>
<td>98.92%</td>
</tr>
<tr>
<td>43</td>
<td>S-transform based features Temporal features Mixture of ST and WT based features</td>
<td>Multilayer Perceptron Neural Network</td>
<td>MIT</td>
<td>N - 95.70%</td>
</tr>
<tr>
<td>35</td>
<td>PCA in 29 samples from QRS Average RR-interval QRS complex width</td>
<td>Mixture of experts (SOM, LVQ)</td>
<td>MIT</td>
<td>95.52%</td>
</tr>
</tbody>
</table>

Dyadic wavelet transform with a quadratic spline wavelet together with the pre/post RR interval ratio is effective for distinguishing normal and premature ventricular contraction (PVC) from other beats.

An approach for personalized ECG heartbeat pattern classification is studied by Jiang, Kong and Peterson [29]. It is based on blockbased NNs, where a 2-D array of modular component NNs with flexible structures and internal configurations is implemented using re-configurable digital hardware. Network structure and connection weights are optimized using local gradient-based search and evolutionary operators with the rates changing according to their effectiveness in the earlier evolution period.

Moody and Mark (1983) [30] compared the performance of multiple classifiers in the detection of atrial fibrillation. Features were based on RR intervals which are known to be irregular when atrial fibrillation is present. Markov process models with different averaging techniques (e.g. filter and interpolation) were tested, as well as a RR predictor array. Improvements were noted with variations from the basic Markov process but all detectors showed room for improvement.

In an attempt to improve the performances achieved in [30], Artis, Mark and Moody (1991) [31] used an artificial neural network to detect atrial fibrillation. The inputs to the network were the cells of a so-called generalized interval transition matrix. Each RR interval was classified as short, normal or long...
and pairs of intervals were assigned to cells of the matrix. A sliding window of intervals was adopted, allowing a beat-by-beat classification. Atrial fibrillation sensitivity and positive predictive accuracy were respectively of 92.86% and 92.34%.

There are several other studies that continues to use RR-based features for ECG classification tasks. In [32] (2005) both beat and arrhythmia classifications were attempted. First, using a set of decision rules based on three RR intervals, beats were classified as normal, premature ventricular contractions, ventricular flutter/fibrillation or 2\textsuperscript{nd} heart block. The classifications were then used as input to detect and classify arrhythmic episodes, using a deterministic automaton. Six arrhythmia were considered: ventricular bigeminy, ventricular trigeminy, ventricular couplet, ventricular tachycardia, ventricular flutter/fibrillation and 2\textsuperscript{nd} heart block. In this study 98% accuracy for arrhythmic beat classification and 94% accuracy for arrhythmic episode detection and classification were achieved.

Kaiser, Kirst and Kunze (2010) [33] attempted to classify 5-minutes ECG records unto atrial fibrillation or no atrial fibrillation (including normal sinus rhythm and different arrhythmia). The authors developed low processing power techniques since the goal was to use the algorithms on mobile device. Features were extracted from the RR interval tachogram, where the maximum difference between any two RR intervals and the variance of the set of RR interval duration were computed. A decision tree based on threshold comparisons was then used to classify the segments. To improve the performance of the classifier a second analysis helped by morphological filters was included. The proposed algorithms correctly identified 436 of 440 five minute episodes of atrial fibrillation or flutter and also correctly identified up to 302 of 342 episodes of no atrial fibrillation and a sensitivity of 99.1% and a specificity of 88.3% were achieved.

Not all type of beats and rhythms can be distinguished so, many authors chose to explore the potential of other kind of features.

Thakor and Zhu (1991) [34] used adaptive filtering techniques applied to ECG signals and stated that arrhythmia detection issues could benefit from this analysis. They concluded that by adaptively cancelling the QRS-T complex one could detect the presence of ectopic beats because of their abnormal morphology or analyze a paced rhythm and monitor pacemaker performance and failure. Atrial fibrillation detection could also benefit from adaptive filtering since the atrial rhythm can be separated and autocorrelation functions may afterwards be used for classification.

In 2005, Chiu, Lin and Liau [35] attempted a distinction between normal beats, premature atrial contractions (PAC) and premature ventricular contractions (PVC). The authors used a template for each one of these beat types and computed the normalized correlation coefficient between QRS complexes of templates and beat to be classified, using a small testing set (9 PACs and 24 PVCs beats). A beat was classified as PVC if its correlation coefficient with the PVC template was higher than the one with the normal beat template. Additional information concerning the RR interval duration was used to define whether or not a beat should be classified as PAC. Positive predictive values of 99.44% for normal beats, 100% for PAC's and 95.35% for PVCs were reached. For sensitivity they reached 99.81, 81.82 and 95.83%, respectively for normal beats, PACs and PVCs.

To distinguish between normal ECGs, atrial premature contractions, premature ventricular contrac-
tions, supraventricular tachycardia, ventricular tachycardia and ventricular fibrillation, Srinivasan and Krishnan (2002) [36] used segments of 1.2 seconds around the R peak. ECGs were modelled using an autoregressive model and the coefficients were used to classify the arrhythmia by means of stage-by-stage generalized linear model. Different stages of the classifier were determined by the Euclidean distances between mean autoregressive coefficients from various classes. Accuracy values varied from 93.2% for normal segments to 100% for ventricular tachycardia segments.

Several authors used feature sets containing both temporal and morphological information.

Chazal, O’Dwyer and Reilly (2004) [37] compared the performance of 12 classifier configurations in distinguishing 5 different beat types. Three categories of features were used: RR interval features, heartbeat interval features and ECG morphology features. Classifier configurations were based on linear discriminant (LDs), a statistical classifier model. The best results were obtained when two leads were used for classification. For ventricular ectopic beats, the sensitivity was 77.7% and the positive predictivity was 81.9%. For supraventricular ectopic beats, sensitivity and positive predictive values of 75.9 and 38.5% were reached.

In [38], 2007, a set of decision rules was established to classify beats as normal or ventricular ectopic, using both morphological and time-domain features. A QRS pattern matrix was constructed with information about amplitude-temporal distribution of QRS and deviation of RR interval from the mean RR interval was also considered. Sensitivity and specificity values were reported to be respectively 99.81 and 98.87%.

Melgani and Bazi [39], 2008, performed several tests to compare the performance of different classifiers on distinguishing between normal beats, atrial premature beats, ventricular premature beats, right bundle branch block, left bundle branch block and paced beats. Two types of features were extracted, morphological and temporal (QRS complex duration, RR interval and average RR interval over the 10 last beats) and it was used support vector machines, the k-nearest neighbor and the neural network and some tests included a feature selection step with principal component analysis. The authors reported an overall accuracy of 89.72%.

The potential of neural networks continues to be one of the main machine learning classifiers used in ECG beat/rhythm analysis.

In Hu et al. [40], 1993, an artificial neural network was used to distinguish between normal and abnormal beat patterns, and to classify 12 different abnormal beat morphology. Inputs to the network consisted of an amplitude-scaled QRS beat pattern. A classification rate of 84.5% was obtained when using a composite classifier where the first network distinguished between normal and abnormal morphology and the accuracy varied between beat types.

Silipo and Marchesi (1998) [41] explored artificial neural networks in three different ECG analysis tasks: beat classification, myocardial ischemia and chronic alterations. Normal beats, ventricular ectopic beats and supraventricular ectopic beats were considered for classification. The authors used an artificial neural network structured as an autoassociator so as to be able to reject unknown patterns (new or ambiguous beats). The input vector to the network consisted of samples of the beat and a measure of its prematurity degree based on RR interval measurements. Tests with different beat types
included in the training set were performed. When the three beat types were considered, normal and ventricular ectopic beats were more easily recognized (99 and 96% respectively) than supraventricular ectopic beats (75%).

If artificial neural networks have drawn considerable attention as classifiers, wavelet transforms have earned their place in the feature extraction process. Compared to other frequency analysis methods, e.g. Fourier transform, wavelet analysis allows a multi-scale decomposition and overcomes some drawbacks in terms of frequency resolution.

Classification of normal sinus rhythm, atrial fibrillation, ventricular fibrillation and ventricular tachycardia ECG records was addressed in Khadra, Al-Fahoum and Al-Nashash, 1997.

Güler and Übeyli [42] (2005) performed a beat-by-beat classification to distinguish between normal, congestive heart failure, ventricular tachyarrhythmia and atrial fibrillation beats. Statistical features were extracted from the wavelet coefficients to try to represent both the frequency distribution of the signal and changes in frequency distribution. Mean of the absolute values of the coefficients in each sub-band, average power of the wavelet coefficients in each sub-band, standard deviation of the coefficients in each sub-band and ratio of the absolute mean values of adjacent subbands were the features used. It was employed a combined neural network model, where the outputs of the first set of networks acts like input to the second level network. An overall classification accuracy of 96.94% was achieved.

Martis et al. [43] (2012) compared the performance of a neural network (NN), a support vector machine (SVM) and a Gaussian mixture model (GMM) in the distinction between normal beats and 12 different beat types. Features were extracted from the wavelet decomposition of the signal. A feature selection method (principal component analysis) was applied to the sub-bands that covered the frequencies of interest. This reduced set of coefficients was used as input to the classifiers. Overall accuracy values of 87.36, 93.41 and 95.60% were obtained for GMM, NN and SVM classifiers respectively.

In [44], 2003, information about RR intervals was used in combination with 23 selected wavelet transform coefficients to distinguish between normal and 12 abnormal beat types. An artificial neural network was used as classifier and an overall accuracy of 96.77% was reached.

Shen et al., 2012, [45] achieved beat classification after applying an adaptive feature extraction method to a large set of features. These included wavelet coefficients, statistical properties of the coefficients and electrophysiological measures (RR interval information, amplitude of waves, ratio between different intervals, etc.). The classification scheme included the use of k-means clustering, one-against-one support vector machine and a modified majority voting mechanism. A recognition rate of 98.92% was reported by the authors but accuracy varied between beat types.

In [46], 2013 a beat classification scheme was proposed to distinguish between normal beats, left and right bundle branch block, atrial and ventricular premature contractions, paced and fusion beats. A combination of temporal, morphological and statistical features was used for classification. Some of these features were extracted after wavelet decomposition of the signal. Prior to classification, a feature selection method was employed. Multiple one-against-one support vector machines were employed and the final classification was given by a maximum voting mechanism. The authors reported an overall accuracy superior to 98%.
The work proposed by Patel [47], 2012, deals with the development of an efficient arrhythmia detection algorithm using ECG signal so that detection of arrhythmia at initial stages is possible using a smart-phone which is readily available anywhere which makes complete system mobile. Subjects for experiments included normal patients, patients with Bradycardia, Tachycardia, atrial premature contraction (APC), patients with ventricular premature contraction (PVC) and patients with Sleep Apnea. Pan-Tompkins algorithm was used to find the locations of QRS complexes and R Peaks. The algorithm to detect different arrhythmia is based on position of P wave, QRS complex, R Peak and T wave and on interval between these waves on android smart-phone. The algorithm was tested using MIT-BIH arrhythmia database. Results revealed that the system is accurate and efficient to classify arrhythmia as high overall performance (97.3%) for the classification of the different categories of arrhythmic beats was achieved.

In 2014, Manab Kumar Das and Samit Ari [48] proposed the design of an efficient system for classification of the normal beat (N), ventricular ectopic beat (V), supraventricular ectopic beat (S), fusion beat (F), and unknown beat (Q) using a mixture of features. In this paper, two different feature extraction methods are proposed for classification of ECG beats: (i) S-transform based features along with temporal features and (ii) mixture of ST and WT based features along with temporal features. The extracted feature set is independently classified using multilayer perceptron neural network (MLPNN). The performances are evaluated on several normal and abnormal ECG signals from 44 recordings of the MIT-BIH arrhythmia database. In this work, the performances of three feature extraction techniques with MLP-NN classifier are compared using five classes of ECG beat recommended by AAMI (Association for the Advancement of Medical Instrumentation) standards. The average sensitivity performances of the proposed feature extraction technique for N, S, F, V, and Q are 95.70%, 78.05%, 49.60%, 89.68%, and 33.89%, respectively.

### 3.3 Wavelet Transform

Recently, wavelet transform has been widely used in several applications, including the processing of non-stationary signals, feature extraction for classification tasks, heart rate variability analysis, ECG data compression, among others. The wavelet transform is similar to the Discrete Fourier transform. Their main difference is the fact that fourier transform (FT) will tell what frequencies are present in the signal but a wavelet transform (WT) will tell what frequencies are present and where (or at what scale). This transform is a multi-resolution analysis and, consequently, overcomes the difficulty of finding an optimal resolution for the analysis of the signal. It allows the examination of the low frequency content over a larger amount of time without compromising the accurate time domain localization of the high frequencies. After the choice of an analyzing wavelet function, ψ(t), the wavelet transform of a continuous time signal x(t) is defined as Equation 3.3:

\[
W_a \psi (b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \psi^* \left( \frac{t - b}{a} \right) dt
\]
where \( a \) and \( b \) are the dilatation and time parameters, respectively.

For a function to be considered an analyzing wavelet, or mother wavelet, it should satisfy some mathematical properties, in particular, the function needs to have finite energy and respect the admissibility condition that implies a zero average [49]. A family of functions can be derived from a mother wavelet by applying translations and dilatation. The wavelet transform can therefore be seen as the decomposition of the signal by this set of basis functions. A discretized version of the continuous wavelet transform described above is often used and a dyadic grid, with \( a = 2^m \) and \( b = 2^m n \) is commonly employed. Then, the basis functions can be written as Equation 3.3:

\[
\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m} t - n)
\]  

A fast implementation of the decomposition algorithm can be achieved by recursively applying high pass and low-pass filters to the wavelet coefficients from the previous scale. The filter coefficients of the high-pass and low-pass decomposition filters characterize the wavelet used.

The redundant discrete wavelet transform (RDWT) can also be used to decompose the signal. This discretized version of the continuous wavelet transform is shift invariant, having the same spatial sampling rate in all scales.

In the last couple of decades, wavelet theory has been widely used for the applications mentioned above [49]. The choice of mother wavelet is an important step in all of these applications and several functions have been tested, including Daubechies, Morlet, spline, raised cosine and quadratic spline wavelets.

### 3.4 R peak Detection

R peak detection is a very important part of this work so an algorithm based on the work of [50] was used to detect the R peaks. The detection method starts with a filter step. First, low pass and high pass filters are used on the signal, with cutoff frequencies of 16 and 8 Hz, respectively. The derivative is then calculated and its absolute value is computed. The last step consists of computing the moving average method, with an 80 ms window over the signal. Once the peaks are detected, they need to be classified wither as a QRS complex or noise. This decision is made using a set of detection rules that rely on peak height, peak location (relative to the last QRS peak) and maximum derivative. The accuracy of the algorithm depends on the computation of a threshold for the detection, defined using QR peaks and noise peaks heights, as shown in equation 3.4

\[
\text{DetectionThreshold} = \text{AverageNoisPeak} + TH \times (\text{AverageQRSPeak} - \text{AveragePeakNoise})
\]  

(3.4)
were TH is the threshold coefficient that represents a compromise between correct and false detections. Decreasing this value will lead to a higher number of correct detections, at the expense of false detections.

The following detection rules are applied in the algorithm:

1. Peaks that precede or follow larger peaks by less than 200 ms are ignored;

2. If there is a peak, there is a check for the presence of both positive and negative slopes in the raw signal. If they don’t exist, the peak represents a baseline shift;

3. If the peak occurred within 360 ms of a previous detection, the value of the maximum derivative in the signal is checked. If it is smaller than half the maximum derivative of the previous one, the peak is considered to be a T-wave;

4. If the peak is larger than the threshold, it is considered a QRS complex. If not, it is considered to be noise;

5. A peak should be considered a QRS complex if:

   • no QRS has been detected within 1.5 RR intervals;
   • there was a peak that was larger than half the detection threshold;
   • and the peak followed the preceding detection by at least 360 milliseconds

### 3.5 Classifiers

Two classifiers were used for the classification task, described below: k-nearest neighbor and support vector machine.

#### 3.5.1 Support Vector Machine (SVM)

A Support Vector Machine is a state-of-the-art maximum margin classification algorithm rooted in statistical learning theory [51]. It is a method for classification of both linear and non-linear data. SVM is a supervised machine learning algorithm that can be used for both classification and regression tasks, being mostly used for classification. In this algorithm, each data item is plotted as a point in an n-dimensional space (where n is the number of features used in task at hand), where the value of each feature is the value of a particular coordinate. Then, the classification is performed by finding the hyperplane that best differentiates the two classes [52] (Figure 3.1). In other words, given labeled training data, the algorithm outputs an optimal hyperplane which categorizes new examples.

A supervised SVM training is done by placing a line (for 2-D dimensional data) or a hyperplane between two different classes by maximizing the margin between all datapoints. To find an optimal separator, the Equation 3.5 with the constraints $a_j \geq 0$ and $\sum_j a_j y_j = 0$ has to be solved. This quadratic programming optimization problem can be solved by a Lagrangian and the Kuhn-Tucker theorem.
As some data is not linearly separable, it may be required to move the data into a higher dimensional space. To obtain high dimensional datapoints $x$, they will be transferred into a high dimensional feature space by $F(x)$. This means that we can replace the $x_j \cdot x_k$ with $F(x_j) \cdot F(x_k)$. A possible function $F$ for the sample could introduce an additional dimension by $F(x) \rightarrow x^2_1, x^2_2, \sqrt{2}x_1x_2$. With this function it can be shown that

$$F(x_i) \cdot F(x_j) = (x_i \cdot x_j)^2$$

The equation $(x_i \cdot x_j)^2$ is called kernel function and is often written as $k(x_i, x_j)$. This also represents the so called kernel trick, which means that all occurrences of $(x_i \cdot x_j)$ are replaced by the kernel function and therefore allow to find optimal solutions in an arbitrary high dimensional feature space. The found solution can be afterwards retransferred into the original space and result into arbitrary shapes.

$$k(x_i, x_j) = e^{-\frac{\|x_i-x_j\|^2}{2\sigma^2}} = e^{-\gamma\|x_i-x_j\|^2} \text{ with } \gamma = \frac{1}{2\sigma^2}$$

For this work, the radial basis function (Equation 3.7) was used as the kernel function, where the kernel parameter $\gamma$ (or $\sigma$), which is defined by the user, is responsible for controlling the kernels width. The penalty parameter of the error term ($C$) must also be input and adapted to the task at hand. $C$ and $\gamma$ were varied between $[10^{-1}, 10^5]$ and $[10^{-5}, 10^0]$, respectively.
3.5.2 K-Nearest Neighbours (KNN)

The concept behind the k-nearest neighbors (kNN) classifier is a very simple one. Assuming we have a dataset for which we know the true labels of the samples, new samples can be classified according to their similarity with labeled samples.

The principle behind nearest neighbor methods is to find a predefined number of training samples closest in distance to the new point, and predict the label from these. Assuming we have a dataset for which we know the true labels of the samples, new samples can be classified according to their similarity with labeled samples. A set of features is chosen to represent each sample and the similarity between samples is measured resorting to a metric - standard Euclidean distance is the most common choice. If desired, different weights can be attributed to the neighbors, assuring for instance that closest neighbors contribute in a larger extent to the fit. The ‘training set’ is used to construct a neighbor base and the test patterns are classified according to the class of the k most similar neighbors. An example of binary classification is shown in Figure 5.12. The $k$-NN does not require an explicit training phase, it classifies new instances “locally” by looking at the category labels of the $k$ most similar training examples [53].

![3-Class classification (k = 15, weights = 'uniform')](image1)

![3-Class classification (k = 15, weights = 'distance')](image2)

Figure 3.2: Example of classification using KNN algorithm [53].

These classifiers were implemented using Python programming language and the scikit-learn Python module, a machine learning kit developed for this language [53].
Chapter 4

Proposed Methodology

This thesis aimed to develop a system for automatic classification of heartbeats. This was done by focusing on two main tasks: denoising and filtering techniques; and classification of arrhythmic heartbeats. Three different databases were used in this work, depending on the type of task at hand. This chapter describes, firstly, the types of heartbeats that were to be classified and then the proposed methodology for both the denoising and the classification tasks. The methodology used for both the denoising and the classification tasks are schematized in figures 4.1 and 4.2, respectively.

4.1 Types of Arrhythmia

In this work, the main goal, as mentioned before, was to develop a system for automatic classification of arrhythmia in ECG recordings. The classes of arrhythmia that were used in this task have as a basis the Association for the Advancement of Medical Instrumentation (AAMI) [54] recommended practice. This combines the MIT-BIH heartbeat types into only five heartbeat classes, described previously in section 2.3. Each class includes heartbeats of one or more types. Table 4.2 describes each class and which types of heartbeats belong to each class. Class N consists of beats that originate in the sinus node (normal and bundle branch block beat types), class S consists of supraventricular ectopic beats, class V are ventricular ectopic beats, F contains beats that result from fusing normal and V beats and, lastly, class Q contains unknown beats, including paced beats. For this work, some changes were made to
this recommended practice (Table 4.1), namely:

- Label for left bundle branch block beat (L) was added to the classification system. The L heartbeat is an arrhythmia type that should also be taken into account because, even if most times it is harmless, it can also be an indication of heart disease. This type also has a similar morphology to the Ventricular ectopic beats (V class) and it is interesting to measure the performance of the system in distinguish these two classes. The right bundle branch block beat (R) was also considered but its classification is not best done using lead I signals, which is the one that was used in this work so, this type of bundle branch block heartbeat was considered to be a class N heartbeat.

- Since the purpose was to identify pathological beats, the Q label was excluded from all the experiments, since they were mainly unknown and unclassifiable beats, which wasn’t of interest for this study.

### Table 4.1: Association for the Advancement of Medical Instrumentation (AAMI) recommended practice based classes. [54]

<table>
<thead>
<tr>
<th>AAMI heartbeat class</th>
<th>N</th>
<th>Q</th>
<th>S</th>
<th>V</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beats originating in the sinus node except for left bundle branch block beat</td>
<td>Unknown Beat</td>
<td>Supraventricular ectopic beat</td>
<td>Ventricular ectopic beat</td>
<td>Fusion beat</td>
<td></td>
</tr>
</tbody>
</table>

### Table 4.2: Classes considered for this thesis.

<table>
<thead>
<tr>
<th>AAMI heartbeat class</th>
<th>N</th>
<th>L</th>
<th>S</th>
<th>V</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beats originating in the sinus node except for left bundle branch block beat</td>
<td></td>
<td>Left bundle branch block beat</td>
<td>Supraventricular ectopic beat</td>
<td>Ventricular ectopic beat</td>
<td>Fusion beat</td>
</tr>
</tbody>
</table>

The amount of heartbeats that exist in the databases for each label type considered for this thesis is presented in Table 4.3.
Table 4.3: Number of samples for each label in the MIT-BIH and StM databases.

<table>
<thead>
<tr>
<th>MIT-BIH (number of samples)</th>
<th>StM (number of samples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>101513</td>
</tr>
<tr>
<td>L</td>
<td>8826</td>
</tr>
<tr>
<td>S</td>
<td>1752</td>
</tr>
<tr>
<td>V</td>
<td>1028</td>
</tr>
<tr>
<td>F</td>
<td>33</td>
</tr>
</tbody>
</table>

4.2 Denoising

The main goal in this part of the denoising and outlier removal method, was to take a signal that was corrupted with noise, from now on referred to as the "Noisy Signal", and use several denoising techniques in order to get the cleanest signal possible, measuring the results against a baseline signal with least possible noise added to it, referred to as the "Clean Signal" hereafter. Both these signals correspond to the same time period and have the same heartbeats, acquired by different sensors. An example of these two signals can be seen in Figure 4.3.

Two approaches for signal denoising are proposed, based on:

1. Data filtering;
2. Wavelet Decomposition.

For filter based approaches, the methods described in 4.3.1 and 4.3.3 are included. 2 includes the method described in 4.3.2.

Before applying all the approaches described below, there was a need to remove the baseline wonder from the signals. For that purpose, 2 median filters were applied both to the noisy signal and the clean signal, using window sizes of 0.2 and 0.3 s.

The general steps for evaluating the denoising method were as follows:

1. R peak detection for both signals, that returns 2 arrays of R peaks positions for both signals;
2. compare the locations from both signal’s peaks and consider a correctly detected peak one that is located in the exact same position on both arrays or with a maximum difference of + or - 20 samples;
3. compute the heartbeat related to each "true" peak. For this, a window of 200 ms to the left and 400 ms to the right of the R-peak was used. This fixed-window definition of the heartbeat was used always throughout the developed work;
4. calculate the Signal-to-Noise Ratio - SNR - (Equation 4.1) and cosine similarity - Cossim - (Equation 4.3) between each computed heartbeat, after a normalization step, comparing the beats related to the similar R-peak from both the denoised and the clean signals;

\[
SNR = \frac{\text{RMS(DenoisedSignal)}}{\text{RMS(Noise)}}, \quad \text{where} \quad \text{Noise} = \text{DenoisedSignal} - \text{CleanSignal} \quad \text{and:}
\]

\[
(4.1)
\]
\[ RMS = \sqrt{\frac{1}{N} \sum_{n} x^2[n]} \]  \hspace{1cm} (4.2)

\[ \text{CosineDistance} = 1 - \text{CosineSimilarity}, \text{ \ where:} \]  \hspace{1cm} (4.3)

\[ \text{CosineSimilarity} = \frac{\text{CleanSignal} \cdot \text{DenoisedSignal}}{\|\text{CleanSignal}\| \|\text{DenoisedSignal}\|} \]  \hspace{1cm} (4.4)

5. for each subject, the mean, mode and median of each measurement are calculated for all the detected beats.

For each approach, the specific steps that were used are described below.

Figure 4.3: Top: Example of the baseline signal used for measuring the results of the denoising technique (Clean Signal); Bottom: Example of signal corrupted by noise (Noisy Signal).

### 4.2.1 40 Hz Low-Pass Filter

After the baseline wander removal, a finite impulse response low-pass filter with cutoff frequency of 40 Hz was used to deal with the high frequency noise. These two signals presented a slight misalignment
so, in order to be able to evaluate the denoising results more accurately, an alignment between the 2 filtered signals was done using an algorithm based on the cross-correlation between them. These 2 filtered and aligned signals served as the baseline for comparison between the several denoising methods. The steps taken to evaluate this method were the ones described in the above section.

### 4.2.2 Denoising using Wavelet Thresholding Techniques

Based on previous work, the following wavelets and levels were used:

- Function = "Sym12" and level = 3 which, according to [23] gives the best performance of denoising for both universal and interval-dependent thresholds;

- Function = "Db4" and level = 5, as in [55] (Figure 4.4);

- Function = "Quadratic spline" and level 6, which has been mainly used for ECG delineation purposes because of the relation between ECG peaks and zero-crossings in the wavelet transform [56]. This wavelet (Figure 4.6) has been mainly described to extract features for ECG classification but was also used for denoising in this work.

![Figure 4.4: Mother wavelet the Db4 wavelet](image)

![Figure 4.5: Mother wavelet of the sym12 wavelet](image)

For each type of wavelet, two different approaches were evaluated:
1. Applying the wavelet denoising algorithm after going through a 40 Hz Low-pass filter step to both the clean and the noisy signals.

2. Applying the wavelet denoising algorithm straight after the baseline wander removal for the noisy signal but after removing the high frequency noise from the clean signal.

The wavelet denoising algorithm, unlike the Low-pass filter which was applied to the whole signal, is applied after step 3 from the general steps, that is, it is applied to each heartbeat. Each heartbeat goes through a wavelet decomposition, resulting in a specific number of coefficients, depending on the level chosen for the wavelet decomposition (specified above). Then, a soft thresholding technique is applied to each coefficient, using the universal threshold proposed by Donoho and Johnstone [57]. Lastly, a wavelet reconstruction is applied, leading to the denoised heartbeat, to each steps 4 and 5 from the general steps are going to be applied.

The applied threshold is given by Equation 4.5:

$$\lambda = \alpha \sqrt{2 \log(n)}$$  \hspace{1cm} (4.5)

where \(n\) is the sample size, which, in this case, is 600 samples (600 milliseconds) and \(\alpha\) is given by Equation 4.6:

$$\alpha = \frac{MAD}{0.6745}$$  \hspace{1cm} (4.6)

where \(\alpha\) is calculated based on the last level of the detail coefficients, according to the median absolute deviation (MAD) [58]. The factor in the denominator is the scale factor.

### 4.2.3 Moving Average

As a last method, the moving average filter was used, following the general steps mentioned above with the difference that this filter was applied to both the noisy and the clean signals straight after step 1,
4.3 Outlier removal

Besides noise contamination, signals may present artifacts (acquisition / connectivity failures; ...) that lead to too corrupted segments, that will be considered as outliers that should be removed from the signals to achieve better results. An outlier removal technique was used during this thesis, and consisted in the development of beat prototypes. After calculating the RR peaks and computing the heartbeat using the same fixed window as described above, the outlier removal was done, comparing each detected beat to each of the prototypes and which considered a sample as an outlier if it’s cosine similarity to every prototype was less than 0.8, a value that was chosen in order to have better results without losing a big part of the signal. Since, as mentioned above, only 5 type of beats are of interest for this experiment, 50 prototypes were developed (10 for each label) using a labeled database. In order to obtain prototypes for each of the classes (pathological or normal), the corresponding available labeled data was first segmented into individual heartbeats, as explained before and K-means clustering was applied over these heartbeats. The value of K was empirically set to 10, the number of clusters per class. Given the dependence of the K-means on the initialization, this algorithm was run 20 times, selecting as final solution the partition corresponding to the smaller mean square error, and the corresponding centroids were chosen as class prototypes.

After removing every possible outliers, the denoising algorithm that performed better from the ones presented in the previous section is applied to the remaining beats and the SNR and cosine similarity of the denoised signals are calculated as described on steps 4 and 5 from the general steps in the above section.

4.4 Feature Extraction

To begin the classification task of the thesis, a feature extraction process needed to be done. Two different type of features were used and also a combination of both. The types used were temporal features and morphological wavelet based features. The classification part of the thesis was done using the SIM and the MIT-BIH databases and the feature extraction was done to both, using the same set of features. These features were chosen based on [59] and [25].

4.4.1 Temporal Features

Four temporal features were extracted from the RR-intervals (distance between 2 consecutive R peaks) of the pre-processed ECG signals. The features used are described below:

- Pre-RR-interval: Distance from a given heartbeat to the following one;
- Pos-RR-interval: Distance between a given heartbeat and the previous one;
• Average RR-interval: The mean of all the RR-intervals from a recording of a subject;

• Local average RR-interval: Mean of the RR-intervals computed over 10 heartbeats centered at the given point.

### 4.4.2 Morphological Features

For this thesis, wavelet-based morphological features were used in the task of classifying ECG beats, which were extracted from each ECG cardiac cycle by selecting the above mentioned window of -200ms to +400ms around the R-peaks. The choice of the mother wavelet and the decomposition level was based on [59] work. So, firstly, the Daubechies wavelet of order 2 (Db2) was chosen due to it’s similar morphological structure with the ECG signals. This wavelet can be seen in Figure 4.7.

![Figure 4.7: Mother wavelet (left) and scaling function (right) of the Db2 wavelet](image)

Also, a level of decomposition of 4 was chosen, which means that the heartbeat signal was decomposed into the detail coefficients D1-D4 and one approximation coefficient A4.

The wavelet coefficients represent the distribution of the signals energy in time and frequency. From these coefficients, the following statistical values were extracted and used as feature:

- Maximum value of each of the wavelet coefficients;
- Mean value of each of the wavelet coefficients;
- Minimum value of each of the wavelet coefficients;
- Standard deviation of each of the wavelet coefficients.

Therefore, 20 morphological wavelet based features were extracted from the ECG heartbeat.

Then, the validation part of the experiments, done in the StM database, quadratic spline level 6 mother wavelet was also tested to see if it improved the classification results.

### 4.4.3 Evaluation and performance measures

The experiments specifications made for the classification task on both databases are specified in Table 4.4. A stratified 5-fold cross-validation was used to test the implemented algorithms for both databases,
where for each one of the five possible combinations, 4 folds were used as training data and one was used as test data. Due to the unbalancing of the number of label samples for the databases seen in table 4.3, a balancing step was performed before running the algorithm, using as criteria for the number of samples from each label the minimum common denominator, i.e., the value that matched the label with the smallest number of samples. In the case of the StM database, since the F label only had 33 samples, the second smallest value was chosen for the other labels, keeping the 33 F samples. This is going to be further discussed in the Results section.

Table 4.4: Experiments specification for Classification Task

<table>
<thead>
<tr>
<th>Experiment</th>
<th>DATABASE</th>
<th>Labels used</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MIT-BIH</td>
<td>N, S, V, F</td>
<td>This first experiment was made to evaluate the algorithm and the 3 different combinations of features for the detection of the location where a certain heartbeat is originated. It’s based on the AAMI recommended practice, taking out the “Q” type, for the reasons mentioned above.</td>
</tr>
<tr>
<td>2</td>
<td>MIT-BIH</td>
<td>N, L, S, V, F</td>
<td>In the second experiment we distinguish the L type from the N type mentioned above, for it is also an important beat to take into account. This label set, as mentioned above, is the one that we aim to optimize and validate after, on the StM Database.</td>
</tr>
<tr>
<td>3</td>
<td>MIT-BIH</td>
<td>L, V</td>
<td>In order to verify if the L and V were being properly identified, since they have a similar morphology, this extra test was made.</td>
</tr>
<tr>
<td>4</td>
<td>StM</td>
<td>N, L, S, V, F</td>
<td>Using the combination of both time and morphological features, this experiment was done to validate the results from point 2 in a different database.</td>
</tr>
<tr>
<td>5</td>
<td>StM</td>
<td>N, L, S, V, F</td>
<td>Experiment 4 is then repeated but using a different mother wavelet, namely the quadratic spline wavelet, with a level of decomposition = 6.</td>
</tr>
<tr>
<td>6</td>
<td>StM</td>
<td>N, L, S, V, F</td>
<td>From experiment 4 and 5, the best set of features was chosen and retested after applying the best denoising method found from the denoising experiments, in order to see the effect of denoising on heartbeat classification.</td>
</tr>
<tr>
<td>7</td>
<td>StM</td>
<td>N, L, S, V, F</td>
<td>Same setup was used but the KNN algorithm was tested instead of the SVM.</td>
</tr>
</tbody>
</table>

The algorithms and feature sets were tested using the MIT-BIH database, which served as benchmark data, and were validated on the StM one. To evaluate the performance of the classifier, accuracy, precision, recall and $F_1$-score metrics were used. Using the typical notation for true positives ($T_P$), true negatives ($T_N$), false positives ($F_P$) and false negatives ($F_N$), these metrics can be defined by Equations 4.7, 4.8, 4.9 and 4.10, where the $F_1$-score is the harmonic mean of precision and recall.

\[
\text{Accuracy} = \frac{T_P + T_N}{N} \tag{4.7}
\]

\[
\text{Precision} = \frac{T_P}{T_P + F_P} \tag{4.8}
\]

\[
\text{Recall} = \frac{T_P}{T_P + F_N} \tag{4.9}
\]

\[
F_1 = \frac{2T_P}{2T_P + F_N + F_P} \tag{4.10}
\]
Chapter 5

Experimental Setup and Results

In this section, the experimental setup used for the thesis is presented and the results are outlined and discussed, for both the denoising and the classification tasks.

5.1 Experimental Setup

The proposed methodology was validated using three different databases that are described next.

5.1.1 BITalino Bicycle (BB) Database

This database was gathered for the study of the influence of physical effort on a bicycle with the integration of a BITalino device [60] in the ECG signal morphology. It had 2 goals:

1. Evaluate the performance of the system and algorithms used for biometric recognition when there are changes in the heart rate (HR) and pattern of the ECG signal;

2. Evaluate the use of heart rate variability (HRV) as a biometric trait.

It collected a number of different biological signals, from ECG signals, to blood volume pulse data and respiration data, among others. For this work, this database was used to test the performance of different denoising methods on an ECG signal. So, only ECG signals were used. Two types of ECG data were used and are described below:

- ECG signals collected on the chest, using a chest band (Figure A.1), with electrodes placed on each side of the sternum and the fifth or sixth intercostal spaces. This was used as the baseline clean signal.

- ECG signals acquired at the hands, using a stationary bicycle with a BITalino device coupled to it (Figure 5.2). Due to amount of movement, sweat, among other things, this signal is highly corrupted by noise and was the signal that was aimed to be denoised.

For both of the used recordings, a sampling rate of 1000 Hz was used. The protocol used for these recordings is specified in [61]. Twenty five subjects were recorded and studied in the present work.
5.1.2 MIT-BIH Arrhythmia (MIT-BIH) Database

The MIT-BIH Arrhythmia database was built from a set of over 4000 long-term Holter ECG recordings, gathered by the Beth Israel Hospital Arrhythmia Laboratory between 1975 and 1979. It contains 48 recordings of slightly over 30 minutes long, where 23 of them are randomly chosen from the set and 25 are selected from the same set to include a variety of rare but clinically important cases that would not be well represented by a small random sample of the Holter recordings [62]. The first 23 records serve as a sample of the variety of waveforms and artifact that an arrhythmia detector might encounter in a routine clinical use. The last 25 records were selected to include complex ventricular, junctional, and supraventricular arrhythmia and conduction abnormalities.

The analog outputs were filtered to limit analog-to-digital converter (ADC) saturation and for anti-aliasing, using a pass-band filter from 0.1 to 100 Hz. These filtered signals were then digitized at a 360 Hz sampling frequency.

More specific characteristics, recording and annotation details can be found in [63].

This database is used in the classification experiments since it is one of the most used databases for this type of tasks and serves as a base for comparison of the algorithms implemented in this thesis and
also to compare to the results from the Santa Marta database, described below.

5.1.3 Santa Marta (StM) Database

The ECG signal acquisition was done by the clinicians of the Santa Marta Hospital in Lisbon, using a Philips Page Writer Trim III recorder, with a sampling frequency of 500 Hz.

This database was used for designing prototypes for the outlier removal task that’s going to be described bellow and for validating the feature set chosen for this thesis, from the benchmark data (MIT-BIH database). In total, 11429 recordings were used for these experiments, a much higher amount than the number of recordings for the benchmark data.

5.2 Results for Denoising Techniques

The denoising experiments were done using the BB database. Due to problems during the recording of the signals, some subjects beginning of the signals were missing, either in the noisy or in the clean version of the signal, which compromised the beat detection because of the fact that the algorithm that was used, based on the work of [50], makes an initial estimate for the detection by detecting maximum peaks in eight consecutive 1-second intervals. Because of this, 3 subjects were taken off the denoising experiment. Data from these subjects, namely subjects 7, 18 and 19, are displayed in Figure 5.3.

![Figure 5.3: ECG recordings from subjects that were removed from the denoising experiment.](image)

The first step was to get an estimate of the mean, mode and median values for CosSim and SNR measures for each subject without the application of any noise reduction technique, that is, using only the 2 median filters for the baseline wander removal. Examples of these first results are presented below, in Table 5.1.
Table 5.1: SNR and CosSim values for the four subjects with the best results (highest values) after only doing a baseline wander removal using 2 median filters.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>2 Median Filters (Baseline wander removal)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUBJECT</td>
</tr>
<tr>
<td>SNR</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>CosSim</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>20</td>
</tr>
</tbody>
</table>

5.2.1 40 Hz Low-Pass Filter Results

This baseline filtering method, described in section 4.2.1 and was applied both to clean signal and to the noisy one. It was used in order to remove the effects of high frequencies on the signals. The four best results for an initial calculation of CosSim and SNR for the signals after applying this technique (for 4 different subjects - subjects 1, 11, 16 and 20) are presented in Table 5.2, where the cosine similarity and SNR means, mode and medians are presented for the each of the subjects after the 40 Hz low pass filter (LP) was applied to both signals. These 4 subjects are going to serve as reference for the results presented bellow for the other techniques.

Table 5.2: SNR and CosSim values for the four subjects with the best results (highest values) before and after 40 Hz LP filter.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>2 Median Filters (Baseline wander removal)</th>
<th>2 Median Filters + 40 Hz LP Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SUBJECT</td>
<td>MEAN</td>
</tr>
<tr>
<td>SNR</td>
<td>1</td>
<td>5,917</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>0,630</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3,652</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>2,997</td>
</tr>
<tr>
<td>CosSim</td>
<td>1</td>
<td>0,810</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>0,671</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>0,785</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0,687</td>
</tr>
</tbody>
</table>

As it can be seen for these 4 subjects, just using the 40 Hz Low pass filter improves both the mean, mode and median values for both the CosSim and SNR metrics.

5.2.2 Wavelet Based Denoising Techniques

Wavelet based denoising was performed to the signals, to try and improve the values obtained for the baseline filtering. Summarizing the process already outlined in the previous section, each heartbeat of the noisy signal, after filtering the whole signal with the 2 Median filters + 40 Hz LP filter, went through a wavelet decomposition process, where a threshold was applied to each of the the extracted coefficients and then reconstructed from these new coefficients. Then CosSim and SNR was calculated for each of
the beats and an mean, mode and median was calculated for each subject's set of heartbeats. In this section, the results for each of the used wavelets are presented.

**Results of Sym12 with decomposition level 3**

Firstly, the Sym12 wavelet with decomposition level 3 was applied directly to each detected heartbeat from the noisy signal without it undergoing a 40 low-pass filter step first. The SNR ans CosSim were calculated comparing these heartbeats to the same heartbeats in the clean signal removing its high frequency noise using the low-pass filtering technique. The results presented in Table 5.3 for the 4 subjects previously chosen. As it can be seen, for the four subjects, the values for the CosSim and SNR metrics decreased in comparison to the low-pass filtered ones.

Table 5.3: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Sym12, is applied to the heartbeats without applying a 40 Hz low pass filtering to the noisy signal first.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>MEAN (40 Hz LP Filter)</th>
<th>MODE (40 Hz LP Filter)</th>
<th>MEDIAN (40 Hz LP Filter)</th>
<th>MEAN (Sym12 Level 3)</th>
<th>MODE (Sym12 Level 3)</th>
<th>MEDIAN (Sym12 Level 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
<td>6,441</td>
<td>9,000</td>
<td>6,625</td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
<td>2,782</td>
<td>3,000</td>
<td>3,029</td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
<td>3,461</td>
<td>6,000</td>
<td>4,079</td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
<td>3,860</td>
<td>4,000</td>
<td>4,121</td>
</tr>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
<td>0,826</td>
<td>0,960</td>
<td>0,892</td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
<td>0,723</td>
<td>0,860</td>
<td>0,809</td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
<td>0,767</td>
<td>0,930</td>
<td>0,868</td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
<td>0,720</td>
<td>0,860</td>
<td>0,790</td>
</tr>
</tbody>
</table>

Then, the same wavelet was tried after applying the 40 low-pass filter also to the noisy signal. The results for the 4 subjects are presented in Table 5.4.

Table 5.4: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Sym12, is applied to the heartbeats after first applying a 40 Hz low pass filtering to the noisy signal.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>MEAN (40 Hz LP Filter)</th>
<th>MODE (40 Hz LP Filter)</th>
<th>MEDIAN (40 Hz LP Filter)</th>
<th>MEAN (Sym12 Level 3)</th>
<th>MODE (Sym12 Level 3)</th>
<th>MEDIAN (Sym12 Level 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
<td>6,905</td>
<td>8,000</td>
<td>7,055</td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
<td>3,607</td>
<td>5,000</td>
<td>3,944</td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
<td>4,592</td>
<td>7,000</td>
<td>5,477</td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
</tr>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
</tr>
</tbody>
</table>

As it can be seen, for the four subjects, no significant change existed in the values for the CosSim and SNR metrics in relation to the low pass filtered ones. Actually, some of the values even decreased slightly, which is undesirable.
Results of Db4 with level 5

The results for the denoising experiments using the Db4 mother wavelet with decomposition level 5 without the noisy signal undergoing a 40 Hz low-pass filtering step are shown in Table 5.5, for the 4 subjects referenced for the previous experiments. As it can be seen, using this mother wavelet lowers the quality of the signal significantly. Both the mean CosSim and SNR values decrease for the 4 studied subjects but the standard deviation for the CosSim also decrease for some subjects, which would be an important and decisive outcome if the decrease of the CosSim and SNR values wasn’t so significant.

Table 5.5: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Db4, is applied to the heartbeats without 40 Hz low pass filtering to the noisy signal.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>2 Median Filters</th>
<th>40 Hz LP Filter</th>
<th>DB4 Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>MODE</td>
<td>MEDIAN</td>
</tr>
<tr>
<td>SNR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
</tr>
<tr>
<td>CosSim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
</tr>
</tbody>
</table>

After applying the Db4 wavelet after the low-pass filter is used on the noisy signal as well, the results are either slightly worst or almost the same as without the filtering step, as it can be seen in table 5.6.

Table 5.6: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Db4, is applied to the heartbeats after the 40 Hz low pass filtering to the noisy signal.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>2 Median Filters</th>
<th>40 Hz LP Filter</th>
<th>DB4 Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEAN</td>
<td>MODE</td>
<td>MEDIAN</td>
</tr>
<tr>
<td>SNR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
</tr>
<tr>
<td>CosSim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
</tr>
</tbody>
</table>

Results of Quadratic Spline with level 6

The quadratic spline, with decomposition level 6, mother wavelet without the low-pass filtering step showed worst results then when using only the low pass filter for the noisy signal. These results can be seen in table 5.7.

Still, combining this wavelet with the 40 Hz low pass filter showed promising results in relation to the other wavelets for the denoising task, as it can be seen in Table 5.8. For the 4 subjects chosen above there is a small increase on both the CosSim and the SNR mean, mode and median values.
Table 5.7: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Quadratic Spline, is applied to the heartbeats without 40 Hz low pass filtering to the noisy signal.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SNR</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>SNR</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 Median Filters</td>
<td>40 Hz LP Filter</td>
<td>Quadratic Spline</td>
<td>Level 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
<td>6,158</td>
<td>7,000</td>
<td>6,554</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
<td>1,930</td>
<td>3,000</td>
<td>2,074</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
<td>3,180</td>
<td>6,000</td>
<td>3,927</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
<td>3,332</td>
<td>4,000</td>
<td>3,598</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>CosSim</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>CosSim</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
<td>0,819</td>
<td>0,950</td>
<td>0,886</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
<td>0,684</td>
<td>0,820</td>
<td>0,759</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
<td>0,759</td>
<td>0,910</td>
<td>0,858</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
<td>0,698</td>
<td>0,860</td>
<td>0,762</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8: SNR and CosSim for the previously selected subjects before and after wavelet denoising technique, Quadratic Spline, is applied to the heartbeats after the 40 Hz low pass filtering to the noisy signal.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>SNR</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>SNR</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 Median Filters</td>
<td>40 Hz LP Filter</td>
<td>Quadratic Spline</td>
<td>Level 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
<td>6,951</td>
<td>8,000</td>
<td>7,179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
<td>3,616</td>
<td>5,000</td>
<td>3,960</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
<td>4,568</td>
<td>7,000</td>
<td>5,301</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
<td>4,066</td>
<td>5,000</td>
<td>4,229</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>CosSim</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
<th>CosSim</th>
<th>Mean</th>
<th>Mode</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
<td>0,835</td>
<td>0,970</td>
<td>0,903</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
<td>0,752</td>
<td>0,900</td>
<td>0,850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
<td>0,794</td>
<td>0,930</td>
<td>0,894</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
<td>0,721</td>
<td>0,910</td>
<td>0,797</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.2.3 Moving Average

For this work, a moving average filter was also used, in order to evaluate its possible use in the denoising of ECG signals. The moving average filter was used on the whole signal for each subject and the procedure was described above, in the methods section. The results obtained for the 4 subjects, using a 28 ms window size, are presented in 5.9.

5.2.4 Outlier Removal

The method used for the removal of outliers, derived from artifacts or beats other than the ones that are being studied in this thesis, was based on prototypes for the heartbeat classes of interest for this thesis, namely the N, L, S, V and F classes described in the sections above. These prototypes were developed using technique described in the methods section, using the StM Database heartbeats as models for their development.

In Figures 5.4 and 5.5 the prototypes for each of the 5 beats (green) and the beat that is closest (in terms of euclidean distance) to each centroid (red) is shown. These prototypes were built using the sampling frequency of the StM database, i.e., 500 Hz. In order to be able to use it in the other databases, an interpolation step had to be done. In the case of the denoising experiments, the sampling frequency of the signals from the BB database are, as mentioned before, at a 1000 Hz sampling frequency, so each
Table 5.9: SNR and CosSim values for the four subjects with the best results (highest values) before and after moving average is applied to the signals.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>MEAN</th>
<th>MODE</th>
<th>MEDIAN</th>
<th>MEAN</th>
<th>MODE</th>
<th>MEDIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>6,904</td>
<td>8,000</td>
<td>7,055</td>
<td>7,375</td>
<td>2,000</td>
<td>4,303</td>
</tr>
<tr>
<td>11</td>
<td>3,604</td>
<td>5,000</td>
<td>3,941</td>
<td>3,259</td>
<td>1,000</td>
<td>2,983</td>
</tr>
<tr>
<td>16</td>
<td>4,593</td>
<td>7,000</td>
<td>5,482</td>
<td>3,301</td>
<td>3,000</td>
<td>3,061</td>
</tr>
<tr>
<td>20</td>
<td>4,010</td>
<td>5,000</td>
<td>4,196</td>
<td>3,121</td>
<td>2,000</td>
<td>2,512</td>
</tr>
<tr>
<td>CosSim</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0,834</td>
<td>0,970</td>
<td>0,900</td>
<td>0,804</td>
<td>0,970</td>
<td>0,881</td>
</tr>
<tr>
<td>11</td>
<td>0,752</td>
<td>0,890</td>
<td>0,848</td>
<td>0,755</td>
<td>0,910</td>
<td>0,862</td>
</tr>
<tr>
<td>16</td>
<td>0,794</td>
<td>0,930</td>
<td>0,895</td>
<td>0,826</td>
<td>0,940</td>
<td>0,898</td>
</tr>
<tr>
<td>20</td>
<td>0,717</td>
<td>0,880</td>
<td>0,794</td>
<td>0,718</td>
<td>0,910</td>
<td>0,781</td>
</tr>
</tbody>
</table>

A validation step was performed next, using the MIT-BIH database. This is the standard database used for most researches made in this field, so, an outlier removal as described in the methods section was applied to this database, to see if most heartbeats were considered non-outliers, using the 0.8 cosine similarity criteria described before.

The results for this validation step are presented in Table 5.10 and after the interpolation and re-sampling at a 360 Hz sampling frequency, the same as the validation database. As it can be seen, from the 48 total subjects, only 16 of them showed a difference of more than 100 heartbeats from the initial number of heartbeats to their number after the outlier removal. In the table it can also be seen the maximum and minimum differences in the number of heartbeats before and after the outlier removal algorithm. This means that at least one subject lost 2776 beats but there were also subjects that kept all of their heartbeats after the outlier removal, since the minimum difference is 0. In the Appendix A, the list of heartbeats before and after outlier removal for all subjects from the MIT-BIH Database is detailed. All the subjects with a difference higher than 100 are represented with the red highlight.

The higher differences can have multiple causes, like:

- The MIT-BIH database is used for diagnosing more than the 4 classes represented in the prototypes, so this technique will remove all of those heartbeats;

- Each subjects is different and might have different heartbeat characteristics that were not represented in the Sta Marta database;

- The re-sampling can affect the prototypes and lead to some misrepresentation of the heartbeats.

Still, these results were quite promising, so this outlier removal technique was applied also to the BB database to try to improve the quality of the signals.

The outlier removal was then applied and combined to the best denoising method from the ones studied above (Quadratic Spline, level 6, wavelet denoising), to the BB database signals, in order to measure the improvement on the denoising task if outlier removal is used. These results, for the 4
Figure 5.4: Set 1 (of two) of Final Prototypes for Outlier Removal. Green heartbeats represents the centroid and the red ones represent the heartbeat with the least euclidean distance to the centroid.

subjects used for all the results showed above, are shown in Table 5.11. On this table it’s possible to see that the outlier removal is an important part of the denoising task. In this case, a significant number of possible heartbeats are lost, which will affect the number of samples for the classification task but, in terms of clean heartbeats, it may help the improvement of this task and get better results, due to the better readability of the signals.

5.2.5 Denoising Task Discussion

From the above results, it can be seen that, before the outlier removal, the denoising technique that presented the best results in terms of mean, mode and median for both the CosSim and SNR metrics was the Quadratic Spline wavelet based method, after applying a 40 Hz low pass filter to both the noisy and the clean signals. The moving average also had very promising results, especially in terms of the
mean of CosSim. To further show these results, Figures 5.6 and 5.7 present more subjects used in these experiments, that presented a mean value for the CosSim higher than 0.5 for only the 40 low pass filter method. Some subjects have very low SNR and CosSim values, probably because to recording conditions. During the bike recording, some subjects moved too much, sweated more than others, among other things, which can justify these low values. Even so, after the LP filter, these worst signals also presented a significant increase in the values of CosSim and SNR, which shows the importance of using these filtering technique on ECG signals. These subjects are not represented in the figures.

In terms of balance between the CosSim values and the SNR values, the algorithm that performed better than the rest was the quadratic spline wavelet based one. The moving average performed really well in terms of cosine similarity but was highly outperformed in terms of SNR by all the other techniques. For the purpose of this work, for all the following denoising needs, the quadratic spline wavelet based method was adopted.

Also an outlier removal method was applied to the signals, combined with this wavelet based method, to measure the improvement to the signal denoising results. Figures 5.8 and 5.9 show this improvement also for all subjects with a CosSim higher then 0.5 for only the 40 low pass filter method.

As it can be seen in the figures, practically most subjects present an improvement on both the
Figure 5.6: CosSim mean, mode and median values for subjects whose CosSim mean was higher than 0.5.
Figure 5.7: SNR mean, mode and median values for subjects whose CosSim mean was higher than 0.5.
Figure 5.8: CosSim mean, mode and median values of quadratic spline wavelet based technique with and without outlier removal for subjects whose CosSim mean was higher than 0.5.
Figure 5.9: SNR mean, mode and median values of quadratic spline wavelet based technique with and without outlier removal for subjects whose CosSim mean was higher than 0.5.
Table 5.10: Results for validation of outlier removal method using the MIT-BIH database. This shows the number of subjects with a difference in the initial and final number of heartbeats higher than 100 and also the maximum and minimum values of the difference from the initial and final number of heartbeats, i.e., before and after outlier removal. More detailed results are present in Appendix A.

<table>
<thead>
<tr>
<th>Total number of Subjects</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of subjects with &gt;100 difference</td>
<td>16</td>
</tr>
<tr>
<td>Maximum difference (n heartbeats)</td>
<td>2776</td>
</tr>
<tr>
<td>Minimum difference (n heartbeats)</td>
<td>0</td>
</tr>
</tbody>
</table>

CosSim and SNR metrics. This shows the importance of an outlier removal step and that the developed prototype method can be a promising method to further developed.

5.3 Classification Results

The classification task was done firstly in the benchmark data, namely the MIT-BIH database, being validated after in the StM Database, where a different set of features was also tested. The experiments made are described in 4.4.3 and their results are analyzed in this section.

5.3.1 MIT-BIH - Benchmark Data Experimental Results

Three different experiments were done using the MIT-BIH database, which all started with the same important step: balancing of the feature set. As it could be seen above in Table 4.3, the dataset is highly unbalanced in terms of the number of different types of classes. To ensure that this unbalancing is not wrongly influencing the results, a balancing step was done, where 783 samples from each label (the number of samples of the class with less samples) were chosen randomly, using always the same 783 for all 3 experiments. For this part of the experiment only SVM was used as a classification algorithm. The experiments for this dataset are described below. The $C$ and $\gamma$ parameters for the several feature sets were chosen based on the 4 metrics, the chosen parameters were the ones that presented the best values for the 4 metrics. If some experiment had 2 sets of parameters that had 2 metrics better for one and the other 2 for another, the one that had the highest accuracy value was chosen. This criteria for choosing the parameters was used for all the classification experiments.

Experiment 1 - AAMI Recommended Practices Based Experiment

This first experiment was done to test the algorithm and the classes used for most multiclass heartbeat detection tasks, taking into account the AAMI best practises, less the "Q" class for the reasons mentioned above. This experiment also served to test the different sets of features to use (time, morphological - wavelet Db2, level 4 - or a combination of both feature sets). The results are shown in table 5.12.

As it can be seen, time features only give results above 83% for all performance evaluation indices. This is explained by the fact that most of the pathological heartbeats that are being studied are prema-
ture, which highly influences the types of temporal features that are being used, specially the pre-RR and the pos-RR, that ultimately end up also affecting the average and the local average. Using only the wavelet based features achieved better results then only the time features (>84%).

This result proves the potential that wavelet features have for this type of classification task, specially due to the morphological difference between the different classes. As it was expected and already proven in several previous studies, the combination of morphological and temporal features achieved the best results (>87%). These results show that taking into account both the morphological features of a heartbeat and the distance between heartbeats, good performances can be achieved in automatic classification of heartbeats.

**Experiment 2 - 5 Proposed Labels (N, L, S, V and F)**

In this experiment, the benchmark for the main classification task was set. This consisted in testing the SVM algorithm for the 5 labels (N, L, S, V and F), after the balancing step takes place. These results are presented in Table 5.13.

In this task, the results for the time features slightly decreased. This can be due to the fact that both N and L beats are generated in the SA node, which makes their timings quite similar and it can lead to a wrong classification by the algorithm. Using only wavelet based features achieved a better result than before (increase of about 2%).

One of the problems with also using the L class was that its morphology is quite similar to the V type heartbeats, which was expected to lead to a decrease in accuracy. The fact that the decrease didn’t happen can also be due to the separating of the L and N beats. Since the N class had one less sub-type included, the N class turned less variant in terms of morphology, which improved its classification. As expected, the mixture of both types of features also increased the values of the metrics used to measure the results.

**Experiment 3 - L vs. V Classification**

L and V heartbeats have similar shapes and it is important to test the algorithm’s performance when distinguish this two labels. For this, a third experiment was made, using only L and V labels. The balanced set used for this experiment was the same as the previous experiments, with 783 samples for each of the two classes. The same 3 types of features where tested and the results are showed in Table 5.14. As it can be seen, there’s not much differences between the 3 types of features but still, the combination of time and wavelet features shows the best results. These results are important to understand the influence of separating the L class from the N one in the classification results. The results for all the metrics and also for all feature set are above 98%, which means that the algorithm performs very well in separating L from V heartbeats, despite their similar shape.
Table 5.11: SNR and CosSim values for the four subjects with the best results (highest values) before and after Outlier Removal + Quadratic Spline, level 6, is applied to the signals.

<table>
<thead>
<tr>
<th>SUBJECT</th>
<th>2 Median Filters</th>
<th>Outlier Removal + Quadratic Spline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SNR</td>
<td>CosSim</td>
</tr>
<tr>
<td></td>
<td>MEAN</td>
<td>MODE</td>
</tr>
<tr>
<td>1</td>
<td>6.904</td>
<td>8.000</td>
</tr>
<tr>
<td>11</td>
<td>3.604</td>
<td>5.000</td>
</tr>
<tr>
<td>16</td>
<td>4.593</td>
<td>7.000</td>
</tr>
<tr>
<td>20</td>
<td>4.010</td>
<td>5.000</td>
</tr>
<tr>
<td></td>
<td>0.834</td>
<td>0.970</td>
</tr>
<tr>
<td>11</td>
<td>0.752</td>
<td>0.890</td>
</tr>
<tr>
<td>16</td>
<td>0.794</td>
<td>0.930</td>
</tr>
<tr>
<td>20</td>
<td>0.717</td>
<td>0.880</td>
</tr>
</tbody>
</table>

Table 5.12: Results for the AAMI recommended practices based classes, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>$F_1$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Features (A)</td>
<td>1/0.001</td>
<td>83.38%</td>
<td>84.09%</td>
<td>83.34%</td>
<td>83.71%</td>
</tr>
<tr>
<td>Wavelet Based Features (B)</td>
<td>10/1</td>
<td>84.19%</td>
<td>84.60%</td>
<td>84.16%</td>
<td>84.38%</td>
</tr>
<tr>
<td>Combination A+B</td>
<td>100000/0.00001</td>
<td>87.06%</td>
<td>87.47%</td>
<td>87.10%</td>
<td>87.28%</td>
</tr>
</tbody>
</table>

Table 5.13: Results for proposed classes for this thesis, namely N, L, S, V and F, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>$F_1$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Features (A)</td>
<td>100/0.001</td>
<td>82.16%</td>
<td>82.36%</td>
<td>82.24%</td>
<td>82.30%</td>
</tr>
<tr>
<td>Wavelet Based Features (B)</td>
<td>10/1</td>
<td>86.31%</td>
<td>86.44%</td>
<td>86.34%</td>
<td>86.39%</td>
</tr>
<tr>
<td>Combination A+B</td>
<td>100000/0.00001</td>
<td>88.70%</td>
<td>89.10%</td>
<td>88.72%</td>
<td>88.91%</td>
</tr>
</tbody>
</table>

Table 5.14: Results for classification of L and V heartbeats, where results for the three types of feature sets are presented, as well as the SVM $C$ and $\gamma$ parameters that achieved the best results.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>$F_1$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time Features (A)</td>
<td>100/0.0001</td>
<td>98.13%</td>
<td>98.13%</td>
<td>98.12%</td>
<td>98.13%</td>
</tr>
<tr>
<td>Wavelet Based Features (B)</td>
<td>10/1</td>
<td>98.38%</td>
<td>98.38%</td>
<td>98.38%</td>
<td>98.38%</td>
</tr>
<tr>
<td>Combination A+B</td>
<td>100000/0.00001</td>
<td>98.64%</td>
<td>98.66%</td>
<td>98.63%</td>
<td>98.65%</td>
</tr>
</tbody>
</table>
5.3.2 StM - Validation Setup Experimental Results

To verify the results obtained by using the MIT-BIH database, a database that has been optimized and highly used and studied, the algorithm was tested on the StM database, developed by the IT group, that had 11429 recordings.

The first step was to use the best set of labels from experiment 2, the main aim of the classification task. Even though the number of samples for each label isn’t that big between this dataset and the previous one, the number of different recordings is much higher, which represents a bigger variability in the data both from subject, placement or situation. This variability can result in worst values for the metrics used in this thesis and to the classifier working on recognizing, for example, the subject for the heartbeat in question instead of the type of heartbeat, independently of external conditions. Secondly, a different mother wavelet was tested, namely the Quadratic Spline wavelet with decomposition level 6. After that, two more experiments were done, one using the outlier removal plus best denoising method from the ones studied in the previous experiments and another trying a different classifier, KNN, on the experimental setup that showed the best results for this dataset. In all the experiments a balancing step was performed for the dataset, where the second smallest value for the number of samples was used, due to the fact that the smallest number of samples was for class "F", that had only 33 samples. This is a number too small to be able to train and test an algorithm. So during these classification tasks, there were only 33 "F" samples while the rest of the classes had 1028 samples (or 31 "F" samples and 681 for the rest of the classes in the outlier removal experiment). For this reason, one last test was made, using the method from experiment 1, the one that had the feature set that gave the best results, in other to see the effect of this unbalancing on the experiments results, using both the SVM and the KNN classifiers. All the results for these experiments are outlined and discussed bellow.

Experiment 4 - Classification using SVM Classifier

After the balancing step, 1028 random samples from classes "N", "L", "S" and "V" and the 33 samples for the "F" class were used to train and test the algorithm using a 5 fold cross validation using the combination of time and wavelet based features, the feature set that had the best results for Experiment 2 using the MIT-BIH database. The results for this experiment are showed in Table 5.15.

As it can be seen, the accuracy values still were above 80% for this setup but decreased in relation to the results obtained in experiment 2. Also, while for the MIT-BIH database the precision, recall and $F_1$-score values are very close to the accuracy values, being also above 80%, using the StM Database these 3 metrics decrease significantly. These results can be due to the 33 samples for the "F" label, that may be leading to a large number of false positives for the other labels. To confirm that this is affecting the results, a test was made using only the samples for the "N", "L", "S" and "V" classes and the same algorithm. The results are showed in Figure 5.10.

Even though the improvement in the accuracy values aren’t highly significant, there is a high increase in the other metrics, which shows the importance of the balancing of the dataset, especially if one of the labels has a smaller number of samples as in this case. Without the F-class which has only 33 samples,
the number of false positives decreases and the number of true positives increases, leading to these higher precision, recall and $F_1$ values.

The fact that the values obtained for this dataset are lower than the ones for the MIT-BIH one can be explained by the much higher number of recordings used, which has the advantage of having a higher number of data for training and testing the algorithms but has the disadvantage of having a bigger variability among the recordings as explained above.

Table 5.15: Results for classification of the 5 labels using the Db2 wavelet with a decomposition level 4.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Db2, level=4</td>
<td>10/0.00001</td>
<td>81.06%</td>
<td>74.23%</td>
<td>67.04%</td>
<td>70.45%</td>
</tr>
</tbody>
</table>

Figure 5.10: Comparison between the results for the classification of the 5 different classes (green) and taking the “F” class beats from the dataset (blue).

Experiment 5 - Classification Quadratic Spline Wavelet Based features

Combining the same time features as the previous experience with morphological features extracted using the quadratic spline wavelet with a decomposition level 6 (the same statistical features from the now 7 coefficients, which led to a feature set of 32 features instead of 24), the 5 fold cross validation using the SVM classifier was used and the results are showed in Table 5.18.

The results using this wavelet are much smaller then before, being bellow 70% for all metrics. As in the previous one, the precision, recall and $F_1$ values are much smaller than the accuracy, since the experiment was done for the classification of the 5 labels (including “F”) instead of just the 4.

Taking these results into account, the Db2 wavelet was chosen for the following experiments, since it was the one that had the best results.
Table 5.16: Results for classification of the 5 labels using the Quadratic Spline wavelet with a decomposition level 6.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>( F_1 ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quadratic Spline, level = 6</td>
<td>10/0.0001</td>
<td>65.10%</td>
<td>54.80%</td>
<td>53.74%</td>
<td>54.27%</td>
</tr>
</tbody>
</table>

Experiment 6 - Classification using Outlier Removal and Denoising Step

In the denoising task of this thesis, the method that achieved the best results was the one that used the outlier removal technique plus the quadratic spline wavelet based denoising technique. To investigate the effects of these techniques on the classification task, the same features were extracted from the dataset after a denoising and outlier removal step. After performing the outlier removal on the StM dataset, the number of samples in each class decreased. The final number of each class is presented in Table 5.17.

Table 5.17: Number of samples in each class after outlier removal step.

<table>
<thead>
<tr>
<th>Database\Labels</th>
<th>N</th>
<th>L</th>
<th>S</th>
<th>V</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>StM Dataset</td>
<td>92597</td>
<td>8277</td>
<td>1259</td>
<td>681</td>
<td>31</td>
</tr>
<tr>
<td>After Outlier Removal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Since the second smaller number of samples in a class now is 681, the balancing step was done so that each class had this number of random samples, from the dataset after outlier removal, except for the “F” class which had 31 samples, the only available in the dataset. After this balancing step, the wavelet based denoising method was applied to each heartbeat before extracting the wavelet based features. The results for this experiment are shown in Table 5.18.

The main purpose of this experiment was to see if the algorithm performed better after denoising and outlier removal than without these steps. As it can be seen, the accuracy value is lower then in experiment 4 (the same experiment without denoising and outlier removal) but the other values are slightly better. This can be explained by the decrease in the number of samples from one experiment to another. Since there is less data to train the algorithm, the accuracy is affected but since the difference between the number of samples in the “F” class and the other classes is smaller, the precision, recall and \( F_1 \) values are higher. To better see the results of using the denoising methods, experiment 4 was redone but this time using a balanced dataset that had the same number of samples for each class as the ones for the current experiment, i.e., 681 random samples from the dataset with no outlier removal step for the “N”, “L”, “S”, and “V” classes and 31 for the “F” one. The results are presented in Figure 5.11.

From this figure it can be seen that denoising and outlier removal are important for the classification task. The outlier removal step removes from the dataset all the heartbeats that shown a cosine similarity smaller than 0.8 to the prototypes defined for each type of class. Then, the signals undergo a further denoising step, which increases the quality of the signal. These two combined can explain the improve-
ment of the metrics in relation to the newly balanced dataset that doesn’t have these extra steps. The fact that the 681 samples are being randomly chosen from a dataset bigger than the one after the outlier removal can also explain the decrease of the results from experiment 4 to these ones.

Table 5.18: Results for classification of the 5 labels after an outlier removal and denoising steps.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>C/gamma Parameters</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Db2, level = 4 + Outlier Removal + Denoising</td>
<td>10/0.00001</td>
<td>80.56%</td>
<td>75.21%</td>
<td>67.39%</td>
<td>71.09%</td>
</tr>
</tbody>
</table>

Figure 5.11: Comparison between the results for the classification with 681 samples per class balancing (and 31 for the "F" class) (green) and for classification after outlier removal plus quadratic spline wavelet based denoising step.

Experiment 7 - Classification using KNN Classifier

To also test the performance using a different classifier, this last experiment was made using a KNN algorithm. This experiment used the same setup as Experiment 4, the only difference being the classification algorithm used. The results are shown in Table 5.19 and show worst results (Accuracy = 77.83%) than the ones that used the SVM classifier (Accuracy = 81.06%). The precision, recall and F1 values are again lower than the accuracy ones, probably because of the same reason as before, that is, the under sampling of the "F" class. In Figure 5.12, the results using only 4 classes, without the "F" are showed and compared to the ones with all 5 classes. From this figure, it can be seen that taking out the "F" class performs better than using it, giving higher precision, recall and F1 values, probably due to the decrease in false positives and increase of true positives.

From the experiments’ results, it can be seen that the algorithm that performed better was the SVM one, for both the SIM and the MIT-BIH databases. Also, the testing of both wavelet features showed...
Table 5.19: Results for classification of the 5 labels using the Db2 wavelet with a decomposition level 4 and the KNN classifier.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>K</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall(%)</th>
<th>$F_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB2, level = 4</td>
<td>1</td>
<td>77.83%</td>
<td>64.40%</td>
<td>64.42%</td>
<td>64.41%</td>
</tr>
</tbody>
</table>

Figure 5.12: Comparison between the results for the classification of the 5 different classes (green) and taking the “F” class beats from the dataset (blue), using the KNN classifier.

that the Db2, as it said in the literature, were in fact best for this type of task. Another important result is the problem with ill balanced datasets for these algorithms, showed when the task was done with and without the F type heartbeats, showing an increase in all 4 metrics after removing this label. One last important result was the improvement of the classification task after using the best denoising method tested in the denoising section (outlier removal plus quadratic spline wavelet based denoising), showing the importance of a denoising step before starting classification tasks using ECG signals.
Chapter 6

Conclusions and Future Work

Medical and biomedical advances are continuously being made in the field of cardiac monitoring. This leads to an increase of collected data that needs to be processed and analyzed. Automatic classification and denoising algorithms can help in this task. Denoising methods have been developed in order to improve the quality of ECG signals, facilitating their readability and analysis. Classification algorithms have been studied to help and automatize the recognition of arrhythmia and pathological heartbeats. In this work, several methodologies for the denoising of ECG signals were studied, as well as automatic classification of heartbeats algorithms, using both time and morphological features.

The denoising part of the work studied 3 different types of methods: the baseline one consisted in 2 median filters for baseline wander removal plus a 40 Hz low-pass filter, which achieved good results, namely a cosine similarity up to 0.834. After this, three wavelet based denoising techniques were tested, using the Sym12, Db4 and Quadratic Spline wavelets, with levels of decomposition equal to 3, 5 and 6, respectively. Sym12 wavelet didn’t show any significant improvement, neither in terms of SNR nor cosine similarity. Due to the computational cost of using wavelet decomposition for denoising, taking into account the results, this proved not to be a good method to choose for this task. Next, Db4 wavelet was used and achieved worst results than the previous two methods, reaching a maximum of 0.765 for the cosine similarity value. This method was also discarded as a choice for the denoising of ECG signals. Lastly, the quadratic spline wavelet was tested and achieved a slightly better result then the baseline median filters plus 40 Hz low pass filter, achieving a cosine similarity of up to 0.835. Even thought the improvement wasn’t very significant, this method still performed slightly better for every subject. Next, a moving average filter was tested and led to very prosing results as well, achieving a cosine similarity up to 0.826. Lastly, using denoising method that showed the best results, with the quadratic spline mother wavelet, an outlier removal step was added to the denoising process, in order to test its effect on the final results. For this, 50 prototypes were developed, 10 for each type of heartbeat that was to be recognized in the classification experience, that is, “N”, “L”, “S”, “V” and “F” heartbeats, using the K-means algorithm. After developing the prototypes, the heartbeats detected by the R-peak detector used were considered outliers if their cosine similarity to all the prototypes was bellow 0.8. After this, the quadratic spline method was applied to each heartbeat and the results were measured, achieving...
values up to 0.930 for the cosine similarity, a very significant improvement when compared to all other previous methods.

The second part of the work consisted in developing a feature set and an algorithm for automatic classification of heartbeats and test them against the database collected by the IT group. For this, firstly the different feature sets were tested in the benchmark data, the MIT-BIH database, widely used in the studies of arrhythmia and heartbeat classification. The first step consisted on the balancing of the dataset, based on the minimum common denominator between the number of samples for each class. To evaluate the performance of the algorithm, an experiment was made using only four classes of heartbeats, based on the AAMI best practices, taking out the "Q" class, with 3 different sets of features: Time features(4), wavelet based morphological features (using the Db2 with decomposition level 3 which gave a total of 20 features, 5 for each level) and the combination of both. For this experiment, the combination of both types of features achieved the best results, with an overall accuracy of 87.06%, a value that is similar to the values found in the literature. After this first step to see how the algorithm performed taking into account the classes used in most studies, an extra label was included in the classification task, the "L" label. In that way, the main focus of this part of the work was to classify the heartbeats into these 5 labels: "N", "L", "S", "V" and "F". This classification task was the second experiment, which achieved the best results also for the combination of both features, reaching an 88.70% accuracy and an 89.10% precision. One extra test was performed using the MIT-BIH database because of the morphological similarities between the "L" and the "V" heartbeats to assess how the algorithm performed in separating these two classes. This test resulted in a 98.64% accuracy using both features, which proved that despite their morphological similarities, the algorithm performed quite well in separating them.

The results for the mains focus of the classification task, i.e., classification of the 5 different types of heartbeats mentioned before, were then validated using the StM database. For this dataset, a balancing first step was also performed but, due to the fact that the "F" class, which was the minimum common denominator, only had 33 samples while the second smallest class had 1028, the balancing was done for the 1028 samples per class, except for the "F" class that was represented only by those 33 samples. After this, firstly, the feature set that proved to have the best results for the MIT-BIH experiments, that is, the combination of both time and morphological Db2 wavelet based features. This resulted in an accuracy of 81.06% and a precision of 74.23%. The difference between the results using this database can be due to the fact that the MIT-BIH database consists of recordings from 48 different subjects while the StM one had 11429 recordings, which represented a bigger variability in the data that could affect the results. Also, the fact that the balancing was done using only 33 "F" samples and 1028 for the rest of the classes was thought to have affected the results. To see if this was indeed happening, a test was made taking out the "F" label from the dataset. For this test, accuracy rose up to 82.30%, which is a small increase, but precision rose to 82.36%, which proves the influence of the balancing of the dataset in these experiments. Another experiment made on this dataset, for classifying the 5 different heartbeat classes, was to use the quadratic spline wavelet, with decomposition level 6, which led to a feature set of 32 features, 4 time ones and 32 morphological ones. This tested worst then the Db2 wavelet features,
achieving an accuracy of 65.10%.

To evaluate the effects of denoising on the dataset, an experiment that used the best method from the previous task was done. First, an outlier removal step was done on the dataset, removing all the heartbeats that had a cosine similarity lower than 0.8 to the developed prototypes. Next, the balancing step was done the same ways as before for the StM database, seeing as the minimum common denominator continued to be the "F" class, with 31 samples after the outlier removal. So the balancing was done choosing 681 samples from each of the other classes and using the 31 "F" ones. This resulted in a 80.56% accuracy. Since the number of samples decreased, to be able to compare the results using this denoising method and without it, the same balancing was done to the dataset with no outlier removal and the features were extracted without the denoising method. This showed results of 78.31% for accuracy, 2.25% less, which proves the importance of removing the outliers and denoising the signals before using the classification algorithms.

Since there are several classification algorithms, another one was tested for this task to classify between the 5 classes, namely KNN classifier. The results for this test showed an accuracy of 77.83%, lower than for the SVM ones. The precision was also lower than this value (64.40%) so an extra test was made for this classifier taking out the "F" class, reaching a precision value of 79.01% and an accuracy of 78.98%. Still these values were smaller than the SVM results, which makes SVM a better choice for this type of classification tasks.

There is still a lot of room for improvement in this field, with different features to be explored and different classifiers. As a future development, these experiments could be done taking into account other ECG leads besides the Lead I used for this thesis, that also give important information about the signals and can help in classifying the heartbeats into more than just these five labels. Other features could also be used that gather more morphological information from the signal. The time features already prove to have good results for classification so, if combined with the right set of morphological features, they can lead to better results than the ones presented in this thesis. Also, one improvement that can be made is the use of other classifiers. Recently, deep neural networks have been showing promising results for this type of tasks and can be a good improvement to these methodologies, despite their high computational cost.
Bibliography


### Appendix A

## Outlier removal in MIT-BIH Database

<table>
<thead>
<tr>
<th>Subject</th>
<th>Initial number of Heartbeats</th>
<th>Number of Heartbeats after Outlier Removal</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2272</td>
<td>2272</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1864</td>
<td>1862</td>
<td>-2</td>
</tr>
<tr>
<td>3</td>
<td>2187</td>
<td>2242</td>
<td>-55</td>
</tr>
<tr>
<td>4</td>
<td>2082</td>
<td>2081</td>
<td>-1</td>
</tr>
<tr>
<td>5</td>
<td>2224</td>
<td>758</td>
<td>-1466</td>
</tr>
<tr>
<td>6</td>
<td>1951</td>
<td>1416</td>
<td>-535</td>
</tr>
<tr>
<td>7</td>
<td>1924</td>
<td>1919</td>
<td>-5</td>
</tr>
<tr>
<td>8</td>
<td>2112</td>
<td>1587</td>
<td>-525</td>
</tr>
<tr>
<td>9</td>
<td>1784</td>
<td>1426</td>
<td>-358</td>
</tr>
<tr>
<td>10</td>
<td>2524</td>
<td>2523</td>
<td>-1</td>
</tr>
<tr>
<td>11</td>
<td>1793</td>
<td>1712</td>
<td>-81</td>
</tr>
<tr>
<td>12</td>
<td>2233</td>
<td>1599</td>
<td>-634</td>
</tr>
<tr>
<td>13</td>
<td>1794</td>
<td>1593</td>
<td>-201</td>
</tr>
<tr>
<td>14</td>
<td>1837</td>
<td>1900</td>
<td>63</td>
</tr>
<tr>
<td>15</td>
<td>1953</td>
<td>1952</td>
<td>-1</td>
</tr>
<tr>
<td>16</td>
<td>2990</td>
<td>2366</td>
<td>-624</td>
</tr>
<tr>
<td>17</td>
<td>1536</td>
<td>564</td>
<td>-972</td>
</tr>
<tr>
<td>18</td>
<td>2278</td>
<td>1317</td>
<td>-961</td>
</tr>
<tr>
<td>19</td>
<td>1987</td>
<td>1987</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>1862</td>
<td>1855</td>
<td>-7</td>
</tr>
<tr>
<td>21</td>
<td>2475</td>
<td>2475</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>1511</td>
<td>1514</td>
<td>-3</td>
</tr>
<tr>
<td>23</td>
<td>1655</td>
<td>1599</td>
<td>-56</td>
</tr>
<tr>
<td>24</td>
<td>2600</td>
<td>2569</td>
<td>-31</td>
</tr>
<tr>
<td>25</td>
<td>1889</td>
<td>1859</td>
<td>-30</td>
</tr>
<tr>
<td>26</td>
<td>2113</td>
<td>2095</td>
<td>-18</td>
</tr>
<tr>
<td>27</td>
<td>2808</td>
<td>2235</td>
<td>-573</td>
</tr>
<tr>
<td>28</td>
<td>2646</td>
<td>2606</td>
<td>-40</td>
</tr>
<tr>
<td>29</td>
<td>2986</td>
<td>2944</td>
<td>-42</td>
</tr>
<tr>
<td>30</td>
<td>2409</td>
<td>2423</td>
<td>14</td>
</tr>
<tr>
<td>31</td>
<td>3005</td>
<td>2809</td>
<td>-196</td>
</tr>
<tr>
<td>32</td>
<td>2548</td>
<td>2499</td>
<td>-49</td>
</tr>
<tr>
<td>33</td>
<td>2742</td>
<td>2742</td>
<td>0</td>
</tr>
<tr>
<td>34</td>
<td>3238</td>
<td>3229</td>
<td>-9</td>
</tr>
<tr>
<td>35</td>
<td>2753</td>
<td>2166</td>
<td>-587</td>
</tr>
<tr>
<td>36</td>
<td>3869</td>
<td>584</td>
<td>-3275</td>
</tr>
<tr>
<td>37</td>
<td>2134</td>
<td>2134</td>
<td>0</td>
</tr>
<tr>
<td>38</td>
<td>2147</td>
<td>2143</td>
<td>-4</td>
</tr>
<tr>
<td>39</td>
<td>2048</td>
<td>2041</td>
<td>-7</td>
</tr>
<tr>
<td>40</td>
<td>2197</td>
<td>2196</td>
<td>-1</td>
</tr>
<tr>
<td>41</td>
<td>2486</td>
<td>3734</td>
<td>-1248</td>
</tr>
<tr>
<td>42</td>
<td>2590</td>
<td>2564</td>
<td>-26</td>
</tr>
<tr>
<td>43</td>
<td>1942</td>
<td>1860</td>
<td>-82</td>
</tr>
<tr>
<td>44</td>
<td>2256</td>
<td>2239</td>
<td>-17</td>
</tr>
<tr>
<td>45</td>
<td>1570</td>
<td>1568</td>
<td>-2</td>
</tr>
<tr>
<td>46</td>
<td>1782</td>
<td>1774</td>
<td>-8</td>
</tr>
<tr>
<td>47</td>
<td>2786</td>
<td>2725</td>
<td>-61</td>
</tr>
<tr>
<td>48</td>
<td>2749</td>
<td>2748</td>
<td>-1</td>
</tr>
</tbody>
</table>

Figure A.1: Results of Outlier removal in MIT-BIH Database.