Bruno Guerreiro Brázio

Analysis of heart rate variability on diabetic patients

UNIVERSIDADE DO ALGARVE

FACULDADE DE CIÊNCIAS E TECNOLOGIAS

2017

Bruno Guerreiro Brázio

Analysis of heart rate variability on diabetic patients

Mestrado Integrado em Engenharia Eletrónica e Telecomunicações
Trabalho efetuado sob a orientação de:
Professora Doutora Maria da Graça Ruano

UNIVERSIDADE DO ALGARVE

FACULDADE DE CIÊNCIAS E TECNOLOGIAS

2017

Analysis of heart rate variability on diabetic patients

Declaração de Autoria

Declaro ser o(a) autor(a) deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

Assinatura do candidato:____

Bruno Guerreiro Brázio

© Copyright ***Bruno Guerreiro Brázio***

A Universidade do Algarve tem o direito, perpétuo e sem limites geográficos, de arquivar e publicitar este trabalho através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, de o divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

ACKLOWLEDGEMENT

Foremost, I would like to express my sincere gratitude to my advisor Prof.ª M. Graça Ruano for giving me the opportunity to develop this project, for the continuous support on this journey, her availability, motivation, enthusiasm and knowledge.

I also would like to thank my family: my parents Joaquim and Lilia, for their efforts to providing me the best education possible, for supporting me through my academic course and for loving me no matter what, my grandfather Manuel which inspired me so much during my thesis and my girlfriend Marlene for the support, encouragement and helping me when I needed the most.

I would like to also thank my colleagues who provided me great moments on this journey.

Finally, I would like to thank the sponsorship of LINK - Linking Excellence in Biomedical knowledge and Computational Intelligence Research for Personalized Management of Cardiovascular Diseases within Personalized Health Care, 692023 H2020, (2016-2019).

ABSTRACT

Diabetes mellitus (DM*)* is a chronic condition in which the body produces insufficient insulin, or it cannot be used properly. This condition induces abnormal cardiovascular behaviour due to the irregular pattern of glucose levels in blood, being responsible for an increased morbidity within DM patients. So, researching non-invasive methods of early detection of cardiovascular pathologies is a valuable help for clinical diagnose.

This work concentrates on the analysis of the electrocardiogram (ECG) of DM patients with different cardiac pathologies. The signal processing methodology adopted is to consider the ECG signal as a time-series. The identification of signals' pattern for a specific pathology is searched by analysing the similarity between time-series representations of the same type of pathology and verifying the difference among differentiated pathologies. Searching for time-series similarity of non-stationary signals may be performed in time, frequency or transformed domains. Each of these similarity methods present pros and against which have to be evaluated within the cohorts considered in this study.

A collection of seven similarity methods was assessed on their ability to find the similarity among each cohort, considering the ECG 12 conventional leads' signals together with the 3 Frank leads' signals. The cohorts were composed of ECG signals available at the public database Physionet. Different cohorts were created considering groups of data related to patients with the same diagnosis (myocardial infarction, *diabetes mellitus,* renal insufficiency, hyperuricemia, arterial hypertension and healthy controls), gender and age range. The performance of the similarity measurement methods was evaluated by confronting the signal processing results with the clinical annotations contained in the database.

Also, to broaden the comparison of the obtained results with other researchers who provide conclusions based on the heart rate variability (HRV), an analysis of this parameter will also be reported.

Analysis of the results enabled identification of the best performed similarity method – which was Pearson's correlation coefficient method, to use under specific illness constraints – diabetes mellitus and myocardial infarction, being obtained, in this case, a pattern with 73% similarity. Confronting the obtained results with the published ones

enabled confirmation of the most reliable ECG leads (aVL, L1, V4 and VZ) to identify DM myocardial infarction. In what concerns de HRV analysis we concluded that CVD patients, in overall, have lower HRV in comparison with healthy individuals.

Keywords: *Diabetes mellitus*, time-series, data mining, similarity measures, electrocardiogram (ECG), heart rate variability (HRV).

RESUMO

Diabetes mellitus (DM) é uma condição crónica em que o corpo produz insulina insuficiente, ou a qual não pode ser usada corretamente. Esta condição induz o comportamento cardiovascular anormal devido ao padrão irregular de níveis de glicose no sangue, sendo responsável por uma maior morbidade nos pacientes com DM. Assim, a pesquisa de métodos não invasivos de deteção precoce de patologias cardiovasculares é uma valiosa ajuda para o diagnóstico clínico.

Este trabalho concentra-se na análise do eletrocardiograma (ECG) de pacientes com DM com diferentes patologias cardíacas. A metodologia de processamento de sinal adotada consiste em considerar o sinal de ECG como uma série temporal. A identificação do padrão de sinais para uma patologia específica é pesquisada analisando a semelhança entre representações de séries temporais do mesmo tipo de patologia e verificando a diferença entre patologias diferenciadas. A procura de semelhanças em séries temporais de sinais não estacionários pode ser realizada nos domínios do tempo, frequência ou transformados. Cada um desses métodos de semelhança apresenta prós e contras, os quais devem ser avaliados dentro das coortes consideradas neste estudo.

Uma coleção de sete métodos de similaridade foi testada e avaliada quanto à sua capacidade de encontrar a semelhança entre cada coorte, considerando os 12 sinais convencionais do ECG (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) e ainda os sinais de 3 sensores do tipo Frank (vx, vy, vz). As coortes foram compostas por sinais de ECG disponíveis no banco de dados público Physionet. Coortes diferentes foram criadas considerando grupos de dados relacionados a pacientes com o mesmo tipo de diagnóstico (infarto do miocárdio, diabetes mellitus, insuficiência renal, hiperuricemia, hipertensão arterial e controle de pessoas saudáveis), género e faixa etária.

O desempenho dos métodos de medição de similaridade foi avaliado ao confrontar os resultados do processamento do sinal com as anotações clínicas contidas no banco de dados.

Além disso, para ampliar a comparação dos resultados obtidos com a de outros investigadores que apresentam conclusões com base na variabilidade da frequência cardíaca, uma análise desse parâmetro também será relatada.

A análise dos resultados permitiu a identificação do método de semelhança com melhor desempenho - o método do coeficiente de correlação de Pearson, o qual deve ser usado mediante restrições específicas de doença, isto é, diabetes *mellitus* e infarto do miocárdio, sendo obtido, neste caso, um ciclo cardíaco padrão com 73% de similaridade aos casos analisados. Confrontados os resultados obtidos com os publicados permitiu a confirmação das derivações ECG mais confiáveis (aVL, L1, V4 e VZ) para a identificação do infarto do miocárdio em pacientes com DM. No que diz respeito à análise da variação da frequência cardíaca, concluímos que pessoas com doenças cardiovasculares têm menor variação do ritmo cardíaco em comparação com pessoas saudáveis.

Palavras-chave: *Diabetes mellitus*, séries temporais, mineração de dados, medidas de semelhança, electrocardiograma, variação do ritmo cardíaco.

INDEX

INDEX OF FIGURES

Page

INDEX OF TABLES

Page

ABBREVIATION'S LIST

bpm	beats per minute
CVD	Cardiovascular Disease
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
DM	Diabetes Mellitus
DTW	Dynamic Time Warping
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
ED	Euclidean Distance
FT	Fourier Transform
HPF	High-Pass Filter
HRV	Heart Rate Variability
KLT	Karhunen-Loève Transform
LPF	Low-Pass Filter
ms	millisecond
WHO	World Health Organisation
WT	Wavelet Transform

1. INTRODUCTION

According to World Health Organisation (WHO)

> Diabetes mellitus is a chronic disease caused by inherited and/or acquired deficiency in production of insulin by the pancreas, or by the ineffectiveness of the insulin produced. Such a deficiency results in increased concentrations of glucose in the blood, which in turn damage many of the body's systems, in particular the blood vessels and nerves [1].

Currently, WHO has published several recommendations on diagnostic values for blood glucose concentration (last modified in 1999), for a disease that causes suffering and hardship for approximately 60 million people in the European region for a total of 4422 million all over the world [1, 2, 3].

There are many complications associated with diabetes mellitus [1], such as:

1. Diabetic retinopathy which can lead to blindness and visual disability;
2. Kidney failure, which in advanced stages obliges to haemodialysis;
3. Heart disease, which develops at different types, being hypertension and coronary diseases the most frequent;
4. Diabetic neuropathy can lead to sensory loss and damage to the limbs;
5. Diabetic foot disease with subsequent limb amputation.

In this work, we will concentrate on relating DM with point 3, this is, with heart diseases. We will also follow the line of previous investigations within the research group, that is to say that we will be considering the electrocardiogram (ECG) signals of patients, and processing them as time-series. Time-series are an important class of temporal data that arise from various sources, and to analyse the amount of data from those sets we need to use several data mining techniques. Working with this kind of data representation mostly means that issues such as non-constant sampling rate, noise, enormous amounts of data, etc may be overcome [4].

Measuring similarity within time series plays an important role in finding a pattern, enabling prediction and knowledge discovery. In clinical context if we find a pattern of a specific pathology we can use that knowledge for disease prediction. That might mean allowing medical doctors with additional diagnosis support and therefore improving medical prescriptions, eventually decreasing the number of screening medical exams with their consequent economic savings, besides enabling better disease control with the correspondent social impact.

As mentioned this work concentrates on DM cardiac pathologies ECG analysis. To find a pattern for DM different methods of similarity measures were considered on the conventional 12 leads (i, ii, iii, avr, avl, avf, v1, v2, v3, v4, v5, v6) together with the 3 Frank leads (vx, vy, vz) ECG from different cohorts. Several time-series similarity measuring methods were tested, whose performance were evaluated by confronting the signal processing results with the clinical annotations.

Analysis of the results enables not only the identification of the ECG leads which are more representative of the classes of pathologies under consideration, but also identification of the best similarity measures for this kind of experiments.

Many authors analyse CVD through heart rate variability (HRV). It will also be addressed the evaluation of this clinical parameter.

The structure of this thesis is organized into five chapter.

The present chapter presents a general introduction of this thesis.

Chapter 2 describes some fundamental concepts for the understanding of the upcoming chapters. Such as interpreting an electrocardiogram, what heart rate variability represents, a brief description of time-series and an overview of the similarity measures methods considered.

Chapter 3 exposes the methodology employed on this study. The sequence of experiments and where and how data was gathered to compose the case-study cohorts is explained. The approach followed to establish the range of similarity values to be seek is also detailed.

On Chapter 4 the full description of each implemented experiment is given. Similarity measurements were exhaustively computed for different cohorts to enable a pattern identification of cardiac DM comorbidities.

Chapter 5 concentrates the results obtained for the experiments listed in the previous chapter and the conclusions driven, and as well general conclusions and indicates future research guidelines.

2. REVIEWED CONCEPTS

This chapter presents an explanation about some main concepts, which are fundamental for the understanding of the upcoming chapters.

It will be described how to interpret a cardiac signal, the importance of an ECG and how it relates with HRV, a review of time-series and an explanation of similarity measures approaches. The main idea of this thesis is finding a specific pathology pattern with the DM patients' using different similarity measures.

2.1. CARDIAC SIGNALS

The human heart's electrical system controls all the events that occur when the heart pumps blood. This electrical system is also called cardiac conduction system. We can see a graphical picture of the heart's electrical activity in an ECG. [5, 6]

2.1.1. ELECTROCARDIOGRAM

As it was mentioned, the ECG is the electrical manifestation of the contractile activity of heart, and can be recorded easily with surface electrodes on the limbs and chest. The ECG is one of the most commonly known and used biomedical signal. [6]

Many researches developed methods of ECG analysis over the centuries, which improved significantly our understanding of ECG as a clinical tool [7]. Nowadays, the ECG is an essential part of the initial evaluation of patients presenting cardiac complaints. [8]

It is not hard to understand why this biomedical signal is so recognized and, most likely, the most used biomedical signal. The rhythm of the heart in terms of beats per minute (bpm) is easily estimated by counting the peaks of the signal. But more important is the fact that the ECG shape is altered by CVDs and abnormalities such as myocardial ischemia and infarction, ventricular hypertrophy, and conduction problems. [6]

In Figure 2.1., we have a typical ECG signal of a healthy person.

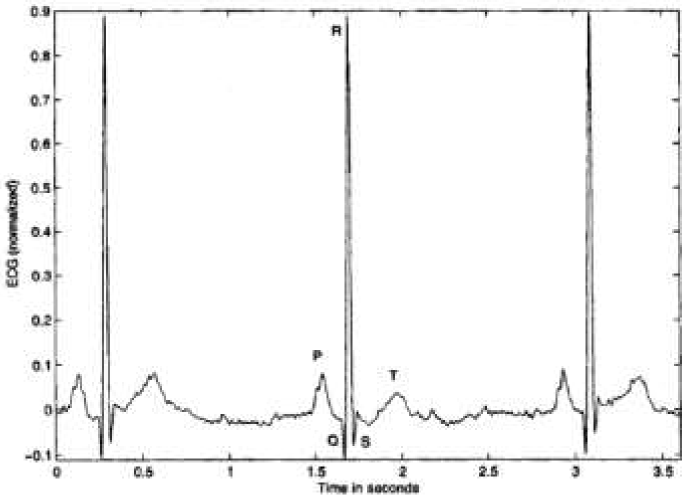


Figure 2.1 A typical ECG signal (male subject of 24 years old) [6]

In the Figure above we have represented an ECG signal, where its components are marked as P wave which records the electrical activity through the atria, as QRS complex which records the movements of electrical impulses through ventricles, as ST segment which shows when the ventricle is contracting but there is no electricity flowing thought it and finally as T wave which shows when the lower heart chambers are resetting electrically and preparing for their next muscle contraction [9].

2.1.1.1. ECG Data Acquisition

2.1.1.1.1. Standard 12-Lead ECG

Usually, in clinical practise, the standard 12-Lead ECG is obtained using four limbs leads and six chest leads in different positions. The right leg is used to place the reference electrode. The left, right arm and left leg are used to get leads I, II and III. A combined reference knows as *Wilson's central terminal* is formed by combining the left arm, right arm and left keg leads, and is used as the reference for chest leads. The augmented limb leads known as aVR, aVL and aVF, where aV stands for the augmented lead, R for the righ arm, L for the left arm and F for the left foot. These leads are obtained by using the exploring electrode on the limb indicated by the lead name, with the reference being *Wilson's central terminal* without the exploring limb lead [6], which can be seen in the Figure 2.2.

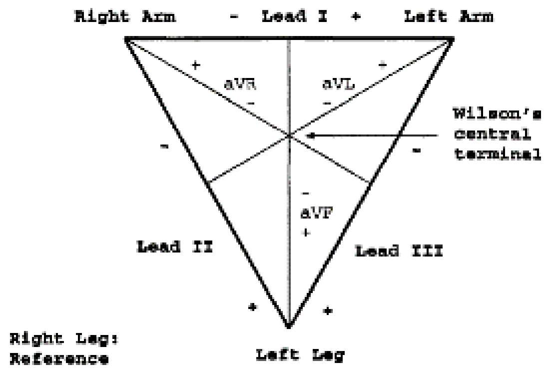


Figure 2.2 Einthoven's triangle and the axes of the six ECG leads formed by using limb leads. [6]

The six chest leads, which are V1-V6, are obtained from six standardized position on the chest with *Wilson's central terminal* as reference [6]. Which is represented in the Figure 2.3.

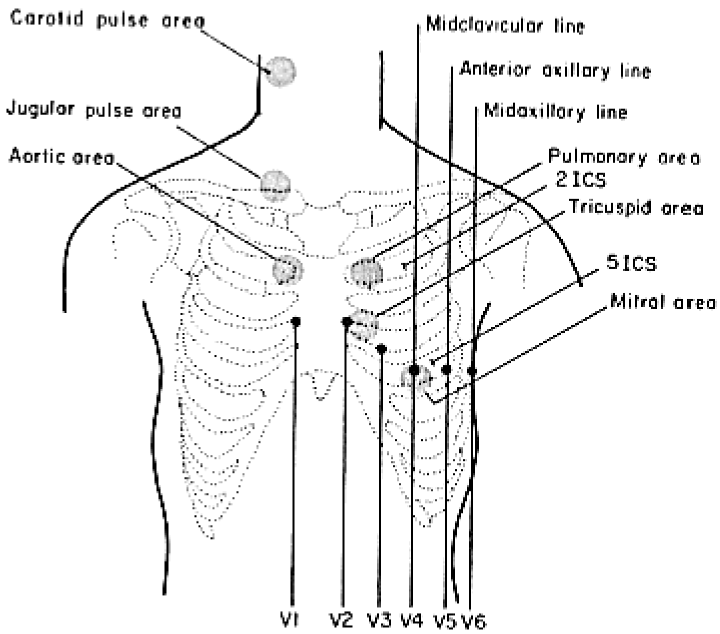


Figure 2.3 Positions for placement of the chest leads V1-V6 for ECG, auscultation areas for heart sounds, and pulse transducer positions for the carotid and jugular pulse signals. [6]

These 12-lead system serves as the basis of the standard clinical ECG, its interpretation in mainly empirical, based on experimental knowledge. In the Figure 2.4, we have an example of a standard 12-lead ECG representation [6].

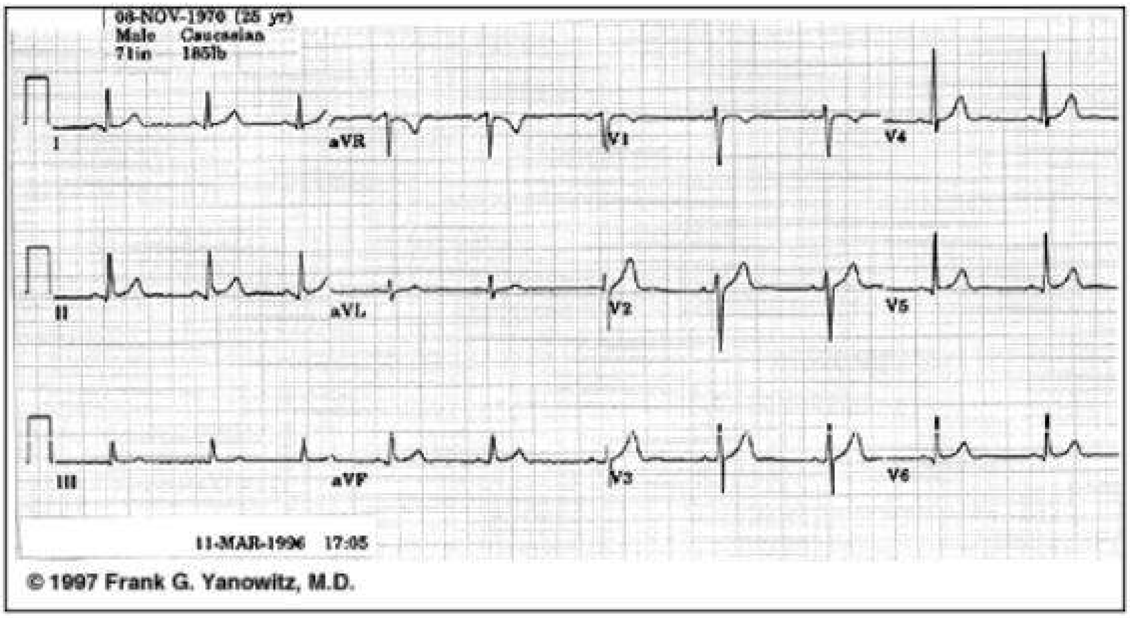


Figure 2.4 Standard 12-lead ECG signals of a healthy male adult. [10]

2.1.1.1.2. Frank Lead system

In 1956 Frank described the heart as a rotating dipole within space. In principle, a rotating dipole works like a battery with a positive and negative pole spinning in space. Frank asked himself how the rotating dipole could be effectively being measured and described. He placed the electrodes on the body so the measured leads X, Y and Z were placed in a row, thereby making a cartesian coordinate system represented in the Figure 2.5. [11]

This system may not substitute but complement Standard 12-Lead ECG. [12]

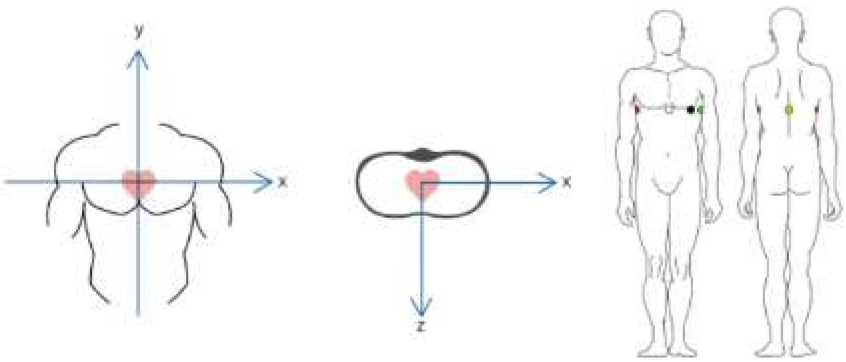


Figure 2.5 The vector ECG views the heart as a rotating dipole. Electrode Position/Vertical Axes. [11]

In the Figure 2.6 we have an example of Frank's Lead ECG signal.

Figure 2.6 Frank Lead ECG signal. [13]

For further reading visit [14].

2.2. HEART RATE VARIABILITY

The heart rate variability measures the specific changes in time between successive heart beats. The time between beats is measured in milliseconds (ms) and is called a "R-R interval". And it is represented in the Figure 2.7 [15]

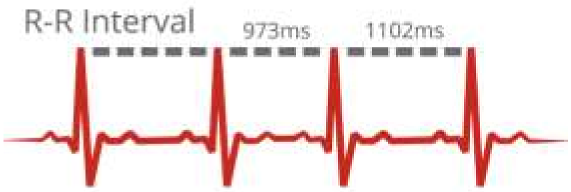


Figure 2.7 Heart rate variability. [15]

The HRV is a non-invasive and sensitive technique to evaluate cardiovascular autonomic control [16]. A low HRV is related with stress, negative psychosocial events and CVD's. It is also associated with a 32-45% increased risk of a first cardiovascular event in populations without known CVD [15, 17].

2.3. TIME-SERIES

A time-series represents a collection of values obtained from sequential measurements over time [18].
Recently, the increasing usage of time-series, has encouraged multiple researches to develop related data mining techniques [19]. Time-series is an important class of temporal data objects, and it can be easily obtained from scientific and financial applications (e.g. ECG, daily temperature, weekly sales totals, and prices of mutual funds and stocks) [4].

In this context, time-series data mining fundamental problem is what method should be used to obtain data classification with precision and accuracy.

To avoid inaccuracies, before any data mining task, pre-processing techniques like normalization and noise removal are required.

Moreover, similarity measure between time-series and segmentation are two core tasks for various time-series mining processes. Based on the time-series representation, different mining tasks can be found in the literature and they can be roughly classified into four fields: pattern discovery and clustering, classification, rule discovery and summarization. Some of the researches concentrates on one of these fields, while the others may focus on more than one of the above processes [4, 20, 19].

In this thesis, the process is to find a pattern in DM patients and consequently perform clustering.

One of the major reasons for time-series representation is to reduce the dimension (i.e. the number of data points) of the original data. The simplest method it might be *sampling.* In this method, a rate of m/n is used, where m is the length of a time series P and n is the dimensionality reduction. However, the sampling method has the disadvantage of distorting the shape of compressed time series (if the sampling rate is too low), which can be seen in [18].

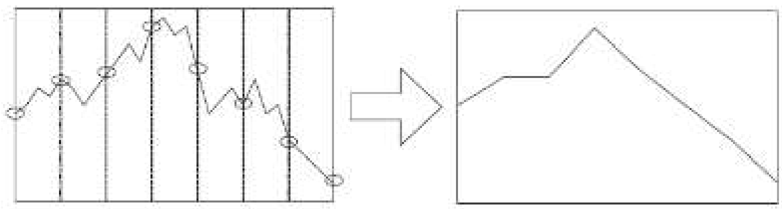


Figure 2.8 Time series dimensionality reduction by sampling [18].

However, there are better option, for instance reducing the dimension by preserving salient points, these points are called as perceptually important points (PIP). We can see the improvement in Figure 2.9 [18].

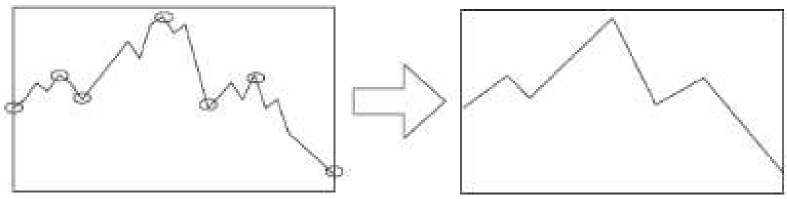


Figure 2.9 Time series compression by data point importance [18].

However, we still have a loss of information, this is the reasoning we also consider frequency transform based methods to measure similarity among different time-series, in the upcoming section, since they can reduce its dimensionality without any significant losses.

2.4. SIMILARITY MEASURES

A usual data mining task is the estimation of similarity among objects. Normally, similarity among series is represented as [0, 1], where "one" it's the absolute maximum for similarity [20].

If we work with an efficient and effective method of measuring similarity, we can find a relation among the time-series. This will greatly increase our accuracy and prediction on our analysis [20].

There are two main groups of similarity measures, which are time domain and transformed based methods, but before choosing one we need to know the characteristics of those methods [20, 21].

2.4.1. TIME DOMAIN METHODS

Usually approaches using time domain methods are the simplest, computationally speaking this doesn't mean that time domain methods are always faster than the transformed based methods, it depends how long and complex the time series are.

In this sub chapter, it is briefly explained methods like *Minkowski distance*, *Euclidean distance* (ED), *Dynamic time warping* (DTW), *Mahalanobis distan*ce and *Correlation coefficient*.

As it was mentioned in 2, the similarity measures presented follow the reasoning presented in [20], since both researches are included in the same research project.

2.4.1.1. Euclidean Distance

If we consider two time-series $T(t) = \{t(1), t(2), \ldots, t(N)\}$ and $S(t) = \{s(1), s(2), \ldots, s(N)\}$ we can estimate the similarity between those series by measuring the distance between each of their pair of points, the lesser the distance the greater the similarity and *vice versa* [20, 21].

So, the Euclidean distance is represented by:

$$D_E(T(t), S(t)) = \left(\sum_{t=1}^{N} |T(t) - S(t)|^2\right)^{\frac{1}{2}} \qquad (1)$$

On the other hand, this method is hard to use in some applications due to its drawbacks. As examples, the distance in this method can only be measured in straight-line, so we can only compare time-series with the same length, it doesn't handle noise and it is very sensitive to signal transformations (Shifting, uniform amplitude scaling, uniform time scaling, uniform bi-scaling, time warping and non-uniform amplitude scaling) [20, 21].

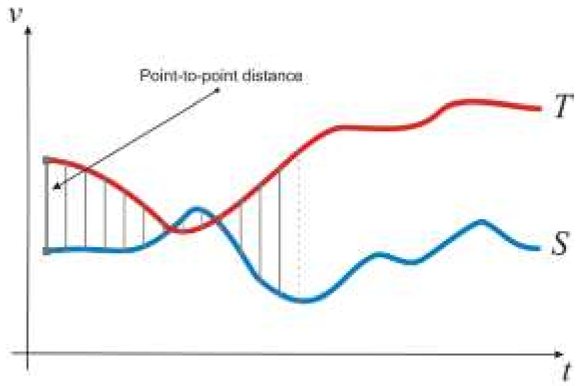


Figure 2.10 *T* and *S* are two time-series of a variable *v*, along the time axis *t*. The Euclidean distance results in the sum of the point-to-point distances, along all the time series [21].

To overcome these issues, changes have been made on the principle of DTW [20, 21].

2.4.1.2. Dynamic Time Warping

Dynamic time warping gives more robustness of the similarity computation, although it is also computationally expensive. With this method, we can compare time-series with different lengths since one-to-one point comparison (which was used in Euclidean distance method) was replaced by a many-to-one (or vice-versa) approach. This improvement allows DTW to recognize shapes, even with signal transformations [21].

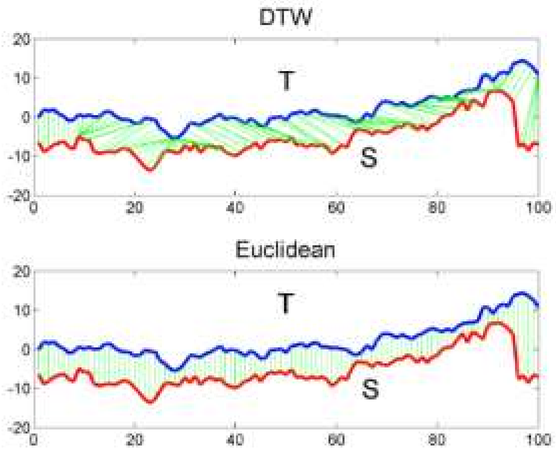


Figure 2.11 Difference between *DTW* distance and Euclidean distance. The former allows many-to-one point comparisons, while Euclidean point-to-point distance (or one-to-one) [21].

Given two time-series $T(t) = \{t(1), t(2), \dots, t(N)\}$ and $S(t) = \{s(1), s(2), \dots, s(M)\}$ where N and M represent respectively the length of the series, DTW method exploits information contained in a NxM distance matrix, as it follows [20, 21] :

$$distMatrix = \begin{pmatrix} d(T_1,S_1) & d(T_1,S_2) \dots & d(T_1,S_M) \\ d(T_2,S_1) & d(T_2,S_2) & \\ \vdots & & \ddots \\ d(T_N,S_1) & & d(T_N,S_M) \end{pmatrix} \qquad (2)$$

where *distMatrix* (i, j) corresponds to the distance of ith point of T and jth point of S.

The *DTW* objective is to find the *warping path* $W = \{w_1, w_2, \dots, wk, \dots, wK\}$ of contiguous elements on *distMatrix* such that it minimizes the following function [21]:

$$DTW(T(t), S(t)) = min\left(\sqrt{\sum_{k=1}^{K} wk}\right) \qquad (3)$$

The warping path can be efficiently computed using dynamic programming. Using this method, a cumulative distant matrix γ of the same dimension as the *distMatrix*, is created to store in the cell (*i, j*) the minimum distance among adjacent cells (optimal path) [20, 21].

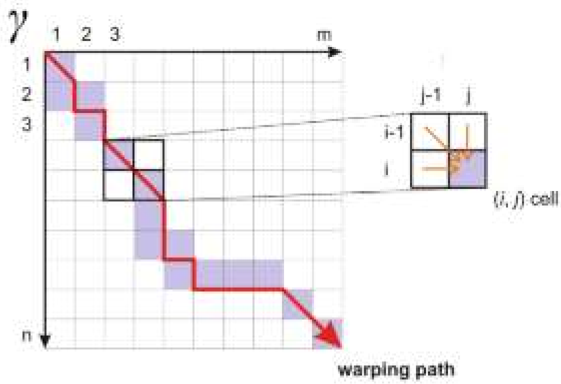


Figure 2.12 Warping path computation using dynamic programming [21].

In many cases, this method can bring unexpected results. For example, when many points of a time-series *T* are mapped to a single point of another series *S*. A common way to fix these events is to restrict the warping path in such a way that it must follow a direction along diagonal [21].

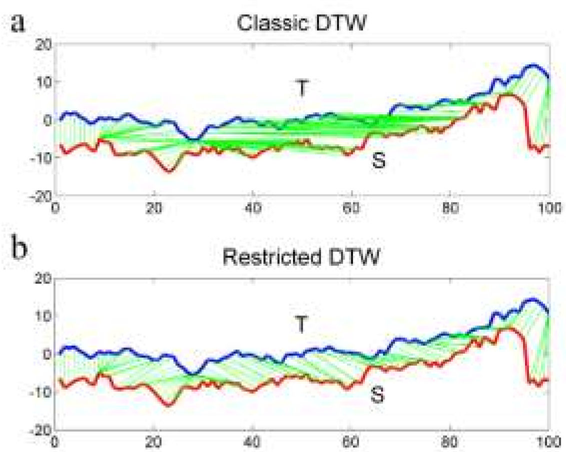


Figure 2.13 Different mappings obtained with the classic implementation of *DTW* (a), and with the restricted path version using a threshold $\delta = 10$ (b). [21].

In Figure 2.13, we fixed our results by restricting the DTW method with the previous method [20, 21]. For further reading please visit [21].

2.4.1.3. Minkowski Distance

This method is one of the simplest time domain methods and can be considered as a generalization of the Euclidean distance [20, 21].

The Minkowski distance is represented as:

$$D_{Minkowski}(T(t), S(t)) = \left(\sum_{t=1}^{N} |T(t) - S(t)|^{\lambda}\right)^{\frac{1}{\lambda}} \qquad (4)$$

Where $\lambda \geq 1$.

In the case of λ=1 we have the same concept of Manhattan distance method, when λ=2 we have Euclidean distance method [20, 21].

2.4.1.4. Mahalanobis Distance

The Mahalanobis distance is defined as a dissimilarity measure between time-series with the same statistical distribution and the covariance matrix C of the multivariate random variable.

It is defined as:

$$D_{Mahalanobis}(T(t), S(t)) = \left(\left(T(t) - S(t)\right)^{T} C^{-1} (T(t) - S(t))\right)^{\frac{1}{2}} \qquad (5)$$

The advantage of using this method is that is takes into consideration the correlations between the time-series stocked in matrix C. Because of this we can identify different patterns and analyse them based on a reference point [20, 21].

2.4.1.5. Pearson's Correlation Coefficient

Pearson's method is a statistical measure which measures the strength of a linear relationship between paired data.

It is invariant to shifting and scaling, being expressed when applied to a sample as [20]:

$$r_{PCC} = \frac{\sum_{i=1}^{N}(T_i - \bar{T})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{N}(T_i - \bar{T})^2}\sqrt{\sum_{i=1}^{N}(S_i - \bar{S})^2}} \qquad (6)$$

Where N is the number of samples, T_i and S_i are single samples indexed with i. Lastly but not least, $\bar{T}$ and $\bar{S}$ are the sample mean, represented as:

$$\bar{T} = \frac{1}{N}\sum_{i=1}^{N} T_i \ and \ \bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_i \quad (7)$$

These samples are constrained by default between -1 and 1. The closer the value is to 1 or -1, the stronger the linear correction is. Positive values denote positive linear correlation, negative values denote negative linear correlation and zero value means that there is no correlation [20, 22].

This method presents the advantage of being unaffected by dispersion differences across linear transformations. [20]

2.4.2. TRANSFORMED BASED METHODS

It was already stated that one of the goals while mining time-series data, is to work with a representation with fewer data points than the raw data, this can be achieved by reducing its dimensionality, while maintaining its main properties [21].

According to the results of previous researches [23, 20] the Transform based methods used in this work were Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT), and they will be briefly explained.

2.4.2.1. Discrete Fourier Transform

The Discrete Fourier Transform (DFT) is a typical data reduction technique which was used to map time-series data from the time domain to the frequency domain [20, 19, 23].

The basic idea of Fourier Transform is to decompose a signal, where any signal can be represented as a sine and cosine basis function, each function being known as a Fourier coefficient. The most important feature of this method is data compression, which allows us to reconstruct the original signal by the corresponding waves with higher Fourier coefficients. By taking into consideration only the first Fourier coefficients for indexing they effectively reduce the search space and speed-up the similarity query [20, 19].

The exponential representation of DFT in frequency domain could be defined as:

$$T(F) = DFT\big(T(t)\big) = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} T(i) e^{-\frac{j2\pi Fi}{N}} \tag{8}$$

Where *F*=0, …, N-1,

$$e^{-\frac{j2\pi \mathrm{Fi}}{N}} = cos\left(\frac{2\pi \mathrm{Fi}}{N}\right) + jsin\left(\frac{2\pi \mathrm{Fi}}{N}\right) \tag{9}$$

From Euler's equation, we can conclude that the Fourier Transform (FT) decompose time-series into periodic signals in the frequency domain, where cosine functions represent the real part of the spectrum and the sine functions the imaginary part of the spectrum [20, 23, 19].

Similarly, to [20], it was used Discrete Cosine Transform (DCT) as a similarity method, where it only uses the real part of the spectrum, which will be briefly explained in the next sub-chapter.

A fundamental property of DFT is guaranteed by Parseval's Theorem, which asserts that the energy calculated on the time-series domain for signal *f* is preserved on the frequency domain. [20, 23, 19]

The energy *E(f)* of a signal *f* is given by:

$$E(f) = \sum_{k=0}^{N-1} |F(k)|^2 = E(F) \tag{10}$$

If we use the Euclidean distance method, by this property, the distance calculated between two signals in time domain will be the same as in the frequency domain. The reduced representation is built by only keeping the first *k* coefficients.

The main drawback of DFT is the choice of the best number of coefficients to keep for a good reconstruction of the original signal [20, 19, 24, 21].

2.4.2.1.1. Discrete Cosine Transform

As mentioned in **2.4.2.1.** DCT is the real part of the FT and for a time-series with length of N, $T(t) = \{t(1), t(2), \ldots, t(N)\}$ is derived from a simplified form of equation (8) that is shown below:

$$T'(t) = p(t) \sum_{k=1}^{N} C_k \cos \langle \frac{\pi(2k-1)(t-1)}{2N} \rangle \qquad (11)$$

In equation (11), t=, …, N, the parameters C_k are scale factors of the cosine wave and $p(t)$ represents a normalization coefficient that could be defined as equation (12):

$$p(t) = \begin{cases} \frac{1}{\sqrt{N}} & , \quad t = 1 \\ \sqrt{\frac{2}{N}} & , \quad 2 \leq t \leq N \end{cases} \qquad (12)$$

For measuring the similarity between two time-series T(t) and S(t) based on DCT coefficients, the first *m* coefficients could represent a good approximation of time-series so this distance could be a good measure of similarity. The template signal, T(t), and the added variation signal, S(t), are decomposed into DCT coefficients and the similarity is measured according equation (13):

$$D_{DCT}\big(T\,(t), S(t)\big) = \sqrt{\sum_{k=1}^{m} \left(C_{k_T} - C_{k_S}\right)^2} \qquad (13)$$

This distance could be the same as the Euclidean distance if we consider all coefficients m=N [20, 23].

Similarly to [20], in this work was considered the first m=4 coefficients to achieve 90 percent of accuracy on the approximation.

2.4.2.2. Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) was proposed to replace DFT. This new technique has several pros over the DFT.

It provides time and frequency information simultaneously, it is more flexible (a wide range of different DWT bases exist, whereas the DFT is just based on *cos* and *sin* with different frequencies) and it has more discrimination power than DFT. The cost of these advantages is greater computational complexity, the flexibility which was an advantage can also be considered as a disadvantage once it can be hard to choose which basis to use. Also, the results are harder to interpret (less intuitive) [20, 19].

The basic idea of Wavelet Transform is data representation in terms of sum and difference of prototype functions, known as *wavelets.* Similarly, to DFT, wavelet coefficients give local contributions to the reconstruction of the signal, while Fourier coefficients always represent global contributions to the signal over time [20, 19, 24, 21].

There are plenty of wavelet's families, although in this work, similarly to [20], we will be using Haar which is the simplest possible wavelet. An example of DWT based on Haar is shown in the

Table 2-1.

The general Haar transform $H_L(T)$ of a time-series T of length n can be formalized as in equation (14):

$$A_{L'+1}(i) = \frac{A_{L'}(2i) + A_{L'}(2i+1)}{2}$$

$$D_{L'+1}(i) = \frac{D_{L'}(2i) - D_{L'}(2i+1)}{2} \tag{14}$$

$$H_L(T) = (A_L, D_L, D_{L-1}, \dots, D_0)$$

Where $0 < L' \leq L$, and $1 \leq i \leq n$.

Level (L)	Averages coefficients (A)	Wavelet Coefficients (D)
1	10,4,6,6	
2	8,6	3,0
3	7	1

Table 2-1 The *Haar* Transform. [21]

In the Table 2-1, we have the Haar transform.of T = {10, 4, 8, 6} depends on the chosen level, and corresponds to merging Averages coefficients (column 2) at the chosen level and all Wavelet coefficients (column 3) in decreasing order among the chosen level. At level 1 the representation is the same as time series. H1(T) = {10, 4, 6, 6} + {} = {10, 4, 6, 6} = T. At level 2, H2(T)= {8, 6} + {3, 0} + {} = {8, 6, 3, 0}. At level 3 is H3(T) = {7} + {1} + {3, 0} = {7, 1, 3 0}. [21]

Decomposing a signal with wavelets, it should be mentioned that two types of filter are used. A high-pass filter (HPF) and a low-pass filter (LPF), as it is represented in the figure below:

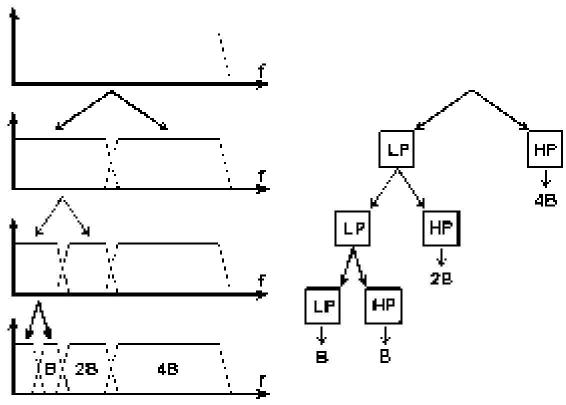


Figure 2.14 Splitting the signal spectrum with an iterated filter bank [25].

If we regard the wavelet transform as a filter bank, we can consider the wavelet decomposing a signal as passing through this filter bank. We split the signal spectrum in two equal parts, a LPF and a HPF part, where the LPF applies a scaling function while the HPF applies the wavelet function. Once the functions were applied what will remain in the LFP part would be an approximation of the signal and in the HPL part would be the details of the signal. We can keep splitting the spectrum until we are satisfied with the detail and scale of the lighter version of the signal, which can be limited by the amount of resources or the computational power available. We can see in the figure below the decomposition tree, where its resolution depends on the different scale and detail (levels) [20, 21, 25].

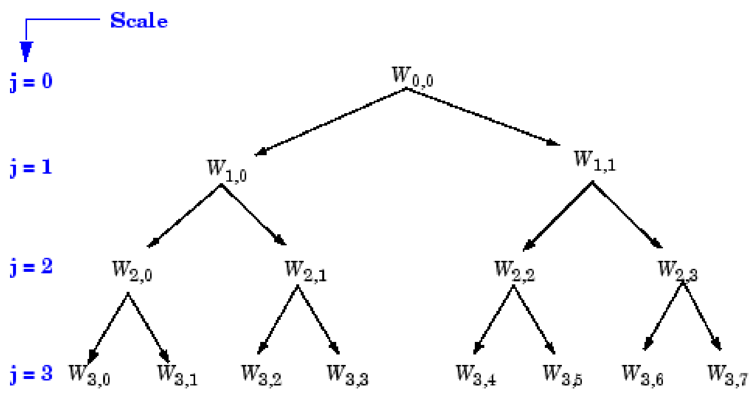


Figure 2.15 Decomposing tree and its respective level of decomposition [26]

Time-series can be decomposed into linear combinations of the basis-functions. So, the signal could be approximated by different resolutions through the following equation:

$$T'(t) = \sum_{j=1}^{J} \varphi_j(t) \tag{15}$$

J represents the level of decomposition and $T'(t)$ is an approximation of the time-series and its accuracy is dependent on the level of the basic functions $\varphi_j(t)$ that are used to reconstruct the signal. These functions are orthogonal and generated by multiplication of the coefficients $d_j \in \mathbb{R}$, which are scalers, with different orthogonal wavelet basis $\psi_j(t)$, so:

$$\varphi_j(t) = d_j \, \psi_j(t) \tag{16}$$

The trend of the input function is captured in approximation to the original function ϕ(t), while localized changes are kept as sets of detailed functions, ranging from coarse to fine ψ(t). If we consider, $\varphi_1(t) = C_{0,0}\phi_{0,0}(t)$ and J as level of decomposition and $j = \log_2 N$, then DWT is computed as it shows:

$$\tilde{T}_j(t) = C_{0,0}\phi_{0,0}(t) + \sum_{j=0}^{j-1} \sum_{k=0}^{2j-1} d_{j,k}\psi_{j,j}(t) \tag{17}$$

Exploring the data reduction ability of DWT for measuring the similarity between time-series, in this work we followed this methodology by combining the Haar wavelet decomposition with the Karhunen-Loève transforms (KLT) to optimally reduce the number of wavelet basis [20, 25].

2.4.2.3. Karhunen-Loève Transform

When we measure the similarity with DWT combined with KLT, the distance between time-series is measured but the reduced number of coefficients are considered according Karhunen-Loève theorem. This method decomposes the time-series into the basic functions which are orthogonal to each other.

Those are obtained as eigenvectors of the covariance matrix composed of the wavelet basis [23]. The approximation of the signal is acquired by reducing the number of basis that have been employed in the similarity measuring instead of reducing the signal. This reduction is obtained from the first highest J eigenvalues of the correspondent covariance matrix [23].

The first step is to decompose the template time-series $T(t)$, with length N, into a linear combination of N wavelet basis $\varphi_j(t)$, equation (18) [23].

$$T\ (t) = \sum_{j=1}^{J} \varphi_j(t) \tag{18}$$

The next step is to decompose the second time-series $S(t)$, with the same length of N, into the same wavelet basis $\varphi_j(t)$, equation (19) [23].

$$S(t) = \sum_{j=1}^{J} \alpha_j\ \varphi_j(t) \tag{19}$$

Where the coefficients α_j could be derived into equation (20) [23].

$$\alpha_j = \frac{\langle S(t), \varphi_j(t) \rangle}{\langle \varphi_j(t), \varphi_j(t) \rangle} \tag{20}$$

Where < > stands for inner product.

As in FT, the distance of these coefficients could show similarity between two time-series, as it is represented in equation (21) [23].

$$D_{DWT}(T(t), S(t)) = \sqrt{\sum_{j=1}^{J} (1 - \alpha_j)^2} \tag{21}$$

If we consider all set of basis J=N, the result would be the same as the Euclidean distance. The most important feature of this method is to reduce noised data and to reduce unnecessary parts of the signal [23].

Similarly to [20], this thesis set all signals' length to N=1024 and J=4 to achieve 92% accuracy in the approximation.

3. METHODS AND EXPERIMENTS

As mentioned before, measuring similarity within time-series plays an important role in finding a pattern, enabling prediction and knowledge discovery.

Since clinical signals are random processes with non-stationary characteristics and each individual has its own, we can consider electrocardiograms like fingerprints where it is literally impossible to achieve similarity of 1, this is 100%. So, in this thesis, we are interested in observing how different similarity measurements methods performs between two time-series varies. With this we can make a statistical study in order to know which methods and ECG leads (below synthetically said leads) have the best performance when it comes to measuring similarity.

A primary experiment was made to know which are the best leads and similarity measures when we are measuring similarity between two time-series. On this experiment, we only took in consideration a cohort with patients with the same diagnosis, gender and age range.

After knowing that, we took a second experiment to find a pattern for a specific cardiac pathology. On this experiment, we have used a cohort where patients have different diagnosis with DM as reference.

In both experiments we are not measuring similarity between whole time-series, but with specific cardiac cycles of both series. Comparing whole ECG signals would result in erroneous results, since in thirty seconds of the time-series the number of cardiac cycles of each patient is variable.

3.1. IMPLEMETATION OF SIMILARITY MEASURING METHODS

To apply similarity measuring methods we need pairs of time-series, where one is the template and the other one is the one we want to measure the similarity with.

The time-series data were collected from the public data base PhysioNet [13]. The similarity measuring methods considered were the ones described in section 2.4.

3.2. DATA ACQUISITION

All data used in this thesis were collected from PhysioNet database [13]. This platform offers free web access to a large amount of biomedical data, many of them including clinical annotations.

In both experiments described in the next sections, the biomedical signals selected were ECGs collected from **The PTB Diagnostic** data base and only thirty seconds of that data was considered, which contains 549 records from 290 subjects (aged 17 to 87, mean 57.2; 209 men, mean age 55.5, and 81 women, mean age 61.6 with different heart diseases. PTB is an abbreviation for Physikalisch-Technische Bundesanstalt, the National Metrology Institute of Germany, which has provided this digitized ECGs for research. The sampling frequency in this database is 1000 Hz [13].

Both experiments required specifically developed software programs, which were implemented using Matlab software [27].

3.3. PRE-PROCESSING

In the real life, all the data collected from devices and sensors are subject to different kinds of noise and artefacts. The first and the most important step is to overcome this issue by performing some pre-processing, which includes noise filtering, normalization, transformations, feature extraction and data selection. Increasing the quality of the data will greatly reduce the probability of misleading results. The noise filtering can be handled by using digital filters or wavelet thresholding. By performing a normalization of the data, all values are adjusted in a common scale into the range [0, 1], this process is also called unity-based normalization which is presented in equation (22).

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{22}$$

Another pre-processing method is the removal of vertical offsets, which is described in equation (23).

$$X' = X - \bar{X} \tag{23}$$

Where $\bar{X}$ is the mean value of the signal.

Another issue to take into consideration is the scaling difference between time-series. In this thesis, we are measuring similarity between ECG's signals whose range of amplitude values varies widely. Since the similarity measuring methods are based on computing the point to point distance between both time-series these variations will produce misleading results. This problem can be fixed using linear transformation on the amplitudes.

Another important issue to consider is that, we must have time-series with the same duration to enable computation of their similarity. So, we must take into consideration the fact that each patient has different cardiac cycles duration. In this thesis, to overcome this problem the ECG's cardiac cycles were centred by QRS complex and the minimum common number of points was considered which means a loss of information.

3.4. EXPERIMENTS

3.4.1. EXPERIMENTS FOR FINDING THE MOST REPRESENTATIVE LEADS IN TERMS OF SIMILARITY VALUES WITHIN COHORTS

As it was mentioned previously on section 2.4, a similarity measuring method should be able to identify similarity between time-series despite the small variations that occur cycle to cycle.

The main goal of this experiment is to measure similarity between time-series from patients with same diagnosis, with this we will able to find which similarity measuring methods and leads are the most effective in identifying similarity among series with a certain pathology in common.

In order to increase the reliability of these experiments, the measurements between time-series were calculated for three different cardiac cycles. The cardiac cycles selected were the maximum and the minimum in terms of duration, plus the cardiac cycle between time-series that would result in less loss of information (closer to each other in terms of duration).

3.4.2. EXPERIMENT FOR FINDING A PATTERN ON DM PATIENTS

After identifying the best leads and similarity measuring methods our goal is to find a pattern of a cardiac disease in DM patients, to do so, several comparisons were made.

Firstly, a performance reference was needed. It is known that each cardiac cycle for a specific patient may vary in form and length, so it was required to know what value of performance (in this case, similarity) would represent the best similarity.

So, we started by computing the similarity between two ECG signals collected from the same patient (this patient has Myocardial infarction and *diabetes mellitus*). By considering measuring similarity between cardiac cycles of the same ECG record of a patient we were aiming to achieve a similarity value close to 1.

The first step was to measure the similarity between two ECG cardiac signals from the same individual collected with two weeks of difference (this patient presented myocardial infarction and *diabetes mellitus*), to find our upper bound.

The second step was to measure the similarity between two ECG signals from different patients but with the same diagnosis (these patients have Myocardial infarction and *diabetes mellitus*).

The third step was to measure the similarity between two ECG cardiac signals from different patients and different diagnosis (this cohort included patients with Myocardial infarction and *diabetes mellitus* in common but with additional different pathologies).

Lastly, the similarity between two ECG cardiac signals was computed between a healthy individual and a patient (this patient has Myocardial infarction and *diabetes mellitus*), and we hypothesised that this would determine the lower bound of the similarity performance range.

In order to increase the reliability of these experiments, the measurements between time-series were calculated with three different types of cardiac cycles' lengths. The cardiac cycles selected were the ones presenting the maximum and the minimum in terms of duration, plus the cardiac cycle length which would result in less loss of information (closer to each other in terms of duration). To be noticed that his procedure was not applied to the above mentioned first experiment.

4. RESULTS AND ANALYSIS

4.1. CASE-STUDIES

The Physionet [13] data considered in this study is listed in Table 4-1 where the name of the database record is specified as well as the characterization of the patients' information.

Number	**Gender**	**Age**	**ECG date**	**Diagnosis**	**Smoker**	**Blood pressure**
S0004	Female	79	14/08/1990	myocardial infarction *diabetes mellitus*	NO	ND
S0010[1]	Female	81	01/10/1990	myocardial infarction *diabetes mellitus*	NO	140/80 mmHg
S0014[2]	Female	81	17/10/1990	myocardial infarction *diabetes mellitus*	NO	140/80 mmHg
S0045	Female	71	14/11/1990	myocardial infarction *diabetes mellitus* renal insufficiency	YES	130/80 mmHg
S0052	Male	63	17/11/1990	myocardial infarction *diabetes mellitus* hyperuricemia	NO	120/70 mmHg
S0088	Female	74	03/01/1991	myocardial infarction *diabetes mellitus*	NO	160/90 mmHg
S0227	Male	59	18/09/1991	myocardial infarction, *diabetes mellitus* arterial hypertension	YES	120/60 mmHg
S0303	Female	32	24/06/1992	Healthy Control	ND	ND
S0311	Female	69	21/07/1992	Healthy Control	ND	ND
S0462[3]	Female	25	17/10/1996	Healthy Control	ND	ND

Table 4-1 Patients' information (ND – no information available).

The records employed were gathered into different cohorts, as described in Table 4.2.

1 - This is our template signal for DM patients.

2 - S0010 and S0014 are the same individual

3 - This is our template signal for Healthy Controls.

Cohort	Characteristics	Sub-division of Cohorts	Patients
1	Same Patient	1.1	S0010
		1.2	S0014
2	Different Patients with the same diagnosis	2.1	S0088
		2.2	S0004
3	Different Patients with different diagnosis	3.1	S0052
		3.2	S0045
		3.3	S0227
4	Healthy controls	4.1	S0462
		4.2	S0303
		4.3	S0311

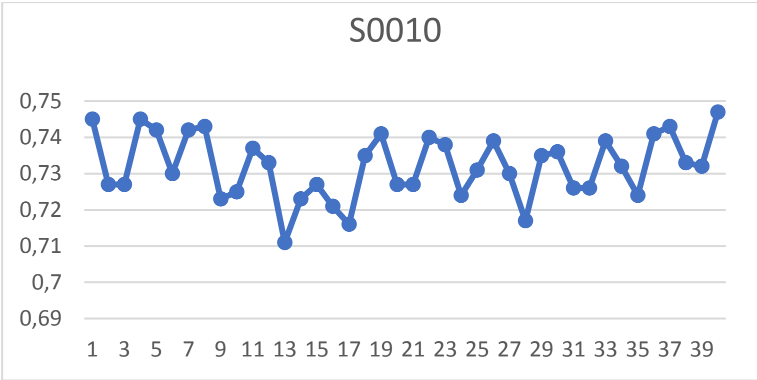
Table 4-2 Cohorts' information.

4.2. EXPERIMENTS FOR FINDING THE MOST REPRESENTATIVE LEADS IN TERMS OF SIMILARITY VALUES WITHIN COHORTS

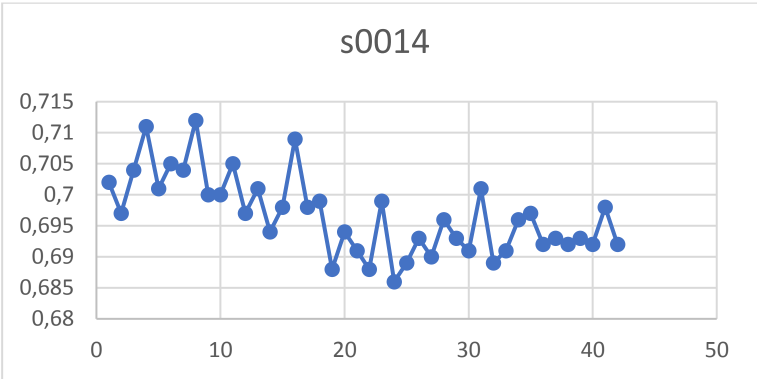
As mentioned in section 3.4.1, similarity measurements of ECG cardiac cycles of patients with the same diagnosis were tested, this is, an evaluation of the most adequate range of similarity performance to be considered in each experiment and the evaluation of similarity measurements between time-series of cardiac cycles of patients with the same diagnosis was performed for each ECG lead. The through description of the experiments is below presented, being graphically exemplified only for some cases, due to the large amount of information available. The table with similarity measurements results may be found in Appendix.

4.2.1. SIMILARITY MEASUREMENTS BETWEEN THE SAME PATIENT

In this experiment signals s0010 and s0014 were collected from the same individual but the ECG signal from s0014 was collected two weeks after s0010. The first step is to identify the cardiac cycles of both ECG signals. In the Figure 4.1, it is represented the cardiac cycles of patient s0010 (a) and patient s0014 (b).



(a)



(b)

Figure 4.1 The cardiac cycles of (a) s0010 patient (b) s0014 patient, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0010 the 13th cardiac cycle was the minimum cardiac cycle while the maximum one was the 40th cardiac cycle, so the HRV is 22ms. For patient s0014 the minimum was found for the 24th cardiac cycle and the maximum for the 8th, thus the HRV is 13ms.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 8th cardiac cycle of the patient s0014. It will be also compared the 13th cardiac cycle of the patient s0010 with the 24th cardiac cycle of the patient s0014. The last comparison will be between the 13th cardiac cycle of the patient s0010 and the 8th cardiac cycle of the patient s0014, this will result in losing only one data point of information.

4.2.1.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with longer data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity. So, the following graphs (Table 4.2) show the comparison between the 40th cardiac cycle of the s0010 patient with the 8th cardiac cycle of the patient s0014.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	30	94	4.2
L2	S_{CC}	38	86	4.3
L3	S_{CC}	16	94	4.4
V1	S_{CC}	24	86	4.5
V2	S_{CC}	23	88.	4.6
V3	S_{CC}	10	93	4.7
V4	S_{CC}	7	92	4.8
V5	S_{CC}	34	90	4.9
V6	S_{CC}	30	81	4.10
Vx	S_{CC}	14	89	4.11
Vy	S_{CC}	51	74	4.12
Vz	S_{CC}	44	86	4.13
aVF	S_{CC}	4	90	4.14
aVL	S_{CC}	12	96	4.15
aVR	S_{CC}	23	80	4.16

Table 4-3 Similarity between the 40th cardiac cycle of the s0010 patient with the 8th cardiac cycle of the patient s0014.

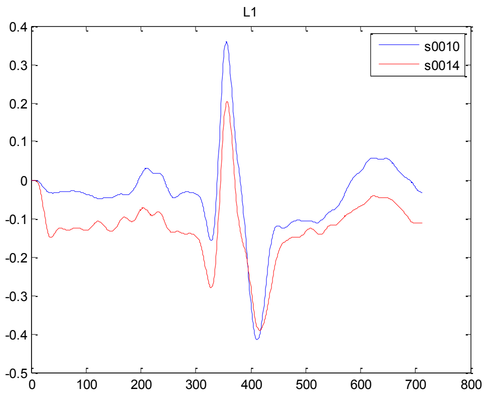


Figure 4.2 L1 lead.

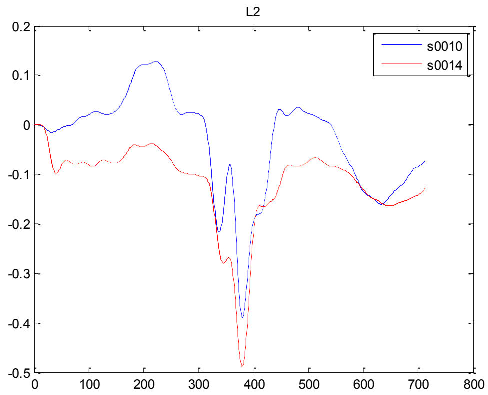


Figure 4.3 L2 lead.

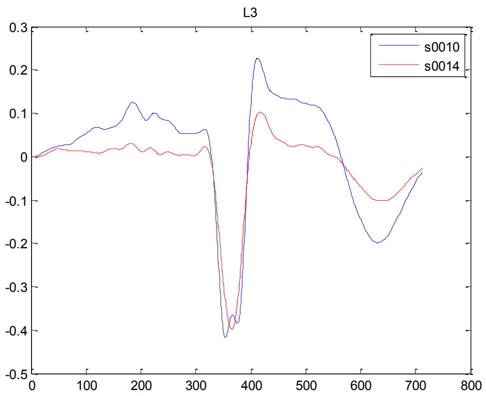


Figure 4.4 - L3 lead.

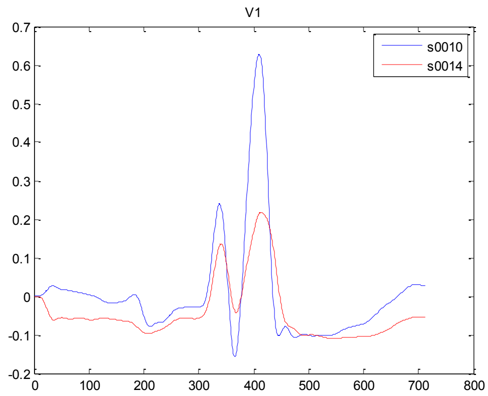


Figure 4.5 - V1 lead.

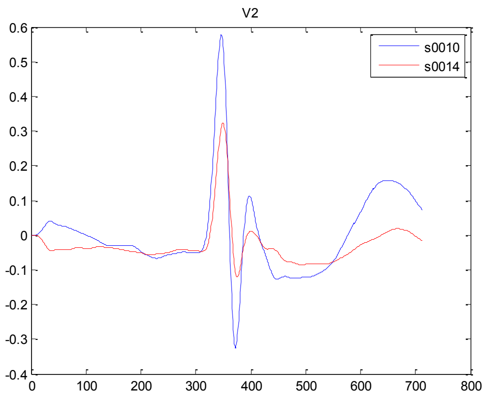


Figure 4.6 V2 lead.

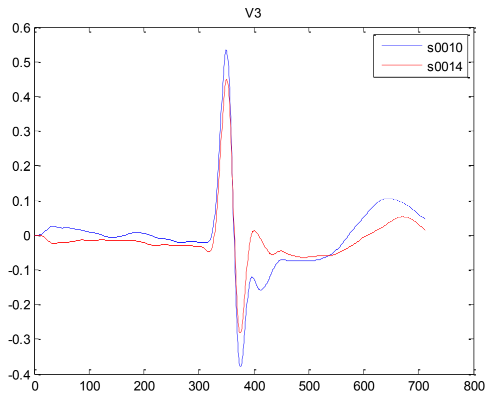


Figure 4.7 - V3 lead.

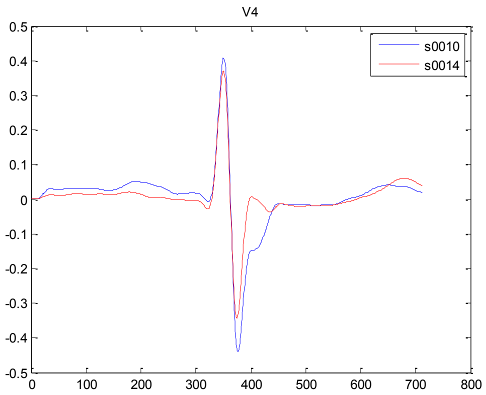


Figure 4.8 V4 lead.

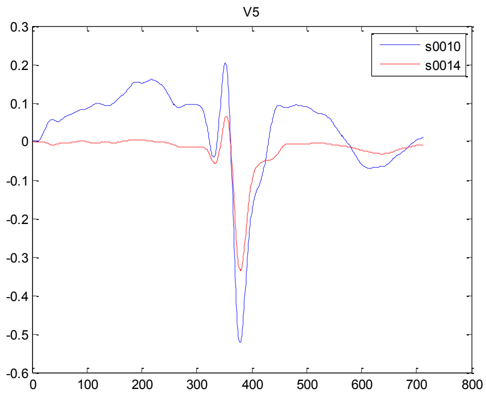


Figure 4.9 V5 lead.

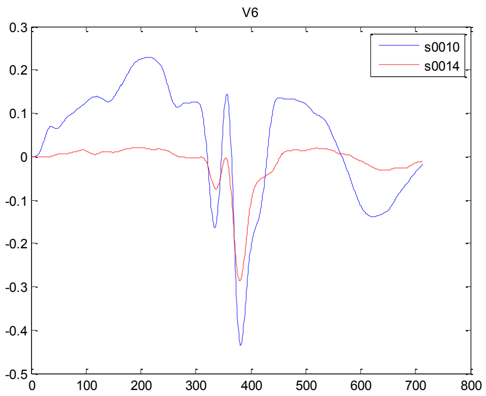


Figure 4.10 V6 lead.

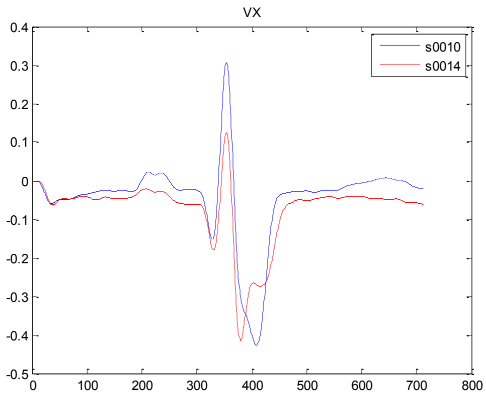


Figure 4.11 VX lead.

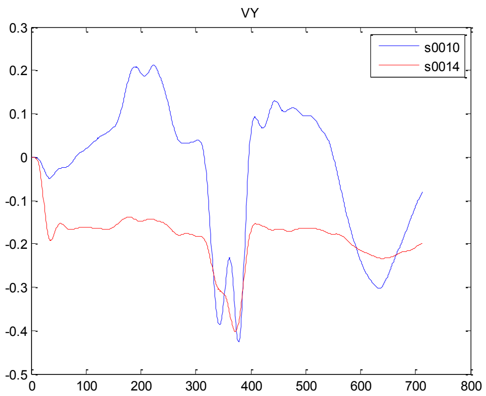


Figure 4.12 VY lead.

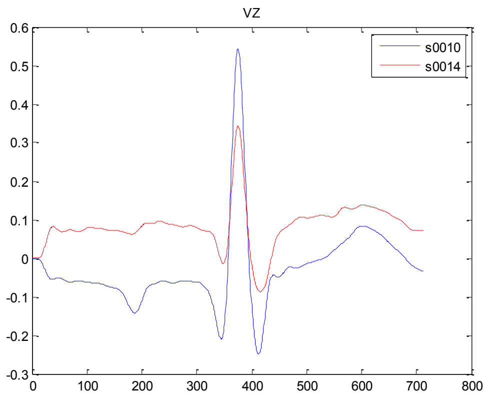


Figure 4.13 VZ lead.

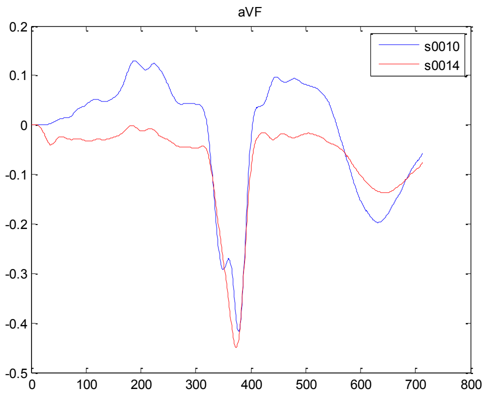


Figure 4.14 aVF lead.

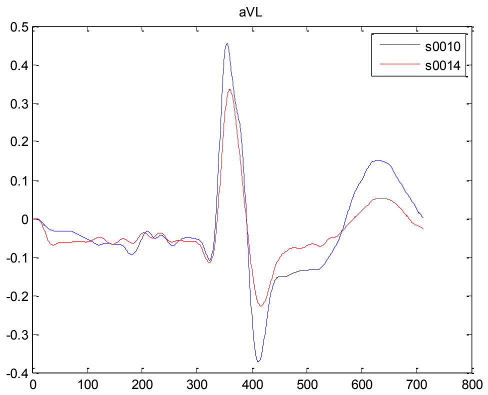


Figure 4.15 aVL lead.

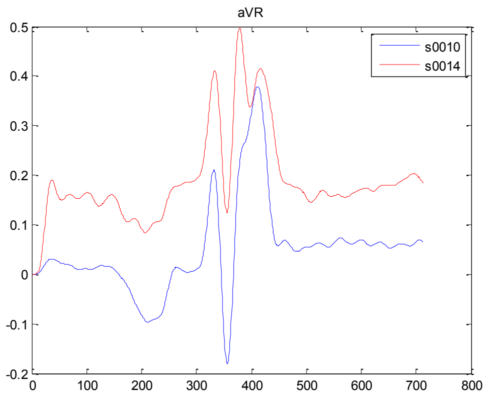


Figure 4.16 aVR lead.

4.2.1.2. Analysis

In this experiment, it was concluded that Pearson's correlation coefficient outperformed other similarity measurement methods for all leads.

It was verified that the highest similarity among time-series was obtained progressively decreasing in the following leads: L1, L3, aVL, V3 and V4.

Since the Wavelet Transform KLT based method has been appointed in previous researches as being an accurate similarity method, we identified that the sequence of the best performed leads, from highest to lower was: VX, aVF, V4, V3, L1 and aVL. The performance of this method was good enough to be taken into consideration.

4.2.2. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT PATIENTS WITH THE SAME DIAGNOSIS - I

In this experiment, signals s0010 and s0088 were collected from different individuals with the same diagnosis, this is, besides having *diabetes mellitus they were diagnosed with* myocardial infarction. In Figure 4.17 is represented the cardiac cycles of patient s0088.

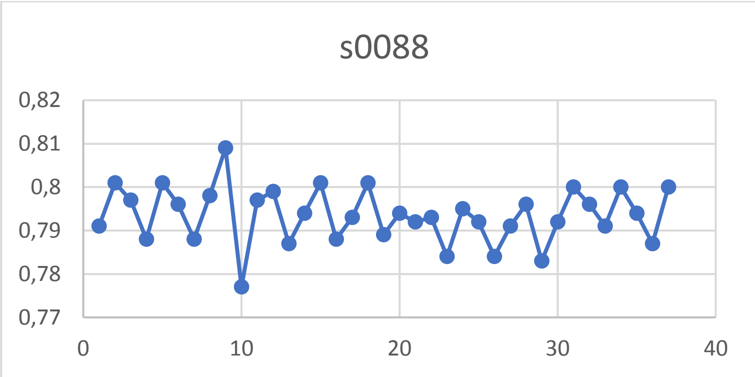


Figure 4.17 The cardiac cycles of s0088 patient, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0088 the 10^{th} cardiac cycle was the minimum cardiac cycle while the maximum one was the 9^{th} cardiac cycle, which represents a HRV of 32ms.

So, to measure similarity in these patients it will be compared the 40^{th} cardiac cycle of the s0010 patient with the 9^{th} cardiac cycle of patient s0088. It will be also compared the 13^{th} cardiac cycle of the patient s0010 with the 10^{th} cardiac cycle of the patient s0088. The last comparison will be between the 40^{th} cardiac cycle of the patient s0010 and the 10^{th} cardiac cycle of the patient s0088, this will result in losing thirty data points of information.

4.2.2.1. Results

Among these three comparisons it was observed that whenever longer data lengths were used for comparing the time-series the better the results were attained, independently of the ECG lead under study. So, only the best results will be presented below for the sake of thesis' simplicity.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{WT}	6	78	4.18
L2	S_{WT}	24	47	4.19
L3	S_{CC}	14	84	4.20
V1	S_{WT}	31	41	4.21
V2	S_{CC}	2	28	4.22
V3	S_{CC}	14	51	4.23
V4	S_{CC}	13	58	4.24
V5	S_{CC}	12	55	4.25
V6	S_{WT}	16	43	4.26
Vx	S_{WT}	13	59	4.27
Vy	S_{CC}	6	45	4.28
Vz	S_{CC}	23	65	4.29
aVF	S_{CC}	16	74	4.30
aVL	S_{CC}	19	84	4.31
aVR	S_{WT}	52	79	4.32

Table 4-4 Similarity between the 40th cardiac cycle of the s0010 patient with the 9th cardiac cycle of the patient s0088.

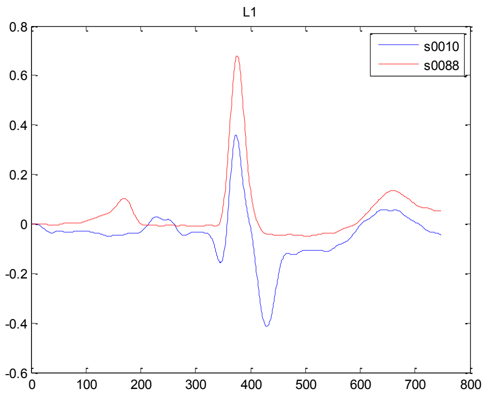


Figure 4.18 L1 lead.

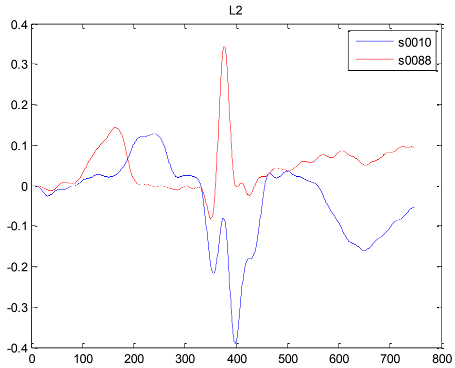


Figure 4.19 L2 lead.

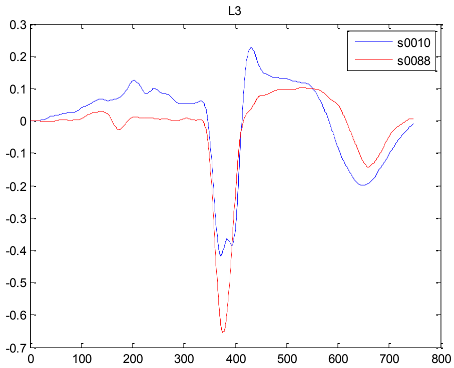


Figure 4.20 L3 lead.

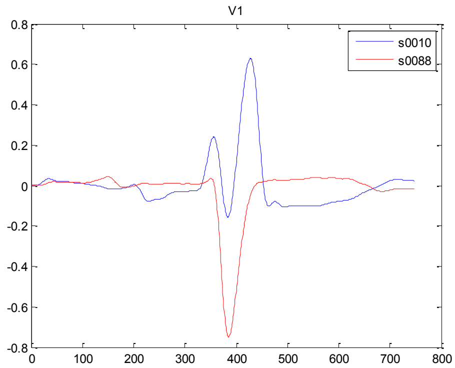


Figure 4.21 V1 lead.

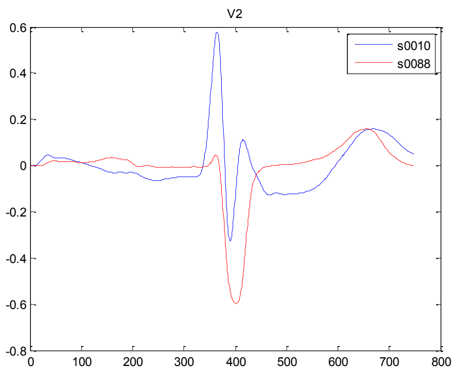


Figure 4.22 V2 lead.

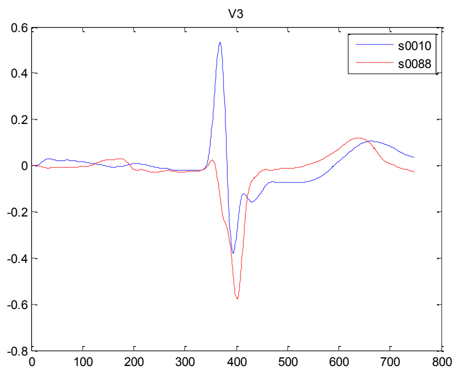


Figure 4.23 V3 lead.

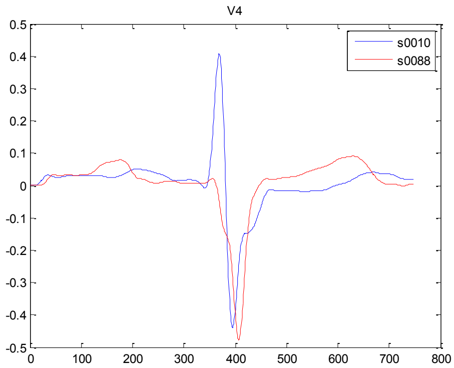


Figure 4.24 V4 lead.

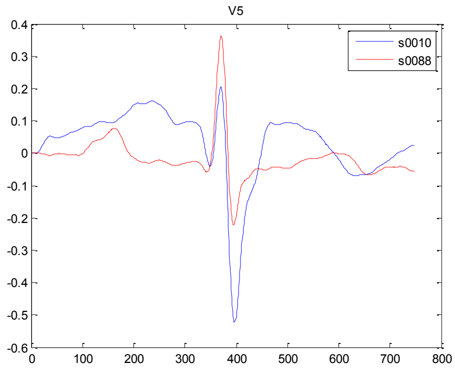


Figure 4.25 V5 lead.

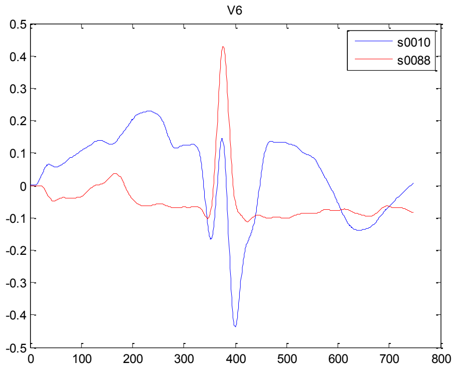


Figure 4.26 V6 lead.

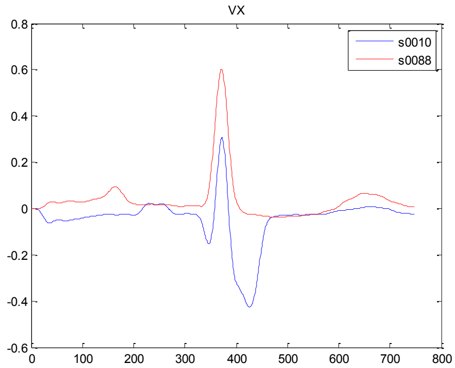


Figure 4.27 VX lead.

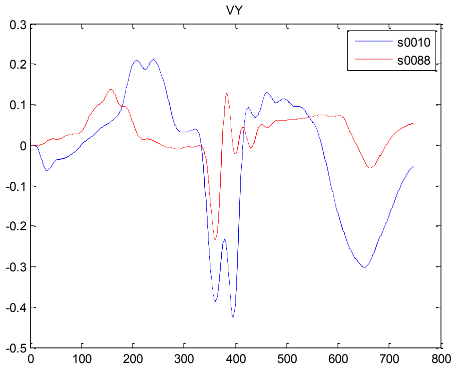


Figure 4.28 VY lead.

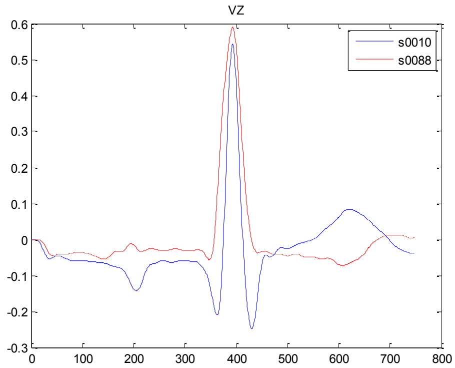


Figure 4.29 VZ lead.

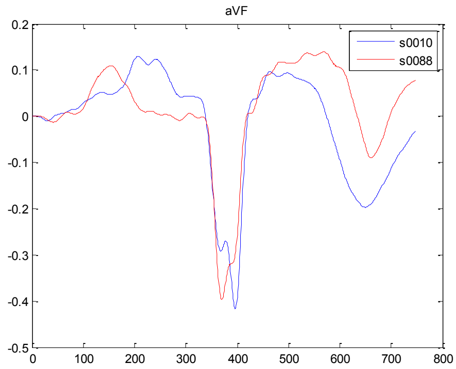


Figure 4.30 aVF lead.

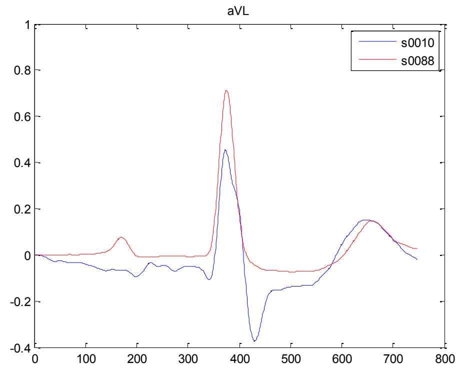


Figure 4.31 aVL lead.

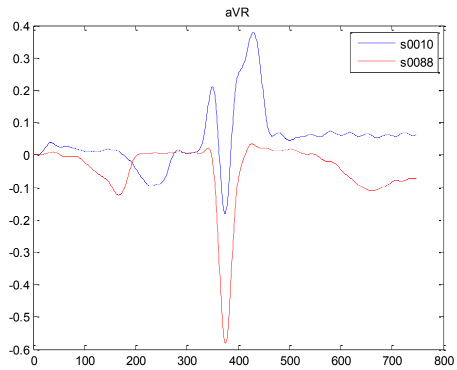


Figure 4.32 aVR lead.

4.2.2.2. Analysis

In this experiment, we can conclude that Pearson's correlation coefficient outperformed other similarities measurement methods in nine out of fifteen leads. However, unlikely the previous experiment, the Wavelet Transform KLT based method's performances was not so far behind the Pearson's correlation coefficient method.

We verified that we have obtained the highest similarity among time-series in the following leads (from highest performance to lower): L3, aVL, aVF, L1, VZ and V4, these were the leads where Pearson's correlation coefficient performed the best.

To be mentioned that the leads where Wavelet Transform KLT based method performed the best, were the following: aVR, L1, aVL, VX, L3 and aVF.

4.2.3. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT PATIENTS WITH THE SAME DIAGNOSIS - II

In these experiments, the tested signals were collected from different patients with the same diagnosis as reported in last section, this is, myocardial infarction besides *diabetes mellitus, but now comparison was performed between patient's time-series s0010 with s0004*. In Figure 4.33, it is represented the cardiac cycles of patient s0004.

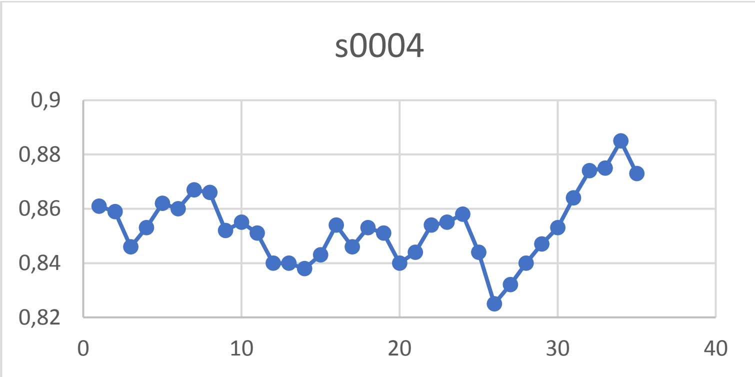


Figure 4.33 The cardiac cycles of s0004 patient, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0004 the 26th cardiac cycle was the minimum cardiac cycle while the maximum one was the 34th cardiac cycle, presenting a 19ms of HRV.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 34th cardiac cycle of the patient s0004. It will be also compared the 13th cardiac cycle of the patient s0010 with the 26th cardiac cycle of the patient s0004. The last comparison it will be between the 40th cardiac cycle of the patient s0010 and the 26th cardiac cycle of the patient s0004, this will result in losing seventy-eight data points of information.

4.2.3.1. Results

Following the same methodology as previously, only the best performed results will be show, in this case, only the comparison between the 40th cardiac cycle of the patient s0010 and the 26th cardiac cycle of the patient s0004 will be presented.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	28	59	4.34
L2	S_{WT}	45	78	4.35
L3	SM_{AH}	9	38	4.36
V1	S_{WT}	1	35	4.37
V2	S_{CC}	10	88	4.38
V3	S_{CC}	18	91	4.39
V4	S_{WT}	9	87	4.40
V5	S_{WT}	17	54	4.41
V6	S_{WT}	32	61	4.42
Vx	S_{WT}	10	64	4.43
Vy	SM_{AH}	2	35	4.44
Vz	S_{WT}	12	80	4.45
aVF	SM_{AH}	11	30	4.46
aVL	S_{CC}	9	63	4.47
aVR	S_{WT}	6	58	4.48

Table 4-5 Similarity between the 40^{th} cardiac cycle of the s0010 patient with the 26^{th} cardiac cycle of the patient s0004.

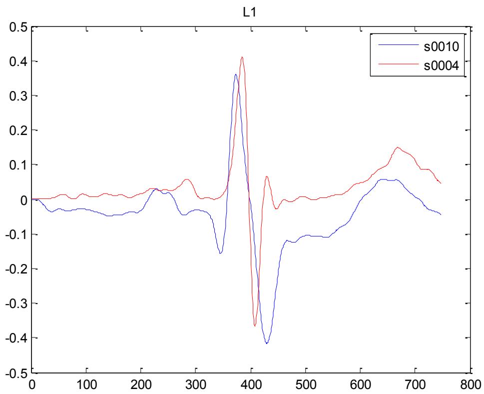


Figure 4.34 L1 lead.

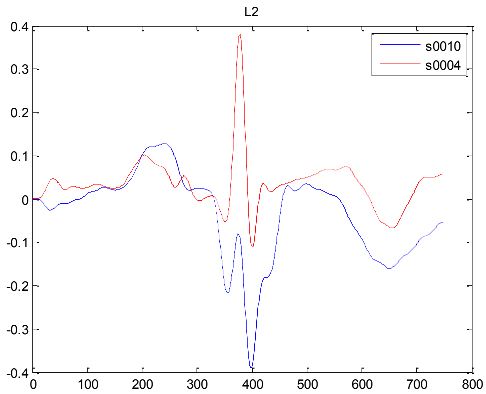


Figure 4.35 L2 lead.

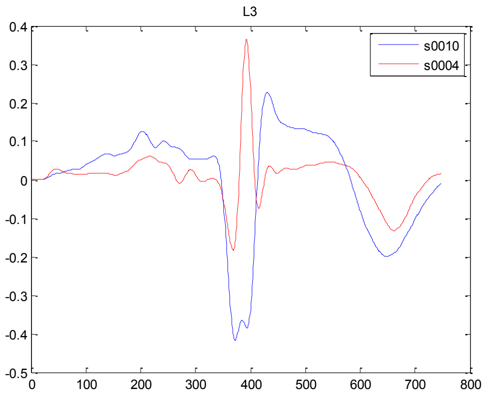


Figure 4.36 L3 lead.

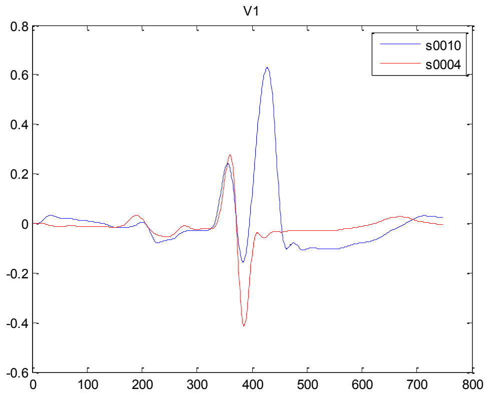


Figure 4.37 V1 lead.

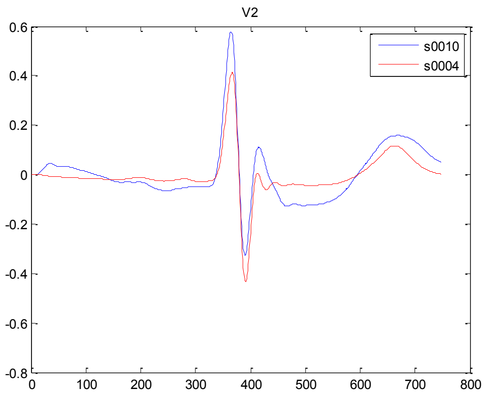


Figure 4.38 V2 lead.

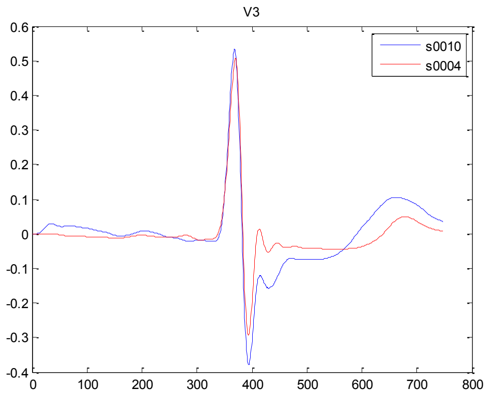


Figure 4.39 V3 lead.

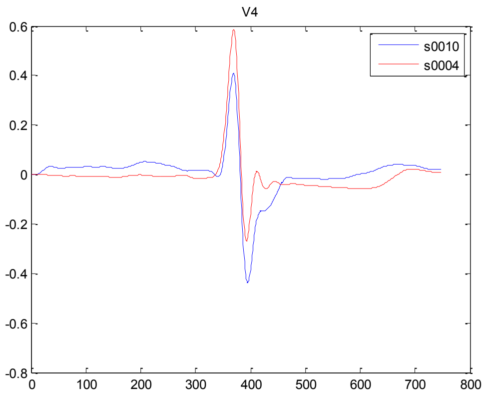


Figure 4.40 V4 lead.

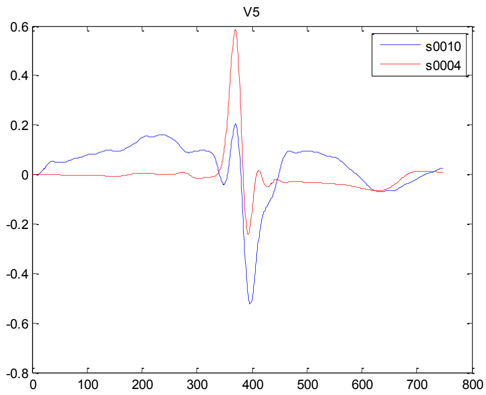


Figure 4.41 V5 lead.

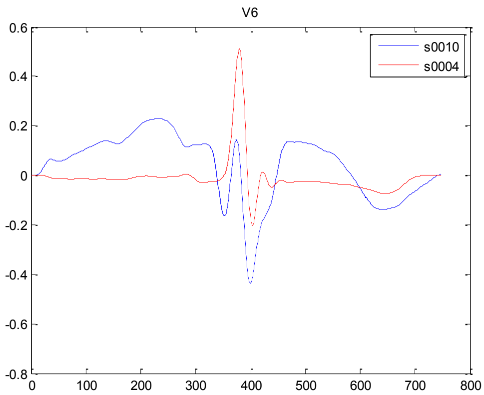


Figure 4.42 V6 lead.

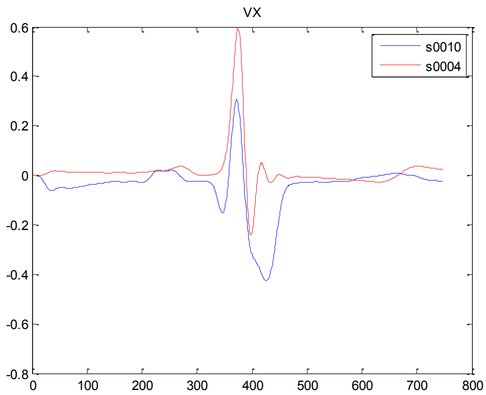


Figure 4.43 VX lead.

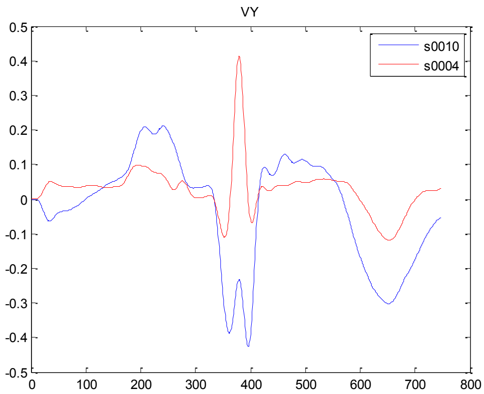


Figure 4.44 VY lead.

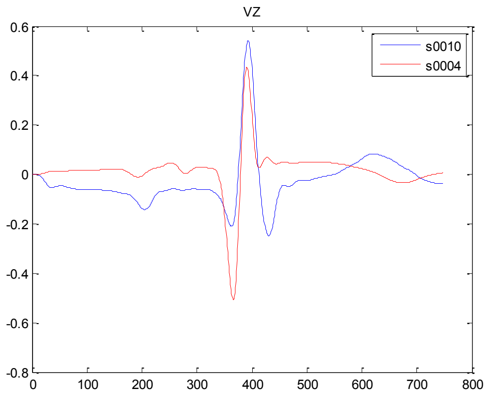


Figure 4.45 VZ lead.

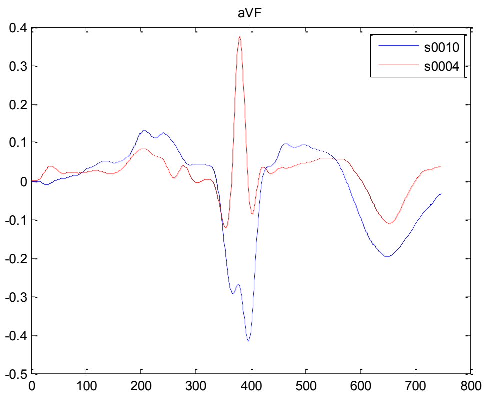


Figure 4.46 aVF lead.

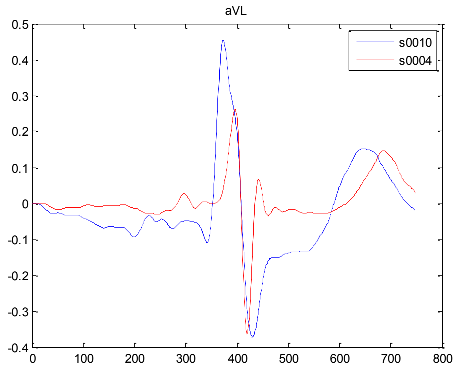


Figure 4.47 aVL lead.

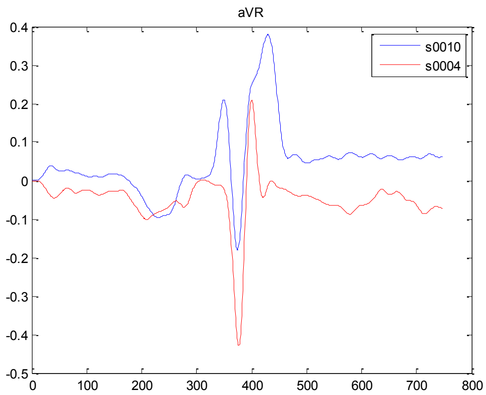


Figure 4.48 aVR lead.

4.2.3.2. Analysis

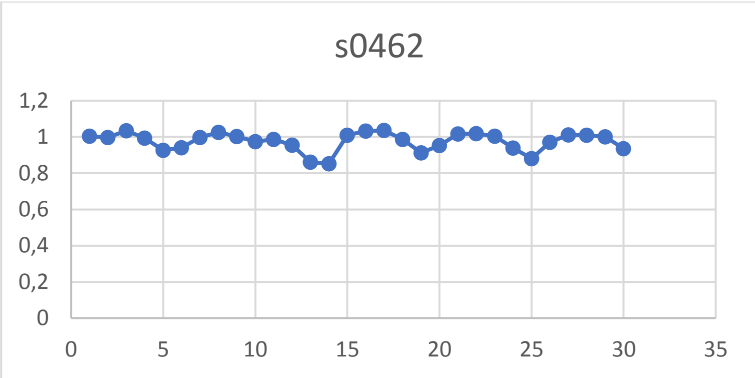
In this experiment, we can conclude that Pearson's correlation coefficient was slightly outperformed by Wavelet Transform KLT based method. It will be once again considered the leads where these two methods performed the best.

We verified that we have obtained the highest similarity among time-series in the following leads: V3, V2, V4, aVL, L1 and VZ, these were the leads where Pearson's correlation coefficient performed the best.

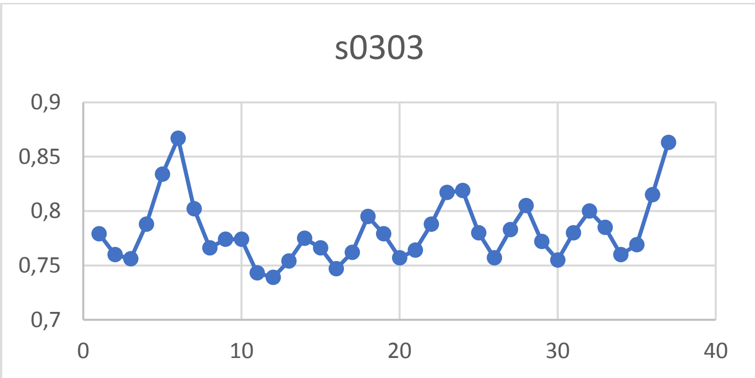
Lastly, the leads where Wavelet Transform KLT based method performed the best, were the following: V4, VZ, V2, L2, V3 and VX.

4.2.4. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT HEALTHY CONTROLS - I

In this experiment signals s0462 and s0303 were collected from distinct healthy individuals. In the Figure 4.49, it is represented the cardiac cycles of individual s0462 (a) and individual s0303 (b).



(a)



(b)

Figure 4.49 The cardiac cycles of (a) s0462 healthy control (b) s0303 healthy control, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for healthy control s0462 the 14th cardiac cycle was the minimum while the maximum was the 17th cardiac cycle, so the HRV of 157ms. For healthy control s0303 the minimum was found for the 12th cardiac cycle and the maximum for the 6th, thus the HRV is 65ms.

So, to measure similarity in these patients it will be compared the 17th cardiac cycle of the s0462 patient with the 6th cardiac cycle of the patient s0303. It will be also compared the 14th cardiac cycle of the patient s0462 with the 12th cardiac cycle of the patient s0303. The last comparison it will be between the 13th cardiac cycle of the patient s0462 and the 37th cardiac cycle of the patient s0303, this will result in losing only two data points of information.

4.2.4.1. Results

The best performed results obtained for the time-series related to the cardiac cycles with closer data lengths are expressed in Table 4-6 and figures 4-50 to 4-64..

Lead	**Best performed method**	**Exceeding the other methods performance by (%)**	**Similarity measure achieved (%)**	**Figure**
L1	S_{CC}	45	77	4.50
L2	S_{CC}	36	81	4.51
L3	S_{CC}	18	62	4.52
V1	S_{CC}	15	82	4.53
V2	S_{CC}	18	72	4.54
V3	S_{CC}	3	60	4.55
V4	S_{CC}	26	76	4.56
V5	S_{CC}	20	86	4.57
V6	S_{CC}	23	87	4.58
Vx	S_{CC}	23	84	4.59
Vy	S_{CC}	33	76	4.60
Vz	S_{CC}	21	94	4.61
aVF	S_{WT}	3	80	4.62
aVL	S_{WT}	15	39	4.63
aVR	S_{CC}	39	82	4.64

Table 4-6 Similarity between the 13th cardiac cycle of the s0462 patient with the 37th cardiac cycle of the patient s0303.

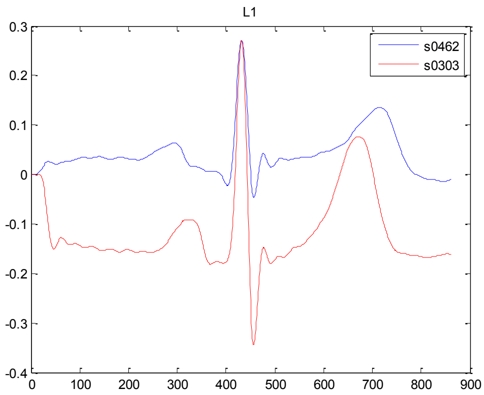


Figure 4.50 L1 lead.

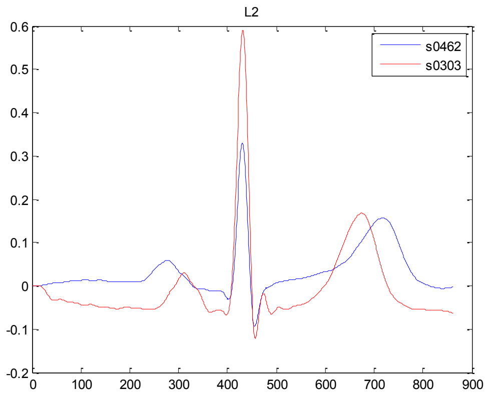


Figure 4.51 L2 lead.

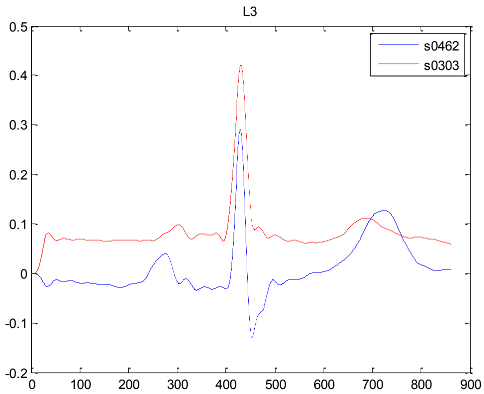


Figure 4.52 L3 lead.

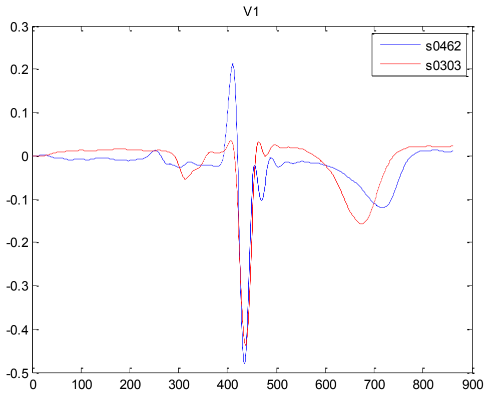


Figure 4.53 V1 lead.

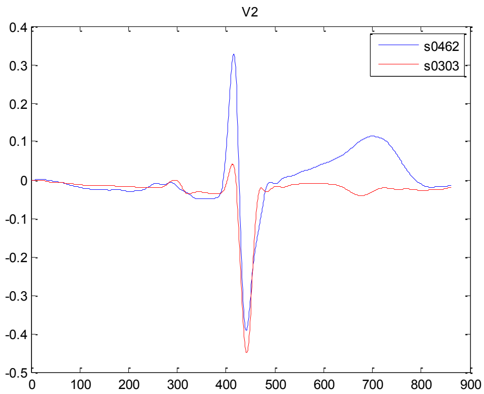


Figure 4.54 V2 lead.

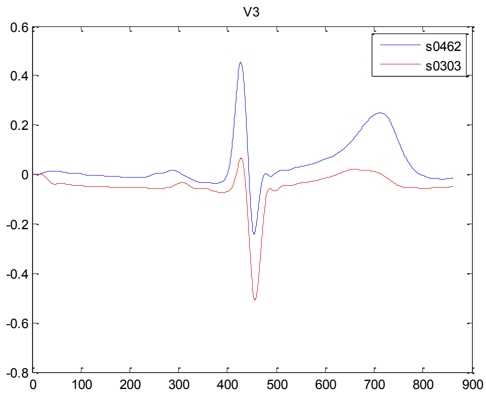


Figure 4.55 V3 lead.

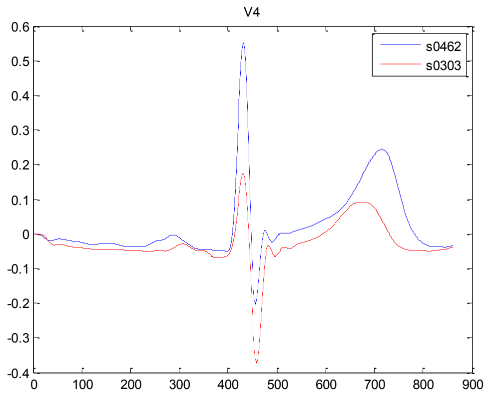


Figure 4.56 V4 lead.

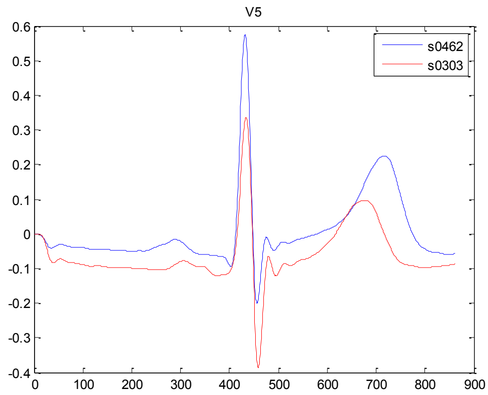


Figure 4.57 V5 lead.

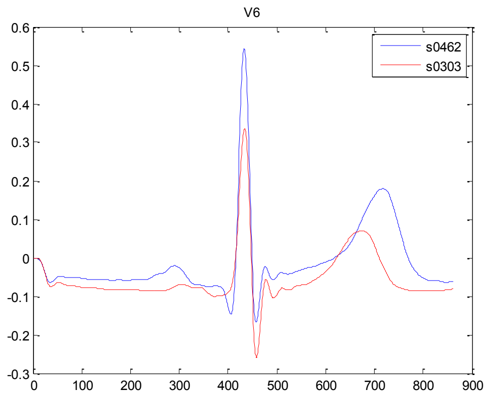


Figure 4.58 V6 lead.

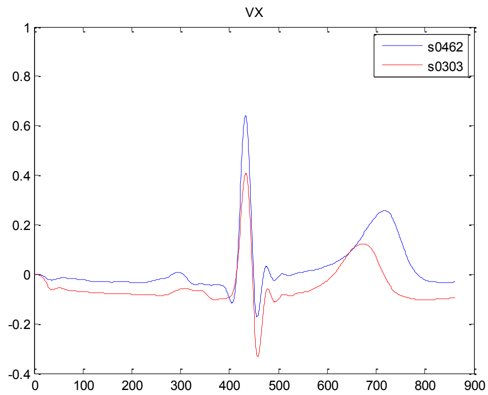


Figure 4.59 VX lead.

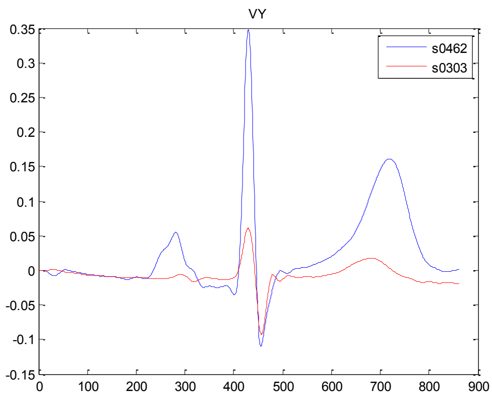


Figure 4.60 VY lead.

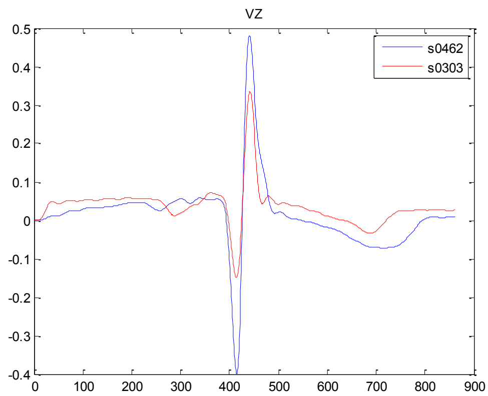


Figure 4.61 VZ lead.

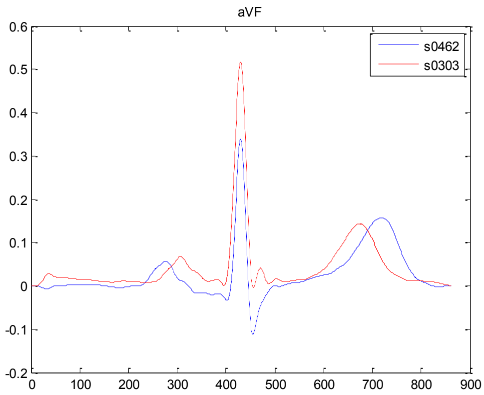


Figure 4.62 aVF lead.

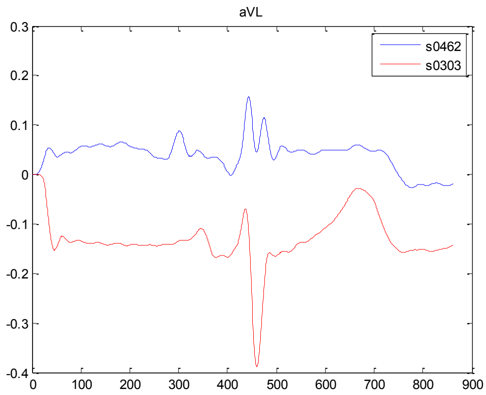


Figure 4.63 aVL lead.

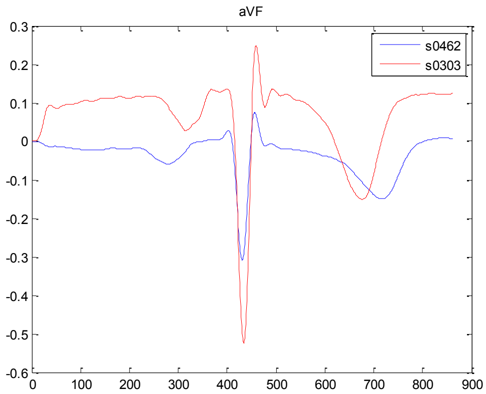


Figure 4.64 aVR lead.

4.2.4.2. Analysis

In this comparison, it can be concluded that Pearson's correlation coefficient outperformed other similarities measurement methods in twelve out of fifteen leads.

We verified that we have obtained the highest similarity among time-series in the following leads: VZ, V6, V5, VX, aVR and V1, these were the leads where Pearson's correlation coefficient performed the best.

Lastly, the leads where Wavelet Transform KLT based method performed the best, are the following: aVF, V5, VX, V3, V6 and V4.

4.2.5. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT HEALTHY CONTROLS - II

In this measurement, the signals that were tested, were collected from different healthy individuals. In Figure 4.65, it is represented the cardiac cycles of patient s0311.

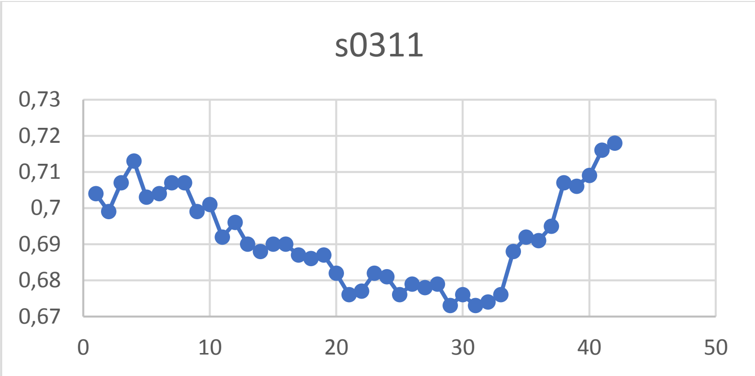


Figure 4.65 The cardiac cycles of healthy control s0311, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for healthy control s0311 the 29th cardiac cycle was the minimum cardiac cycle while the maximum one was the 42th cardiac cycle. The corresponding HRV is 12ms.

So, to measure similarity in these individuals it will be compared the 17th cardiac cycle of the s0462 individual with the 42th cardiac cycle of the individual s0311. It will be also compared the 14th cardiac cycle of the individual s0462 with the 29th cardiac cycle of the individual s0311. The last comparison it will be between the 14th cardiac cycle of the individual s0462 and the 42th cardiac cycle of the individual s0311, this will result in losing one hundred and thirty-five data points of information.

4.2.5.1. Results

The results obtained for the best performed pairs of comparison were between the 14th cardiac cycle of the individual s0462 and the 42th cardiac cycle of the individual s0311. These are the results below presented.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	44	83	4.66
L2	S_{WT}	25	93	4.67
L3	S_{WT}	47	76	4.68
V1	S_{WT}	2	83	4.69
V2	S_{CC}	11	81	4.70
V3	S_{WT}	10	86	4.71
V4	S_{WT}	6	89	4.72
V5	S_{WT}	1	86	4.73
V6	S_{CC}	41	79	4.74
Vx	S_{CC}	30	85	4.75
Vy	S_{CC}	14	80	4.76
Vz	S_{CC}	7	82	4.77
aVF	SM_i	0,5	50	4.78
aVL	S_{CC}	59	78	4.79
aVR	S_{CC}	25	83	4.80

Table 4-7 Similarity between the 14th cardiac cycle of the s0462 patient with the 42th cardiac cycle of the patient s0311.

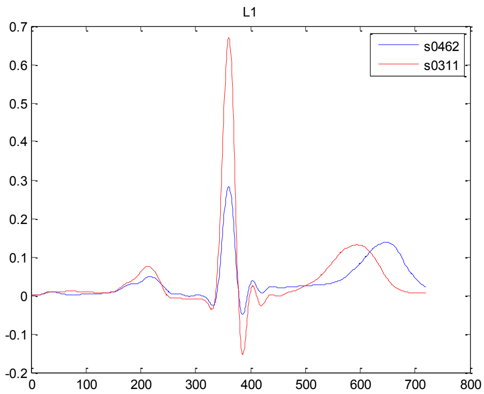


Figure 4.66 L1 lead.

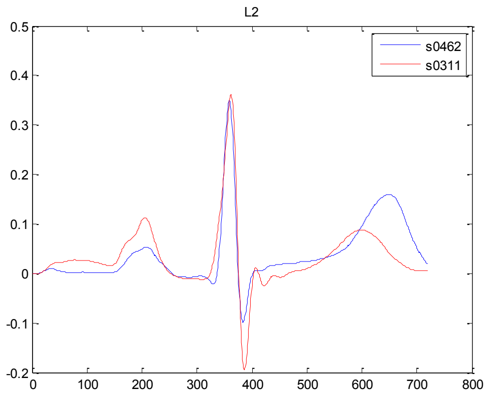


Figure 4.67 L2 lead.

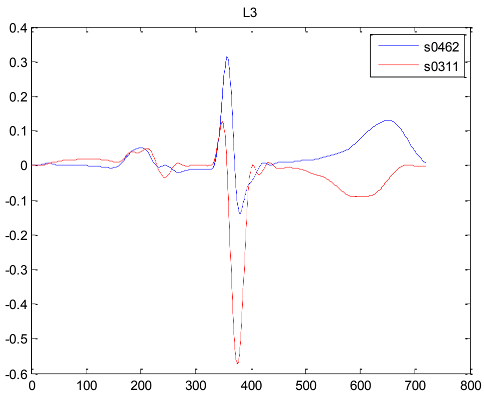


Figure 4.68 L3 lead.

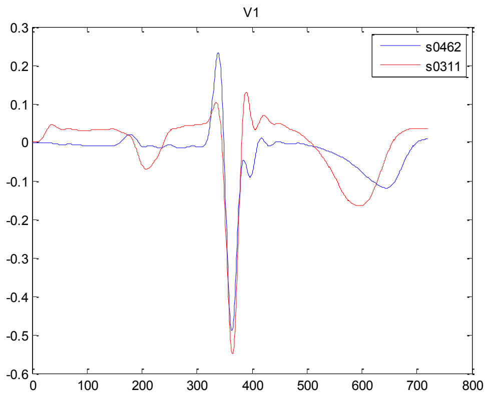


Figure 4.69 V1 lead.

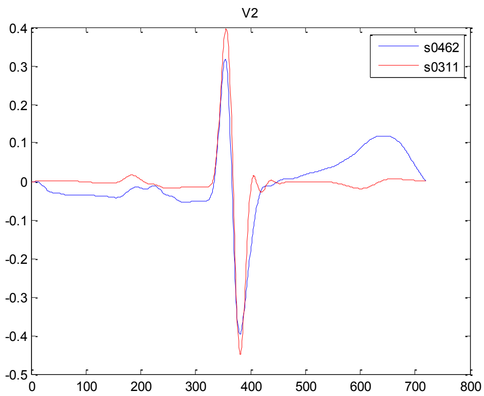


Figure 4.70 V2 lead.

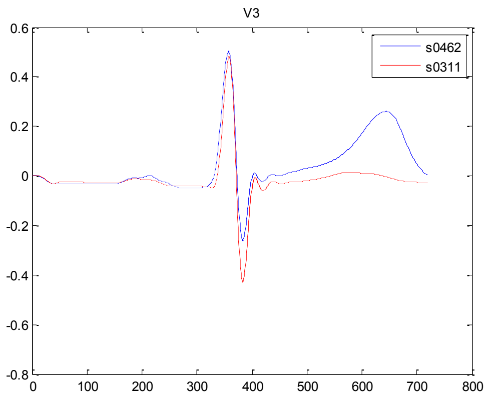


Figure 4.71 V3 lead.

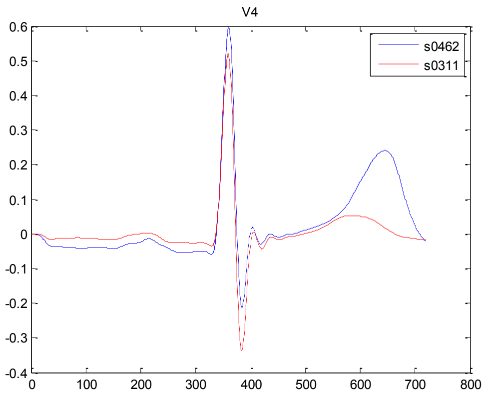


Figure 4.72 V4 lead.

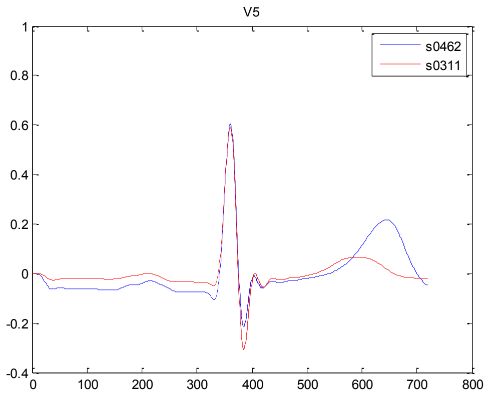


Figure 4.73 V5 lead.

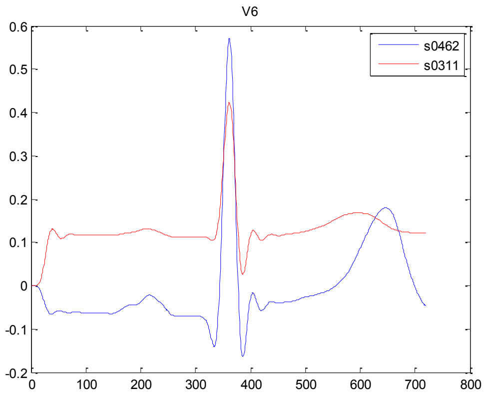


Figure 4.74 V6 lead.

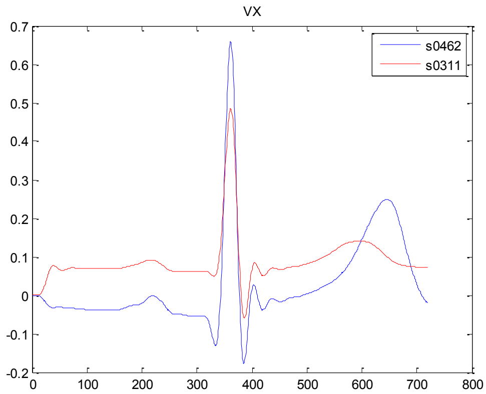


Figure 4.75 VX lead.

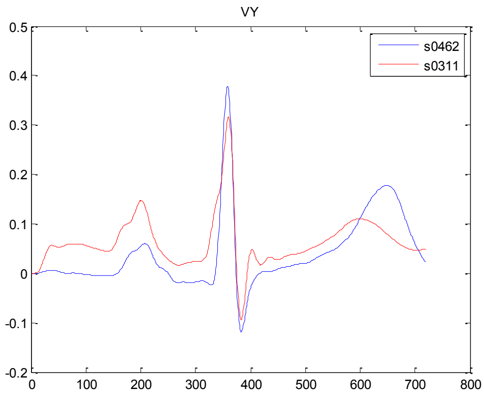


Figure 4.76 VY lead.

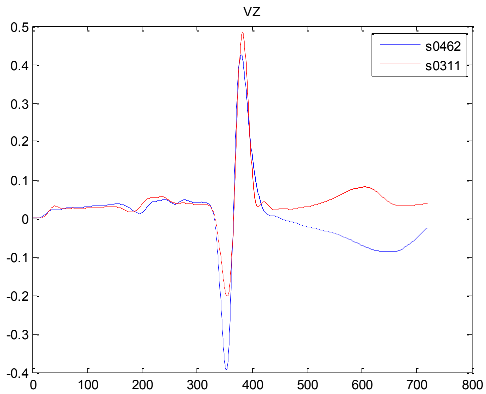


Figure 4.77 VZ lead.

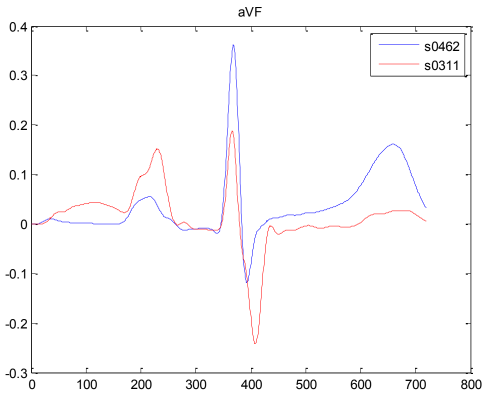


Figure 4.78 aVF lead.

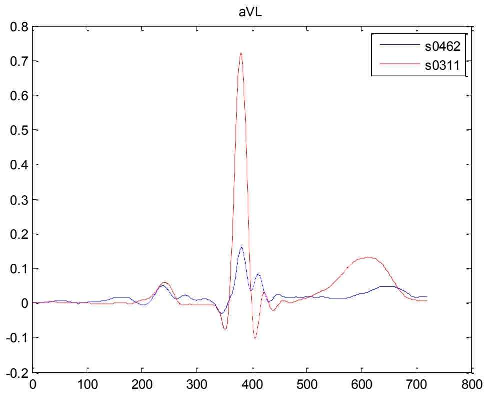


Figure 4.79 aVL lead.

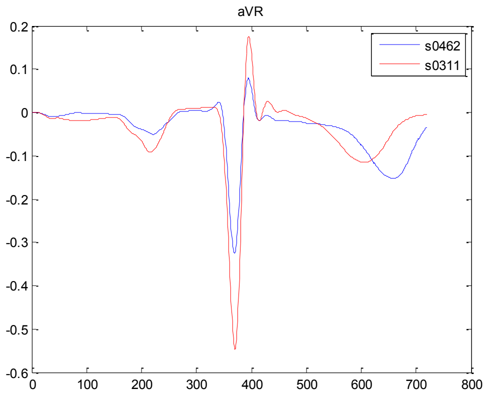


Figure 4.80 aVR lead.

4.2.5.2. Analysis

In this comparison, it can be concluded that Pearson's correlation coefficient outperformed other similarities measurement methods in eight out of fifteen leads.

We verified that we have obtained the highest similarity among time-series in the following leads: V5, VX, V4, aVR, L1 and VZ, these were the leads where Pearson's correlation coefficient performed the best.

Lastly, the leads where Wavelet Transform KLT based method performed the best, were the following: L2, V4, V5, V3, V1 and L3.

4.3. EXPERIMENT FOR FINDING A PATTERN ON DM PATIENTS WITH MYOCARDIAL INFARCTION

4.3.1. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT PATIENTS WITH DIFFERENT DIAGNOSIS - I

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range. In Figure 4.81, it is represented the cardiac cycles of patient s0052.

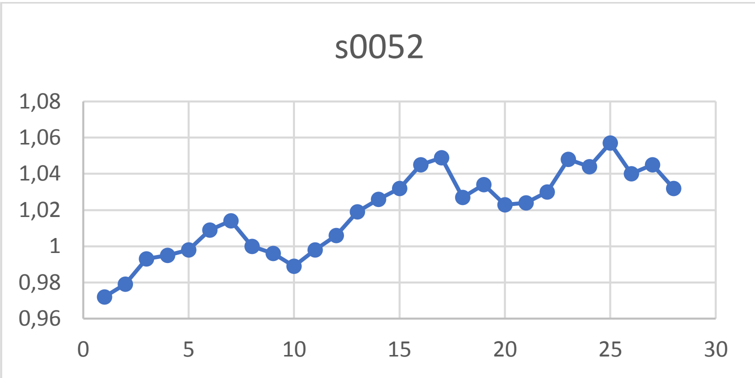


Figure 4.81 The cardiac cycles of patient s0052, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0052 the 1st cardiac cycle was the minimum cardiac cycle while the maximum one was the 25th cardiac cycle, so the HRV is 22ms.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 25th cardiac cycle of the patient s0052. It will be also compared the 13th cardiac cycle of the patient s0010 with the 1st cardiac cycle of the patient s0052. The last comparison it will be between the 40th cardiac cycle of the patient s0010 and the 1st cardiac cycle of the patient s0052, this will result in losing two hundred and twenty-five data points of information.

4.3.1.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with shorter data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{WT}	28	92	4.82
L2	S_{WT}	76	91	4.83
L3	S_{CC}	5	49	4.84
V1	S_{WT}	2	33	4.85
V2	S_{CC}	2	70	4.86
V3	S_{WT}	22	93	4.87
V4	S_{WT}	34	94	4.88
V5	S_{WT}	86	97	4.89
V6	S_{WT}	86	95	4.90
Vx	S_{WT}	38	87	4.91
Vy	S_{WT}	4	28	4.92
Vz	S_{WT}	3	49	4.93
aVF	S_{WT}	7	28	4.94
aVL	S_{CC}	11	65	4.95
aVR	S_{WT}	34	67	4.96

Table 4-8 Similarity between the 13th cardiac cycle of the s0010 patient with the 1st cardiac cycle of the patient s0052.

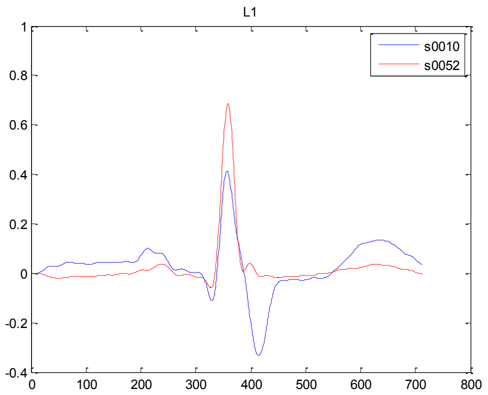


Figure 4.82 L1 lead.

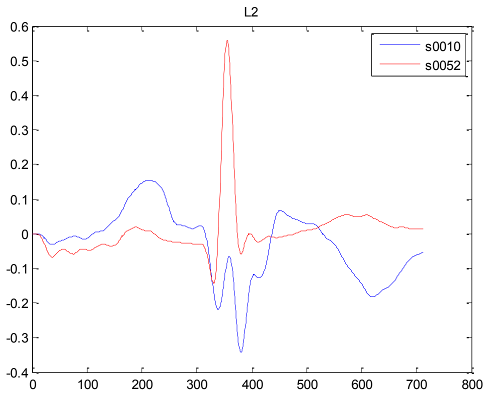


Figure 4.83 L2 lead.

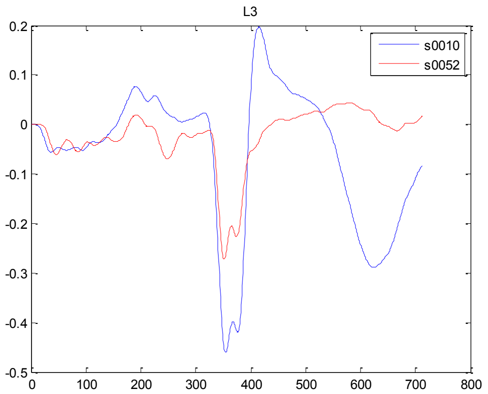


Figure 4.84 L3 lead.

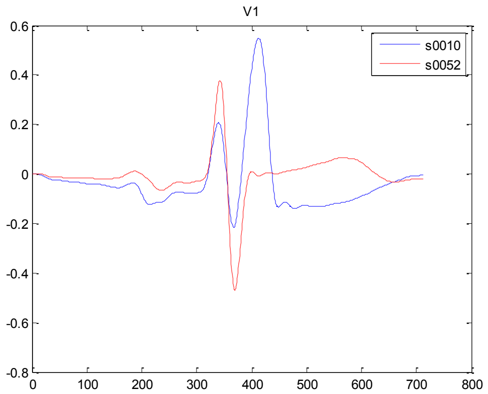


Figure 4.85 V1 lead.

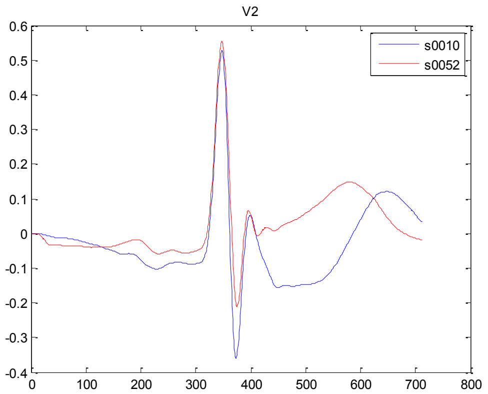


Figure 4.86 V2 lead.

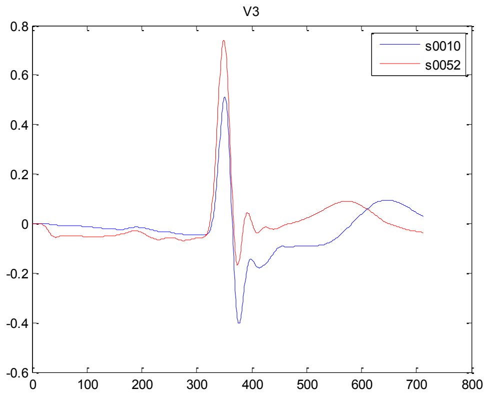


Figure 4.87 V3 lead.

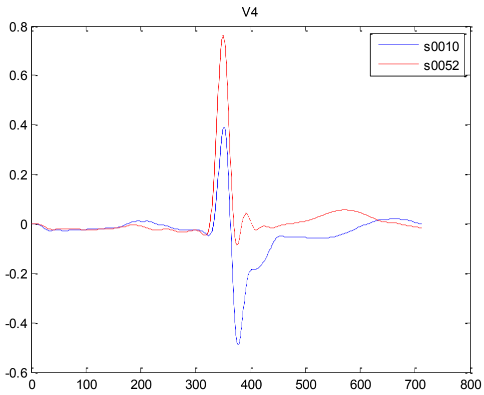


Figure 4.88 V4 lead.

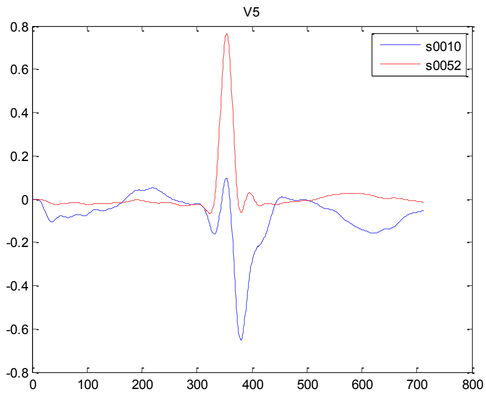


Figure 4.89 V5 lead.

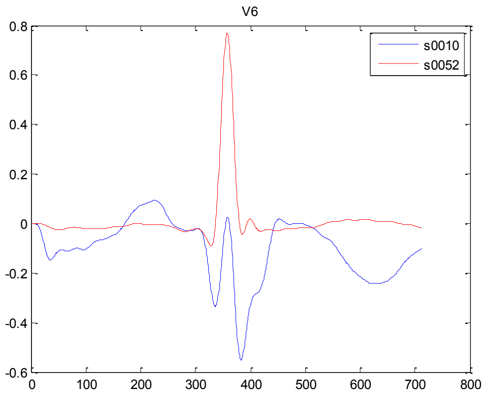


Figure 4.90 V6 lead.

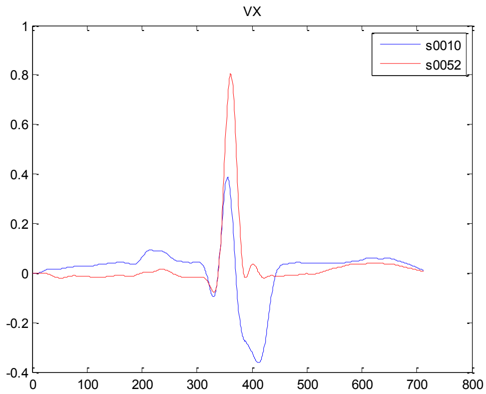


Figure 4.91 VX lead.

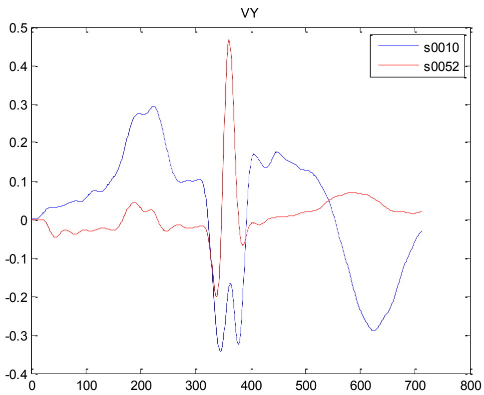


Figure 4.92 VY lead.

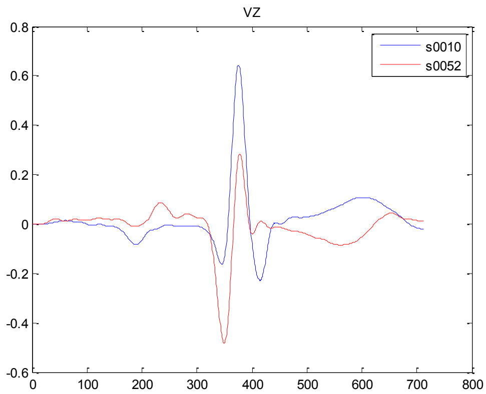


Figure 4.93 VZ lead.

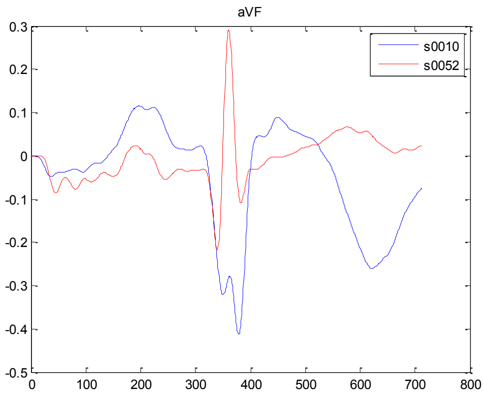


Figure 4.94 aVF lead.

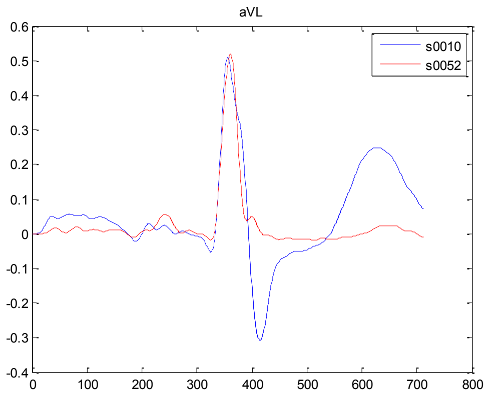


Figure 4.95 aVL lead.

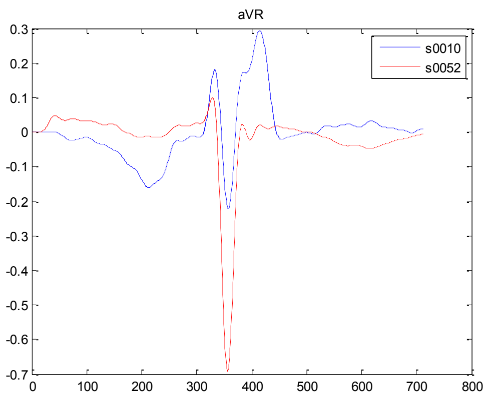


Figure 4.96 aVR lead.

4.3.1.2. Analysis

In this experiment, we can conclude that Wavelet Transform KLT based method outperformed other similarities measurement methods in twelve out of fifteen leads.

We verified that we have obtained the highest similarity among time-series in the following leads: V5, V6, V4, V3, L1 and L2, these were the leads where Wavelet Transform KLT based method performed the best.

Lastly, it will be also considered the leads where Pearson's correlation coefficient performed the best, which were the following: V3, V2, aVL, L1, V4 and L3.

4.3.2. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT PATIENTS WITH DIFFERENT DIAGNOSIS - II

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range. In Figure 4.97, it is represented the cardiac cycles of patient s0045.

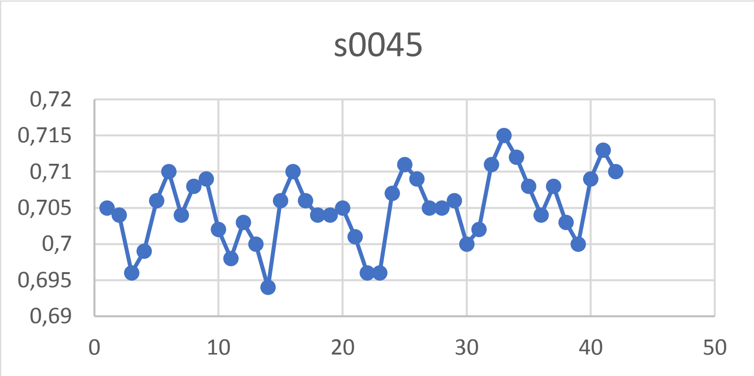


Figure 4.97 The cardiac cycles of patient s0045, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0045 the 14th cardiac cycle was the minimum cardiac cycle while the maximum one was the 33th cardiac cycle, where the HRV is 12ms.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 33th cardiac cycle of the patient s0045. It will be also compared the 13th cardiac cycle of the patient s0010 with the 14th cardiac cycle of the patient s0045. The last comparison it will be between the 13th cardiac cycle of the patient s0010 and the 25th cardiac cycle of the patient s0045, this will result in no losses of information.

4.3.2.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with longer data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	14	72	4.98
L2	S_{WT}	37	73	4.99
L3	S_{CC}	8	60	4.100
V1	S_{WT}	4	26	4.101
V2	S_{WT}	20	37	4.102
V3	S_{WT}	11	48	4.103
V4	S_{CC}	3	76	4.104
V5	S_{WT}	20	69	4.105
V6	S_{WT}	40	68	4.106
Vx	S_{WT}	27	89	4.107
Vy	SM_{AH}	2	28	4.108
Vz	S_{CC}	9	64	4.109
aVF	SM_{AH}	3	41	4.110
aVL	S_{CC}	12	70	4.111
aVR	S_{WT}	21	73	4.112

Table 4-9 Similarity between the 40^{th} cardiac cycle of the s0010 patient with the 33^{th} cardiac cycle of the patient s0045.

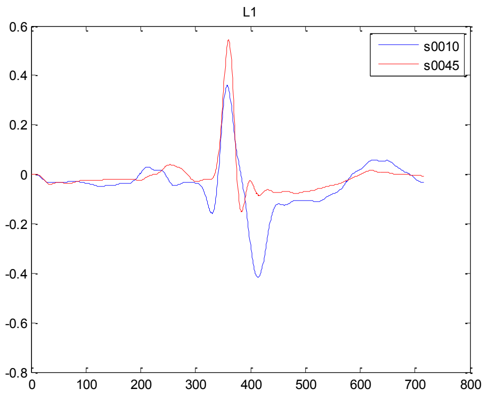


Figure 4.98 L1 lead.

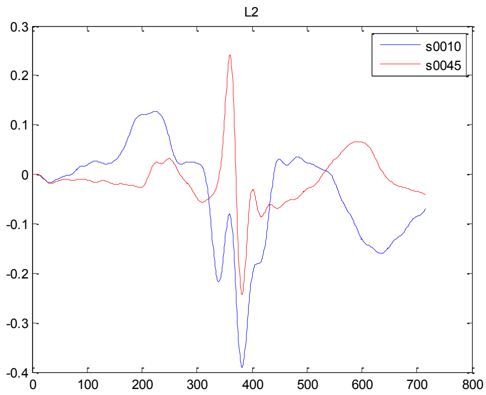


Figure 4.99 L2 lead.

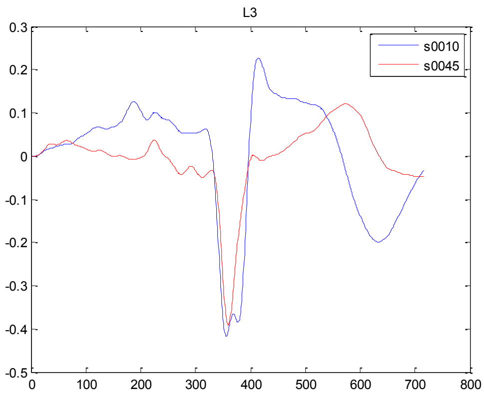


Figure 4.100 L3 lead.

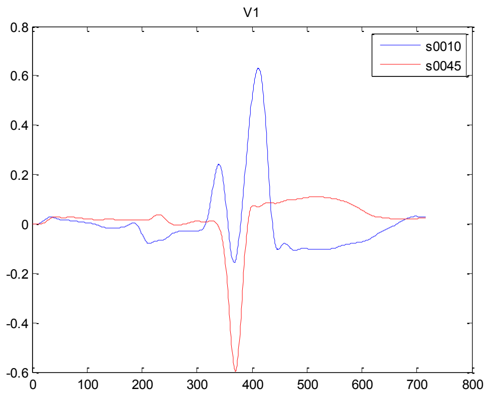


Figure 4.101 V1 lead.

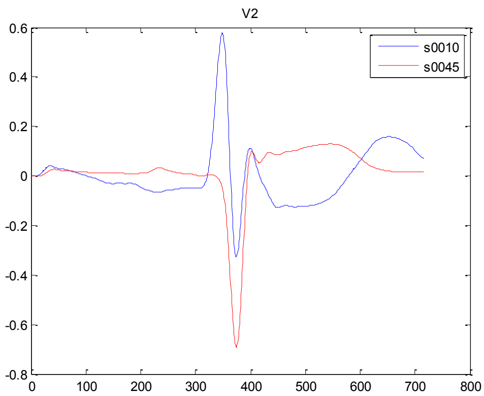


Figure 4.102 V2 lead.

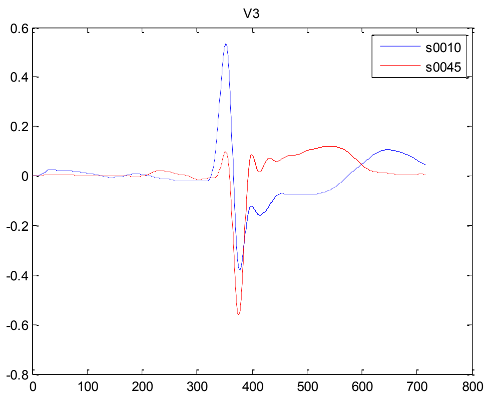


Figure 4.103 V3 lead.

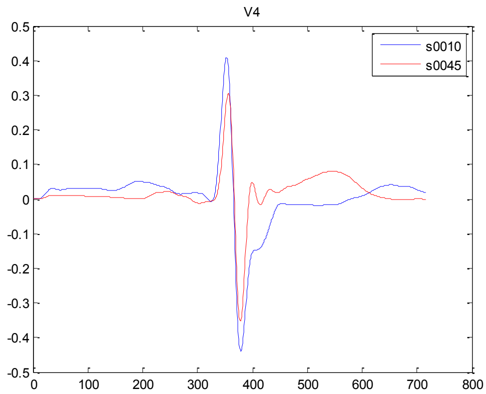


Figure 4.104 V4 lead.

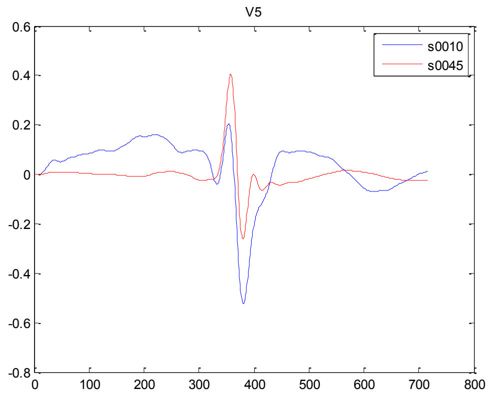


Figure 4.105 V5 lead.

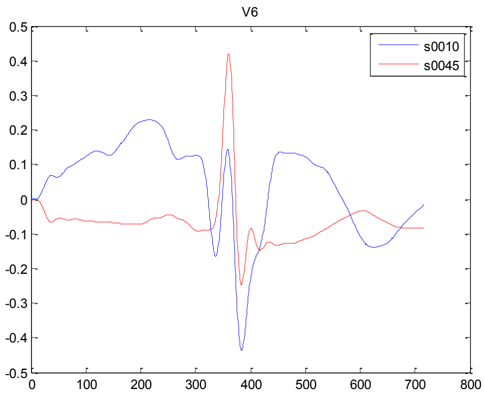


Figure 4.106 V6 lead.

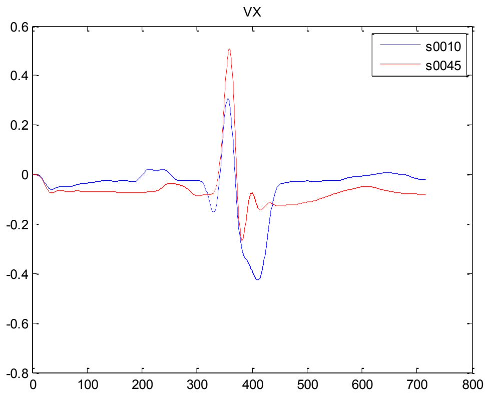


Figure 4.107 VX lead.

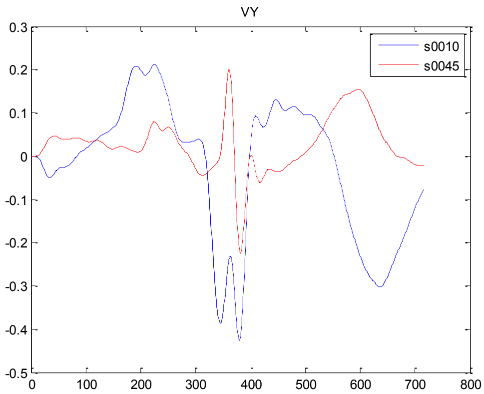


Figure 4. VY lead.

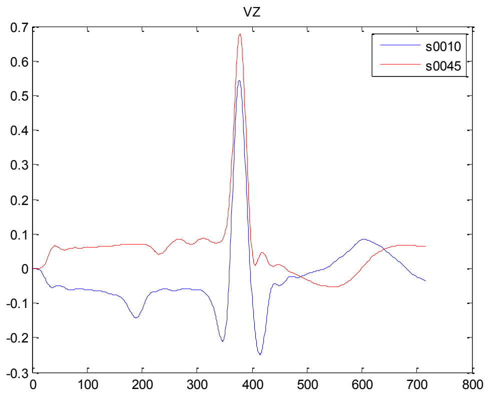


Figure 4.108 VZ lead.

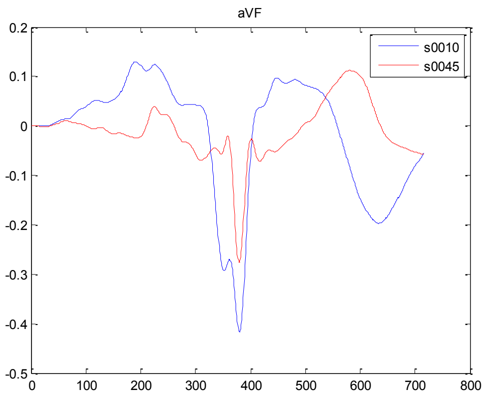


Figure 4.109 aVF lead.

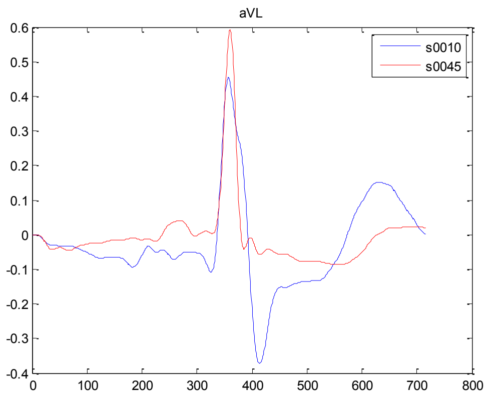


Figure 4.110 aVL lead.

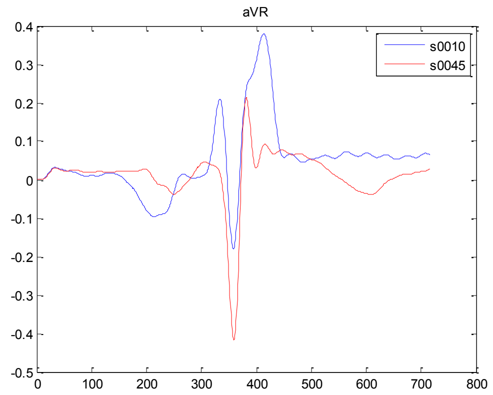


Figure 4.111 aVR lead.

4.3.2.2. Analysis

In this experiment, we can conclude that Wavelet Transform KLT based method outperformed other similarities measurement methods in eight out of fifteen leads, if we calculate its average considering all leads it outperforms the second-best method (Pearson's correlation coefficient) for 10%.

We verified that we have obtained the highest similarity among time-series in the following leads: VX, L2, aVR, V5, V6 and V4, these were the leads where Wavelet Transform KLT based method performed the best.

Lastly, it will be also considered the leads where Pearson's correlation coefficient performed the best, which were the following: V4, L1, aVL, VZ, VX and L3.

4.3.3. SIMILARITY MEASUREMENTS BETWEEN DIFFERENT PATIENTS WITH DIFFERENT DIAGNOSIS - III

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range. In Figure 4.113, it is represented the cardiac cycles of patient s0227.

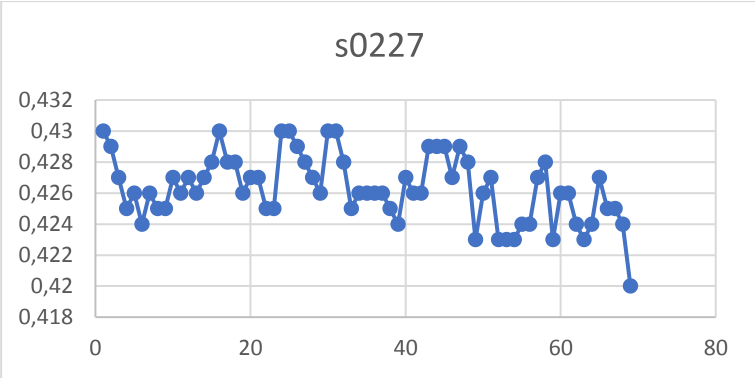


Figure 4.112 The cardiac cycles of patient s0227, where the x-axis represents the number of cardiac cycles and the y-axis the duration of those cycles.

It was observed that for patient s0227 the 69th cardiac cycle was the minimum cardiac cycle while the maximum one was the 31th cardiac cycle, thus the HRV is 5ms.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 31th cardiac cycle of the patient s0227. It will be also compared the 13th cardiac cycle of the patient s0010 with the 69th cardiac cycle of the patient s0227. The last comparison it will be between the 13th cardiac cycle of the patient s0010 and the 31th cardiac cycle of the patient s0227, this will result in a loss of two hundred and eighty-one data points.

4.3.3.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with shorter data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	**Best performed method**	**Exceeding the other methods performance by (%)**	**Similarity measure achieved (%)**	**Figure**
L1	S_{CC}	14	77	4.114
L2	S_{CC}	19	55	4.115
L3	S_{CC}	19	67	4.116
V1	SM_{AH}	9	39	4.117

V2	S_{CC}	13	60	4.118
V3	S_{CC}	15	74	4.119
V4	S_{CC}	29	83	4.120
V5	S_{CC}	33	80	4.121
V6	S_{CC}	9	44	4.122
Vx	S_{CC}	27	70	4.123
Vy	S_{CC}	8	43	4.124
Vz	S_{CC}	15	84	4.125
aVF	S_{CC}	26	61	4.126
aVL	S_{CC}	16	73	4.127
aVR	S_{CC}	9	70	4.128

Table 4-10 Similarity between the 13th cardiac cycle of the s0010 patient with the 69th cardiac cycle of the patient s0227.

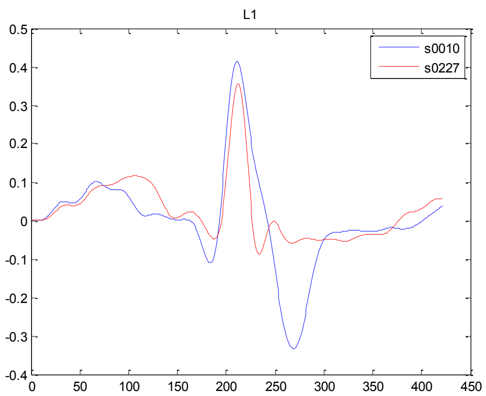


Figure 4.113 L1 lead.

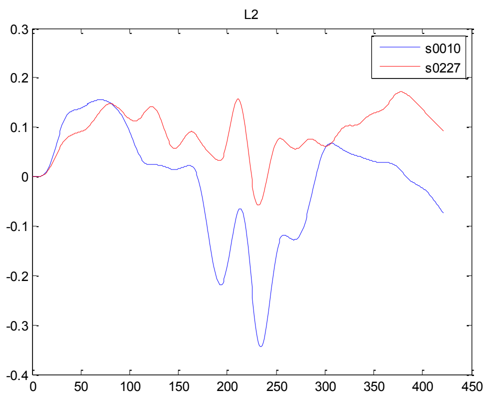


Figure 4.114 L2 lead.

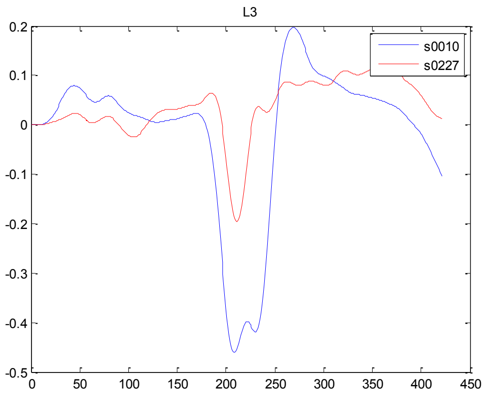


Figure 4.115 L3 lead.

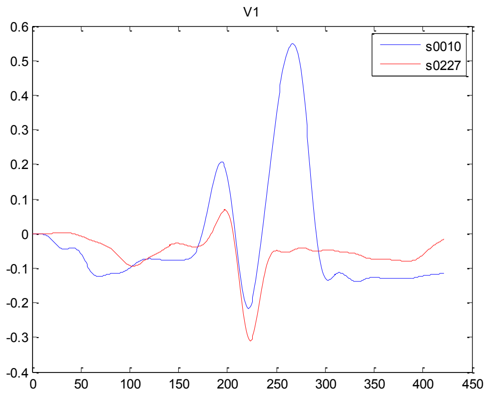


Figure 4.116 V1 lead.

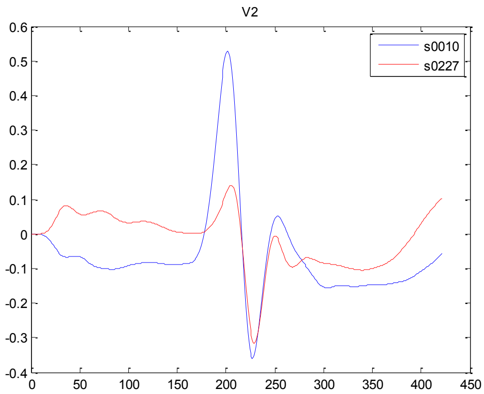


Figure 4.117 V2 lead.

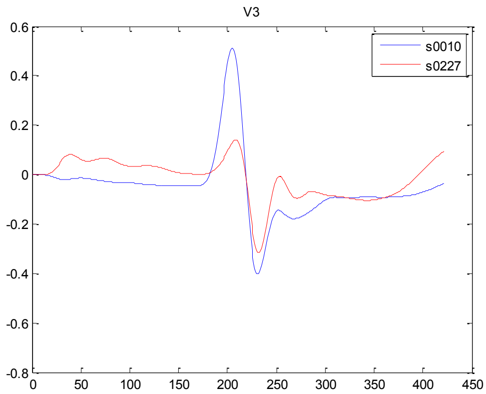


Figure 4.118 V3 lead.

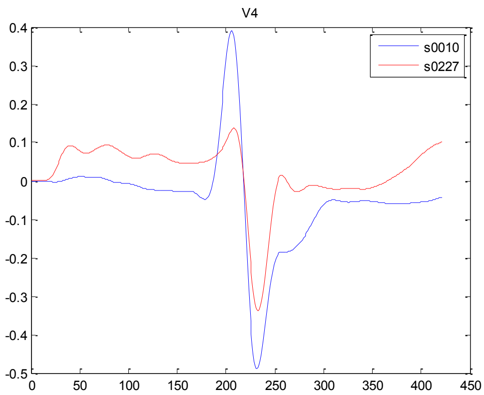


Figure 4.119 V4 lead.

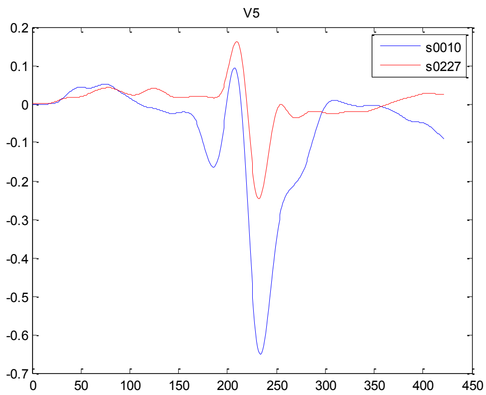


Figure 4.120 V5 lead.

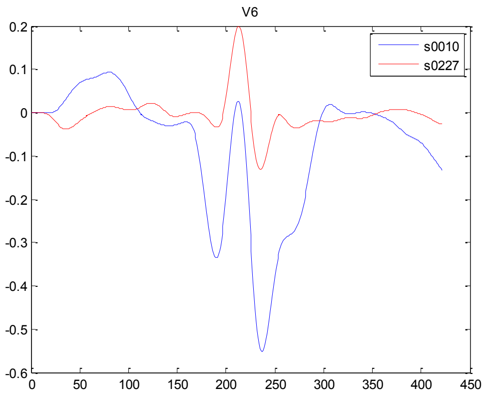


Figure 4.121 V6 lead.

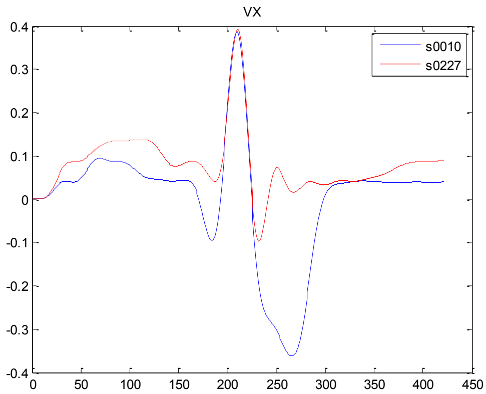


Figure 4.122 VX lead.

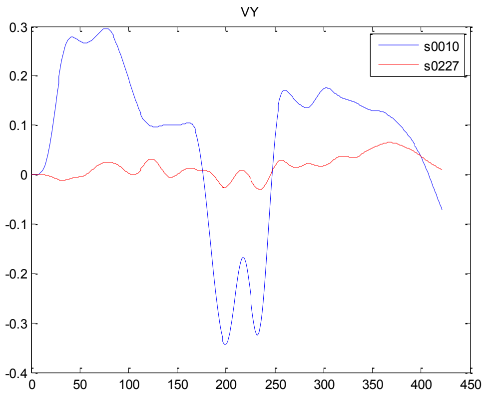


Figure 4.123 VY lead.

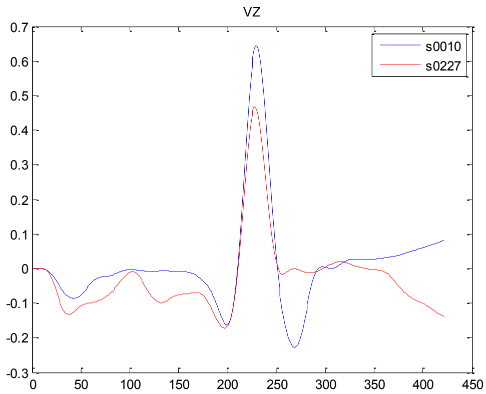


Figure 4.124 VZ lead.

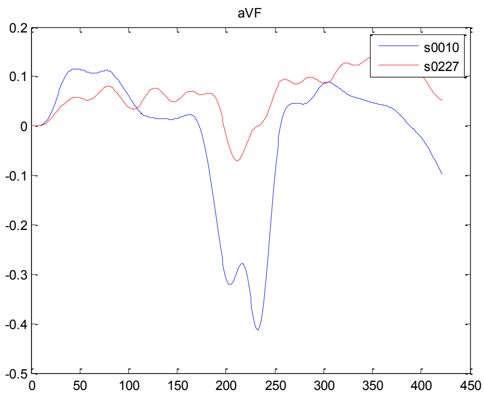


Figure 4.125 aVF lead.

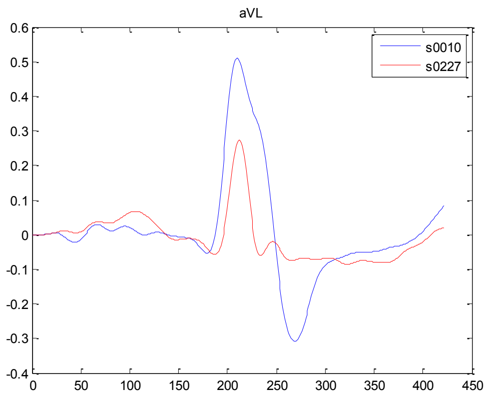


Figure 4.126 aVL lead.

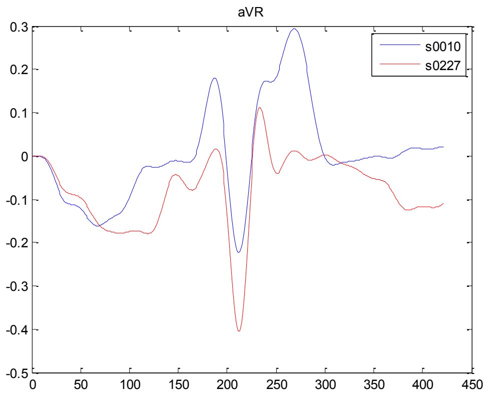


Figure 4.127 aVR lead.

4.3.3.2. Analysis

In this experiment, we can conclude that Pearson's correlation coefficient outperformed other similarities measurement methods in fourteen out of fifteen leads.

We verified that we have obtained the highest similarity among time-series in the following leads: VZ, V4, V5, L1, V3 and aVL, these were the leads where Pearson's correlation coefficient performed the best.

Lastly, it will be also considered the leads where Wavelet Trasnform KLT based method performed the best, which were the following: aVR, VZ, V4, V5, V3 and L1.

4.3.4. SIMILARITY MEASUREMENTS BETWEEN A PATIENT AND A HEALTHY CONTROL - I

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range. In Figure 4.49a, it is represented the cardiac cycles of healthy control s0462.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 17th cardiac cycle of the patient s0462. It will be also compared the 13th cardiac cycle of the patient s0010 with the 14th cardiac cycle of the patient s0462. The last comparison it will be between the 40th cardiac cycle of the patient s0010 and the 14th cardiac cycle of the patient s0462, this will result in a loss of one hundred and six data points.

4.3.4.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with shorter data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	12	68	4.129
L2	S_{WT}	32	56	4.130
L3	SM_{AH}	4	14	4.131
V1	SM_{AH}	6	34	4.132
V2	S_{CC}	13	70	4.133
V3	S_{CC}	7	81	4.134
V4	S_{WT}	34	99	4.135
V5	S_{WT}	80	97	4.136
V6	S_{WT}	74	89	4.137
Vx	S_{WT}	32	84	4.138

Vy	SM_{AH}	1	18	4.139
Vz	S_{CC}	2	59	4.140
aVF	S_{WT}	4	17	4.141
aVL	S_{CC}	8	49	4.142

Table 4-11 Similarity between the 13th cardiac cycle of the s0010 patient with the 14th cardiac cycle of the patient s0462.

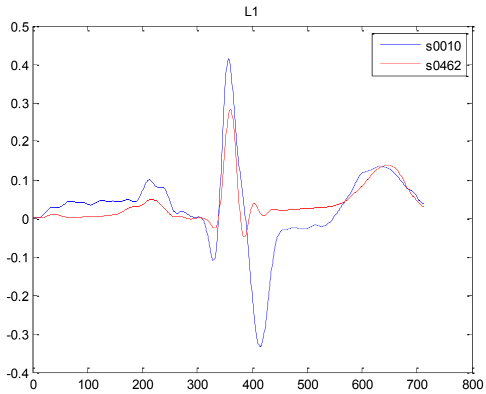


Figure 4.128 L1 lead.

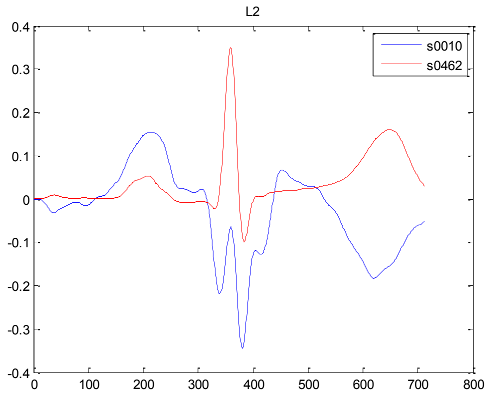


Figure 4.129 L2 lead.

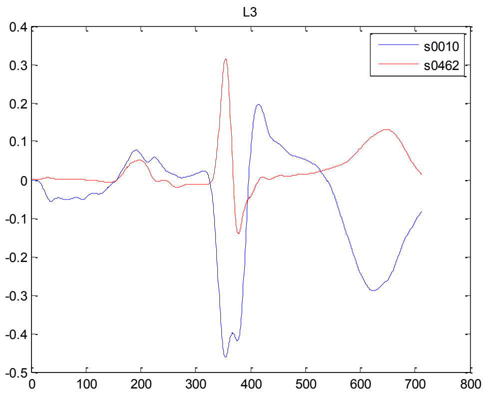


Figure 4.130 L3 lead.

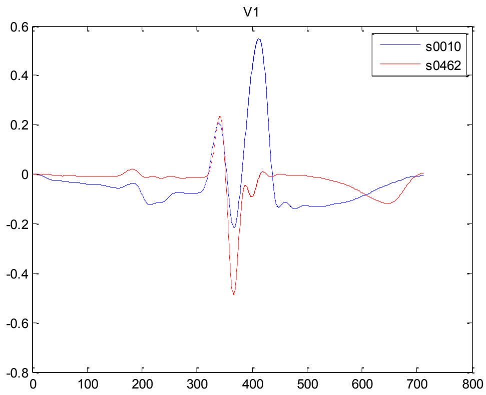


Figure 4.131 V1 lead.

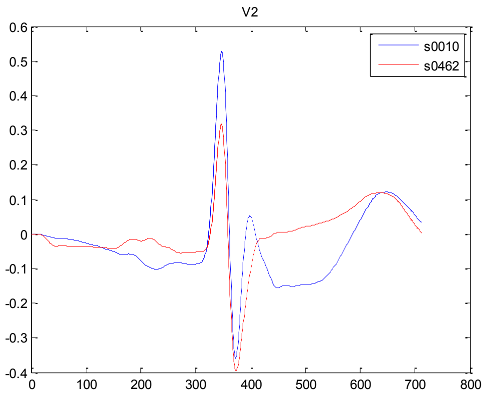


Figure 4.132 V2 lead.

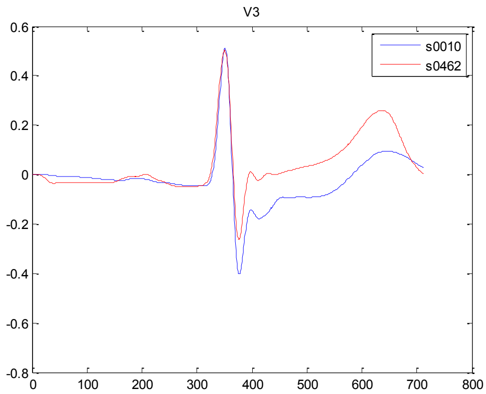


Figure 4.133 V3 lead.

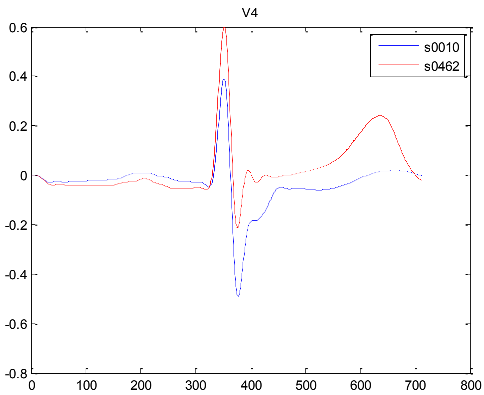


Figure 4.134 V4 lead.

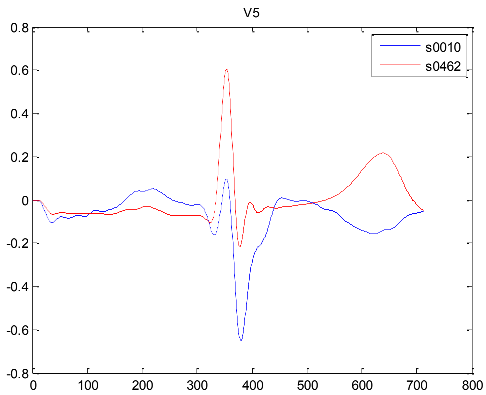


Figure 4.135 V5 lead.

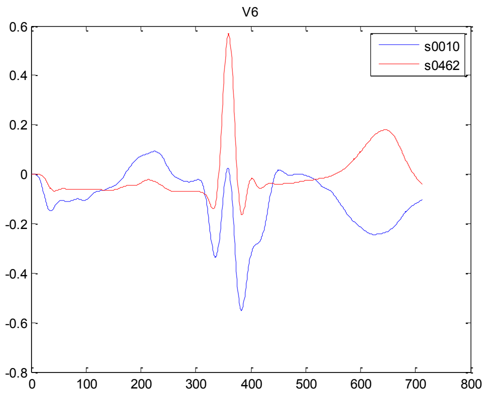


Figure 4.136 V6 lead.

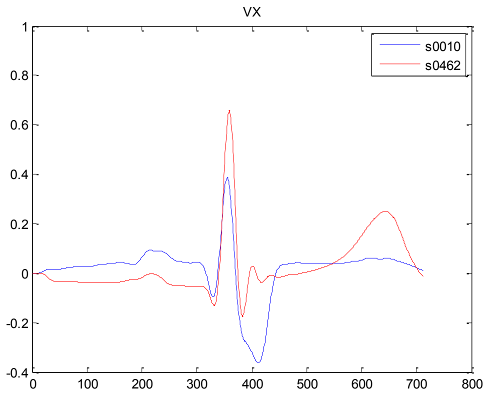


Figure 4.137 VX lead.

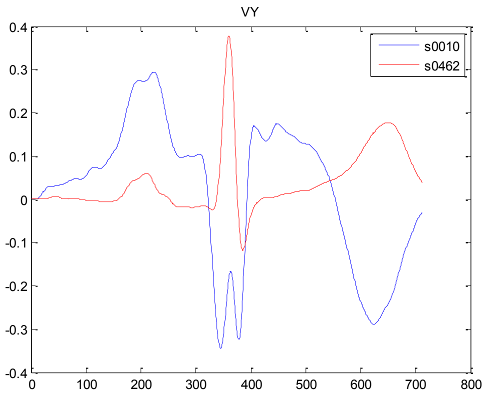


Figure 4.138 VY lead.

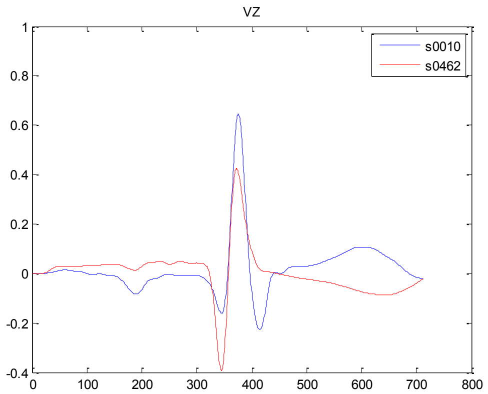


Figure 4.139 VZ lead.

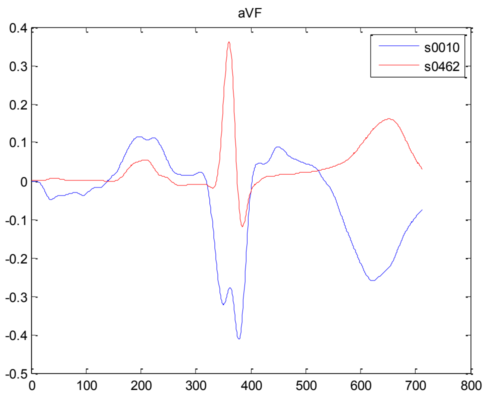


Figure 4.140 aVF lead.

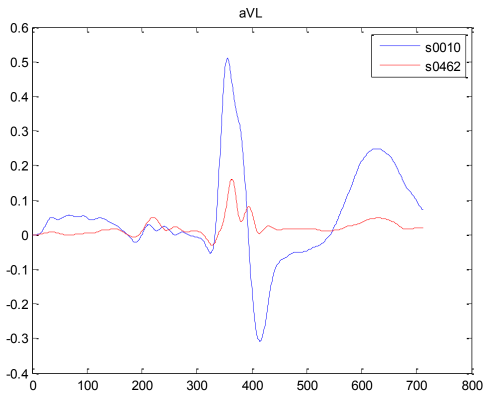


Figure 4.141 aVL lead.

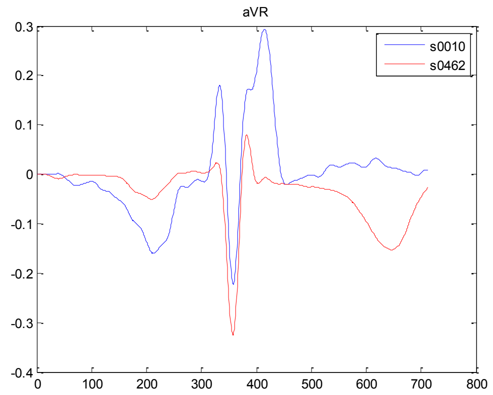


Figure 4.142 aVR lead.

4.3.4.2. Analysis

In this experiment, we can conclude that Wavelet Transform KLT based method outperformed other similarities measurement methods in seven out of fifteen leads, if we calculate its average considering all leads it outperforms the second-best method (Pearson's correlation coefficient) for 17%.

We verified that we have obtained the highest similarity among time-series in the following leads: V4, V5, V6, VX, V3 and VZ, these were the leads where Wavelet Transform KLT based method performed the best.

Lastly, it will be also considered the leads where Pearson's correlation coefficient performed the best, which were the following: V3, V2, L1, V4, VZ and VX.

4.3.5. SIMILARITY MEASUREMENTS BETWEEN A PATIENT AND A HEALTHY CONTROL - II

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range. In Figure 4.49b, it is represented the cardiac cycles of healthy control s0303.

So, to measure similarity in these patients it will be compared the 40th cardiac cycle of the s0010 patient with the 6th cardiac cycle of the patient s0303. It will be also compared the 13th cardiac cycle of the patient s0010 with the 12th cardiac cycle of the patient s0303. The last comparison it will be between the 40th cardiac cycle of the patient s0010 and the

16th cardiac cycle of the patient s0303, this will result in a lossless comparison in terms of data points.

4.3.5.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with shorter data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	Best performed method	Exceeding the other methods performance by (%)	Similarity measure achieved (%)	Figure
L1	S_{CC}	16	57	4.144
L2	S_{WT}	53	89	4.145
L3	SM_{AH}	1	15	4.146
V1	S_{CC}	1	26	4.147
V2	S_{CC}	2	32	4.148
V3	S_{CC}	24	60	4.149
V4	S_{CC}	32	76	4.150
V5	S_{WT}	39	76	4.151
V6	S_{WT}	32	50	4.152
Vx	S_{WT}	16	73	4.153
Vy	SM_i	1	20	4.154
Vz	S_{CC}	39	76	4.155
aVF	SM_i	3	13	4.156
aVL	S_{CC}	12	47	4.157
aVR	S_{WT}	54	87	4.158

Table 4-12 Similarity between the 13th cardiac cycle of the s0010 patient with the 12th cardiac cycle of the patient s0303.

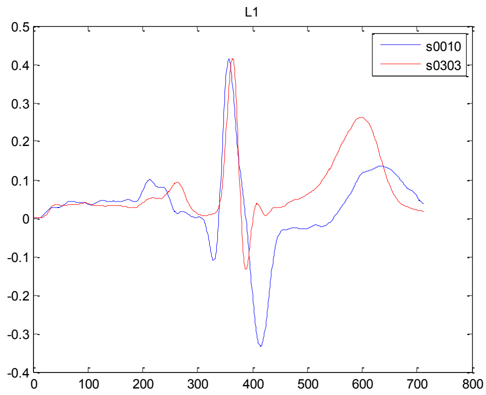


Figure 4.143 L1 lead.

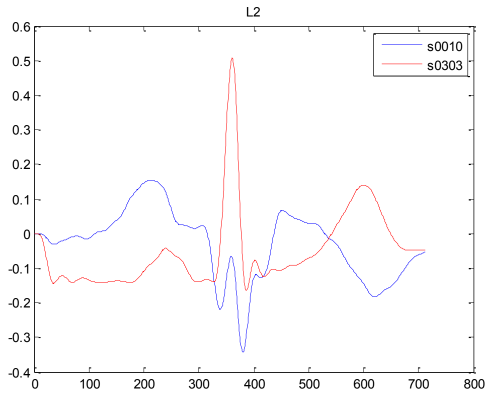


Figure 4.144 L2 lead.

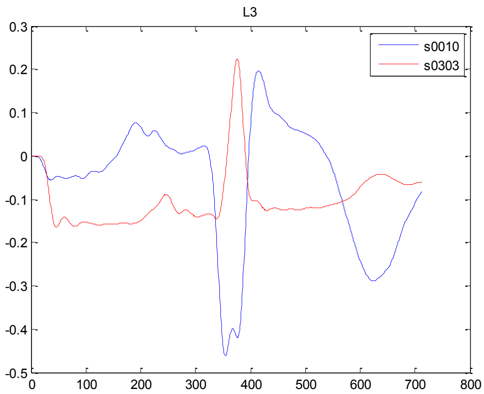


Figure 4.145 L3 lead.

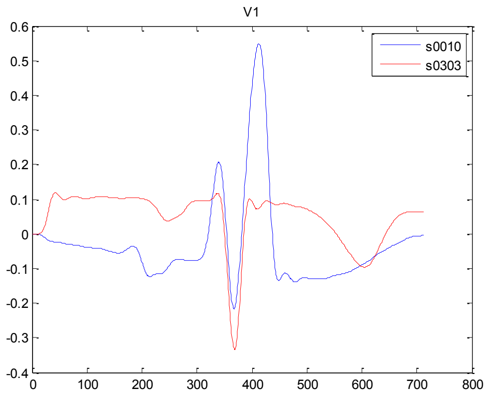


Figure 4.146 V1 lead.

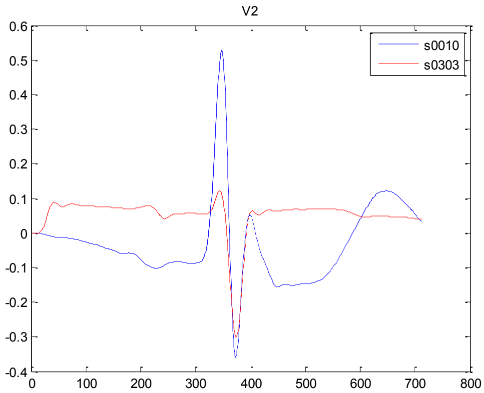


Figure 4.147 V2 lead.

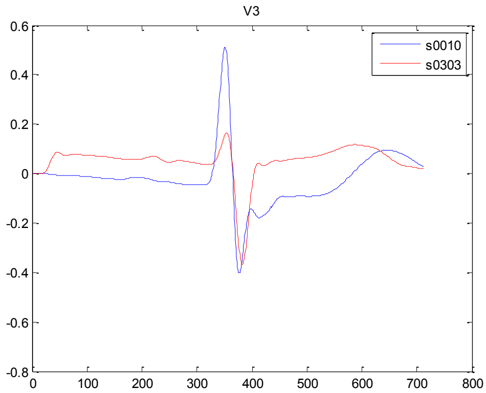


Figure 4.148 V3 lead.

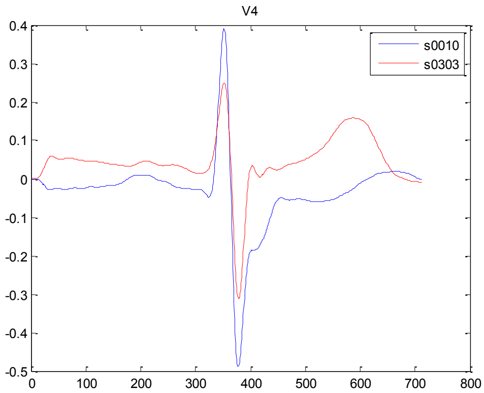


Figure 4.149 V4 lead.

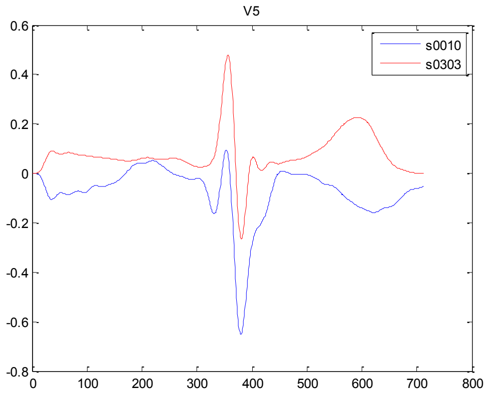


Figure 4.150 V5 lead.

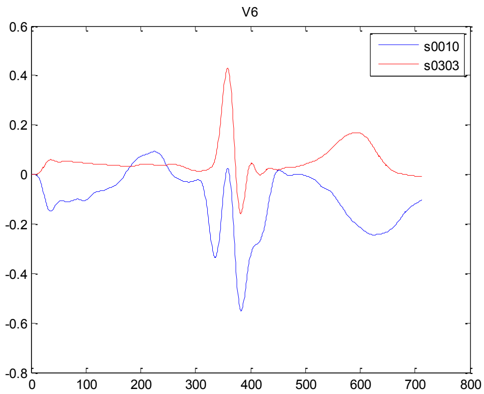


Figure 4.151 V6 lead.

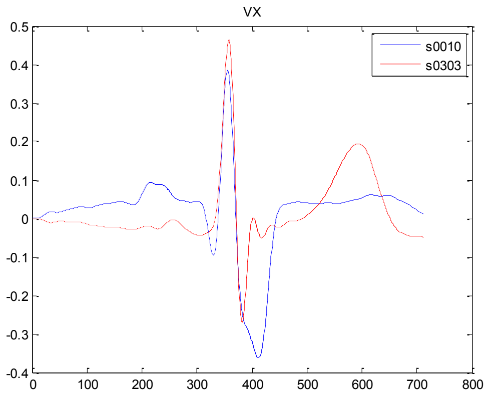


Figure 4.152 VX lead.

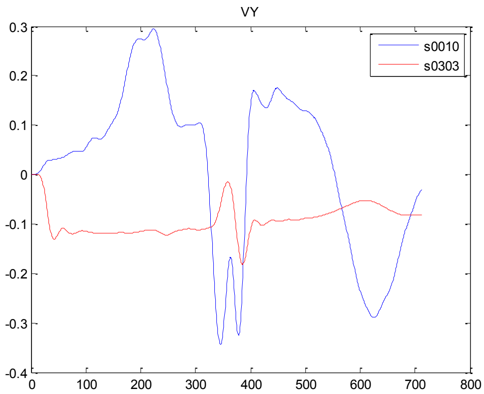


Figure 4.153 VY lead.

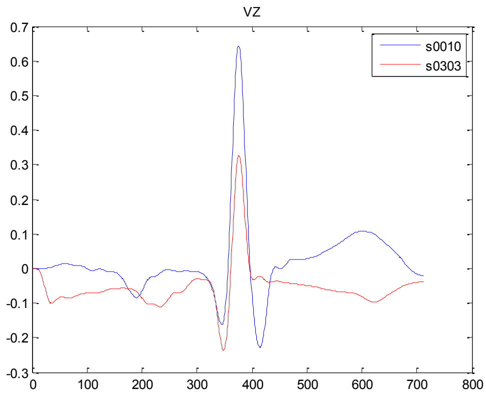


Figure 4.154 VZ lead.

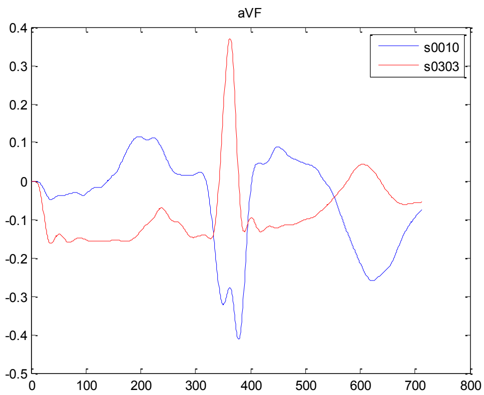


Figure 4.155 aVF lead.

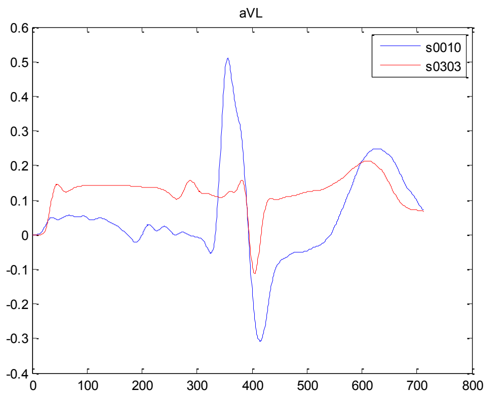


Figure 4.156 aVL lead.

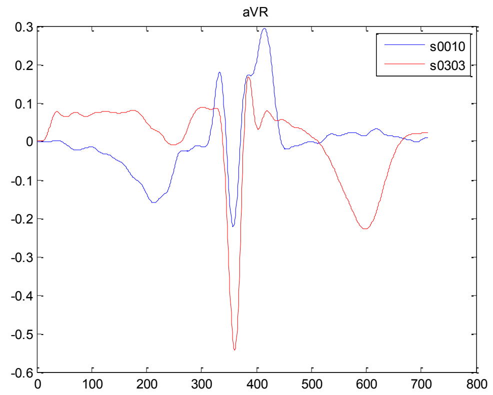


Figure 4.157 aVR lead.

4.3.5.2. Analysis

In this experiment, we can conclude that both methods (Wavelet Transform KLT based and Pearson's correlation coefficient) performed evenly. Pearson's correlation coefficient performed better in seven out of fifteen leads, but if we calculate the average for both methods considering all leads the Wavelet Transform KLT based performs 8% better.

We verified that we have obtained the highest similarity among time-series in the following leads: L2, aVR, V5, VX, V6 and V4, these were the leads where Wavelet Transform KLT based method performed the best.

The leads where Pearson's correlation coefficient performed the best, which were the following: VZ, V4, V3, VX, L1 and aVL.

4.3.6. SIMILARITY MEASUREMENTS BETWEEN A PATIENT AND A HEALTHY CONTROL - III

In this measurement, the cardiac signals that were tested, were collected from a cohort with different diagnosis, gender and age range.

In Figure 4.65, it is represented the cardiac cycles of healthy control s0311.

So, to measure similarity in these patients it will be compared the 40^{th} cardiac cycle of the s0010 patient with the 42^{th} cardiac cycle of the patient s0311. It will be also compared the 13^{th} cardiac cycle of the patient s0010 with the 29^{th} cardiac cycle of the patient s0311. The last comparison it will be between the 17^{th} cardiac cycle of the patient s0010 and the

41th cardiac cycle of the patient s0311, this will result in a lossless comparison in terms of data points.

4.3.6.1. Results

Since among these three comparisons it was observed that when comparing the time-series related to the cardiac cycles with closer data lengths best results were attained, for each ECG lead will only be presented the best performed results for the sake of thesis' simplicity.

Lead	**Best performed method**	**Exceeding the other methods performance by (%)**	**Similarity measure achieved (%)**	**Figure**
L1	S_{WT}	17	84	4.159
L2	S_{WT}	50	72	4.160
L3	S_{CC}	13	66	4.161
V1	S_{CC}	3	39	4.162
V2	S_{CC}	2	75	4.163
V3	S_{WT}	18	92	4.164
V4	S_{WT}	23	95	4.165
V5	S_{WT}	58	95	4.166
V6	S_{WT}	13	33	4.167
Vx	S_{CC}	4	56	4.168
Vy	S_{WT}	0,5	16	4.169
Vz	S_{CC}	15	86	4.170
aVF	S_{CC}	15	43	4.171
aVL	S_{CC}	11	68	4.172
aVR	S_{WT}	48	95	4.173

Table 4-13 Similarity between the 17th cardiac cycle of the s0010 patient with the 31th cardiac cycle of the patient s0311.

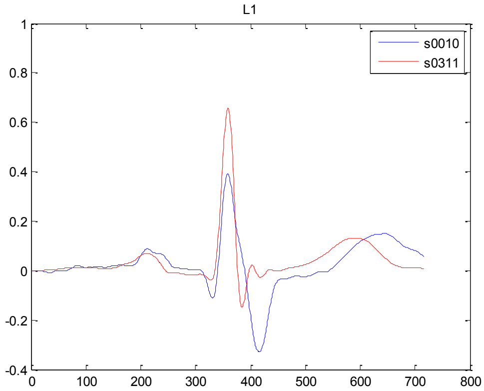


Figure 4.158 L1 lead.

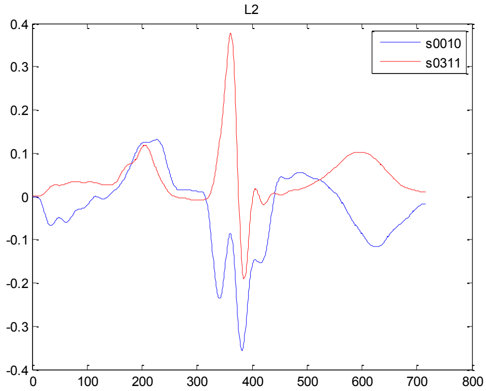


Figure 4.159 L2 lead.

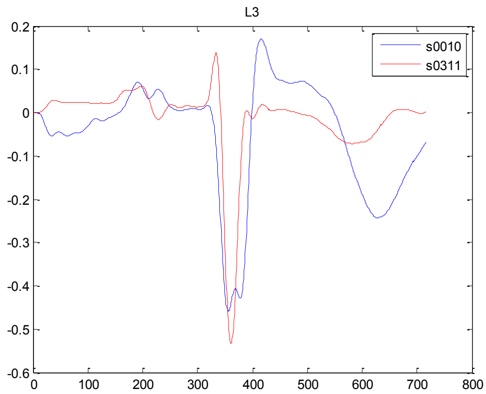


Figure 4.160 L3 lead.

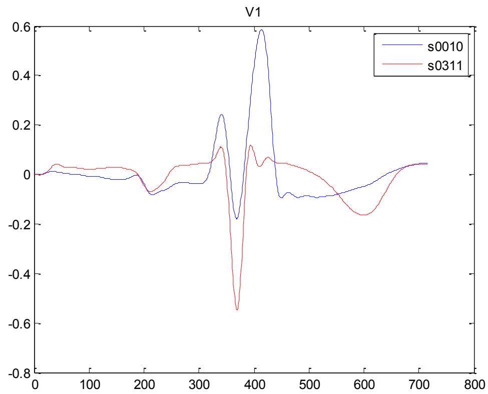


Figure 4.161 V1 lead.

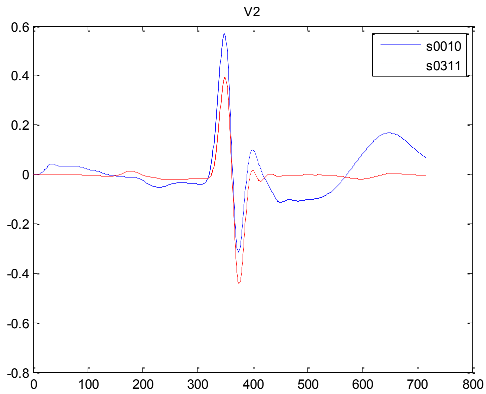


Figure 4.162 V2 lead.

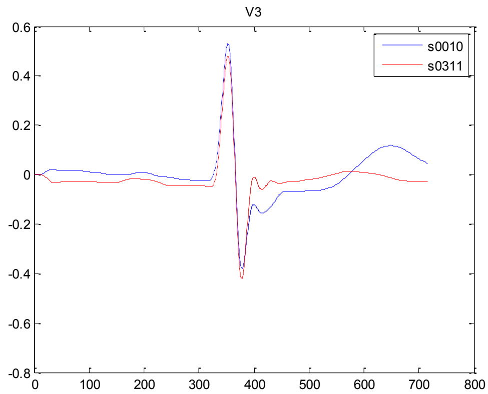


Figure 4.163 V3 lead.

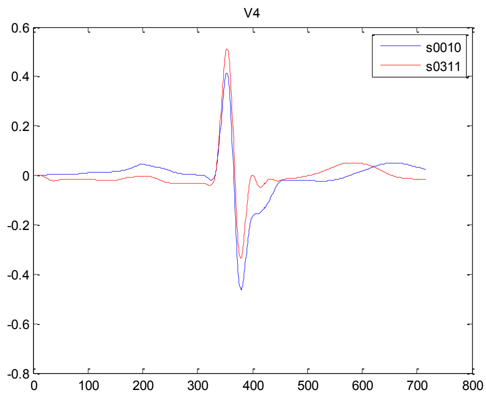


Figure 4.164 V4 lead.

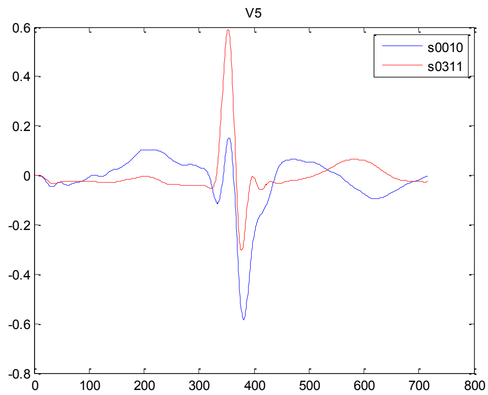


Figure 4.165 V5 lead.

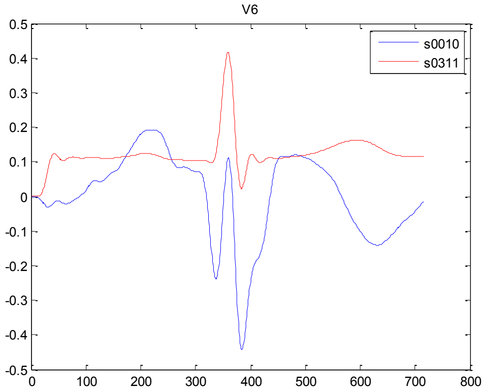


Figure 4.166 V6 lead.

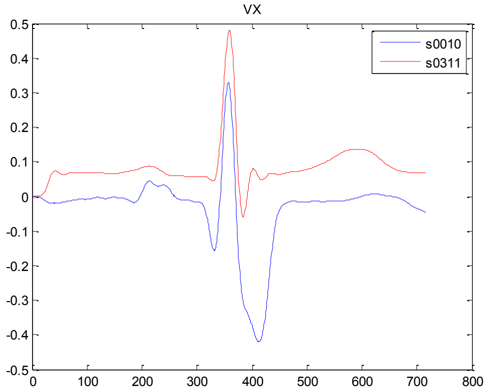


Figure 4.167 VX lead.

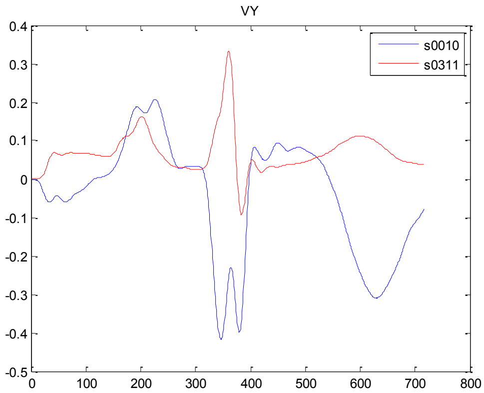


Figure 4.168 VY lead.

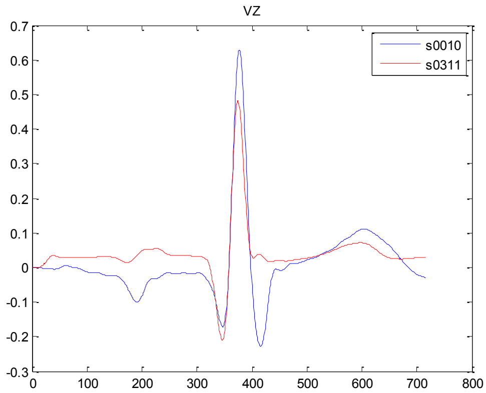


Figure 4.169 VZ lead.

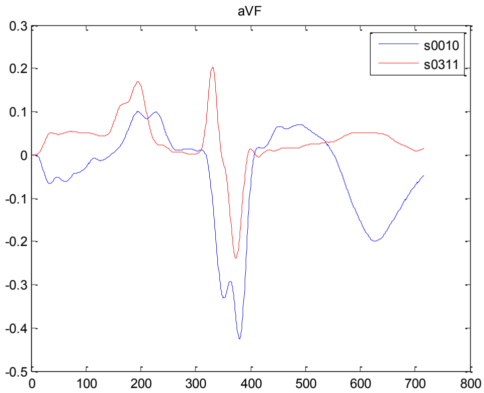


Figure 4.170 aVF lead.

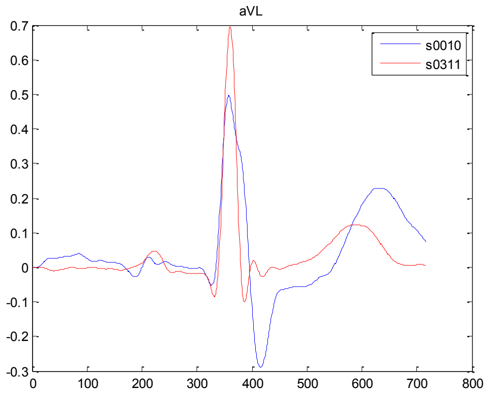


Figure 4.171 aVL lead.

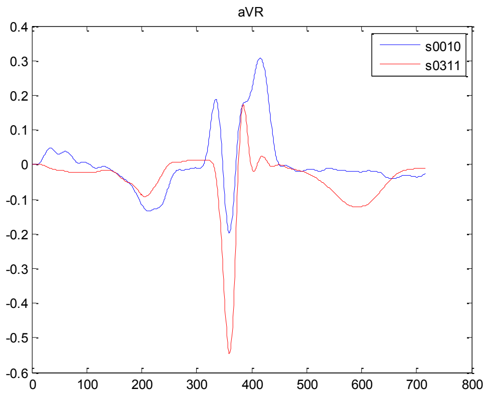


Figure 4.172 aVR lead.

4.3.6.2. Analysis

In this experiment, we can conclude that both methods (Wavelet Transform KLT based and Pearson's correlation coefficient) performed evenly. Wavelet Transform KLT based performed better in eight out of fifteen leads, also if we calculate the average for both methods considering all leads the Wavelet Transform KLT based performs 12% better.

We verified that we have obtained the highest similarity among time-series in the following leads: V5, V4, aVR, V3, L1 and V2, these were the leads where Wavelet Transform KLT based method performed the best.

The leads where Pearson's correlation coefficient performed the best, which were the following: VZ, V3, V2, aVL, L1 and L3.

5. CONCLUDING REMARKS

5.1. CONCLUSION

In this thesis, we can conclude from the first experience (finding the most representative ECG leads in terms of similarity values within cohorts) that the best methods for measuring similarity among time-series from the cohorts with patients with the same diagnosis would be Pearson's correlation coefficient immediately followed by the Wavelet Transform KLT based methods. We also concluded that statistically speaking the more consistent leads for this effect would be L1, V4, VZ and aVL for Pearson's correlation coefficient and V3, V4, VX and aVF for Wavelet Transform KLT based method.

On the second experience, we have measured the similarity among different cohorts with patients with different diagnosis, age range and gender. Using Pearson's correlation coefficient and Wavelet Transform KLT based method and the above referred seven leads, the aim was to find a pattern among DM patients with myocardial infarction. With this methodology we could not find a common pattern, so an average of the performance obtained on the six best performed leads was considered, as well as, an average of the performance obtained for all leads. Analysing all the measurements for Pearson's correlation coefficient we have constructed Table 5-1:

Cohort[4]	**Considering all leads**	**Considering the six best leads**	**Considering the four more consistent leads**
1.1	100%	100%	100%
1.2	88%	93%	91%
2.1	47%	73%	70%
2.2	49%	73%	64%
3.1	36%	63%	58%
3.2	42%	67%	71%
3.3	65%	79%	79%
4.1	34%	66%	60%

[4] Cohorts are described in detail, in section 4.1.

4.2	33%	62%	64%
4.3	47%	75%	61%

Table 5-1 Averaging the results of the measurements considering Pearson's correlation coefficient in different leads.

Analysing this table, we can conclude that the best methodology should be using the six best leads, where there are still results that are unexpected. If we look to cohort 3.3, where the patient has arterial hypertension we can explain the higher similarity in that comparison due to the amount of data that was lost through the pre-processing. The template signal lost 291 data points which is 41% of its signal, where it might be a loss of valuable data points. But this explanation cannot explain the result in cohort 4.3 which is a comparison with a healthy control. To explain these latter results a further study was made, and it was found that there is valuable data in T-waves when it comes to DM patients [28], which are often lost due to centring the signals by QRS complex.

Analysing all the measurements for Wavelet Transform KLT based we have constructed Table 5-2:

Cohort	Considering all leads	Considering the six best leads	Considering the four more consistent leads
1.1	100%	100%	100%
1.2	47%	74%	80%
2.1	46%	65%	37%
2.2	52%	77%	61%
3.1	65%	94%	75%
3.2	52%	73%	58%
3.3	31%	42%	30%
4.1	51%	83%	69%
4.2	41%	70%	40%
4.3	59%	89%	66%

Table 5-2 Averaging the results of the measurements considering Wavelet Transform based method in different leads.

If we analyse how Wavelet Transform KLT based method works, which focusses on the shape of the signal, the result on the cohort in 3.3 is due to the loss of data on template signal during pre-processing. We can also see that we achieved a better measurement in

3.1 which is a patient with myocardial infarction, diabetes mellitus and hyperuricemia than in the measurement between the same patient, which is as well not expected, but once again, the loss of data might be the explanation for this result, since it was lost 261 data points on the signal we wanted to compare with the template, this is a loss of 27% of the signal.

Besides the conclusions of the experiments, the data obtained is also valuable to emphasize some conclusions from other researches. Firstly, by measuring similarity among the same patient we can conclude that Wavelet Transform KLT based method is more sensitive to small variations on the signal than Pearson's correlation coefficient. On the other hand, Pearson's correlation coefficient is more robust (less affected by baseline variations) [23]. In this situation, we can say that Pearson's correlation coefficient would be the desirable choice for finding a pattern, even so more experiments needed to be tested and a larger database should be considered.

In overall, we can also conclude that for comparisons between patients with the same diagnosis Pearson's correlation coefficient will outperforms Wavelet transform based method (cohort 2.2 is an exception, where the difference between measurements is approximately 5%). However, when comparing patients with different diagnosis, we should consider as the best performed the method which presents less similarity since the pathologies are different. In this situation the Pearson's correlation coefficient is still the method to consider however cohort 3.3 is an exception (above explained).

Lastly, we can emphasize that low HRV is associated with CVDs and increasing age, as can be confirmed by other researchers [29]. If we check the results of HRV obtained along this research we verified that the healthy controls have higher HRV than the CVD's patients considered, where the healthy control s0311 is the exception, this might be also explained due to the age of this particular patient (79 years old) [17].

5.2. FUTURE WORK

For future work, it would be desirable to do the same experiments changing the procedure on centring ECG signals by QRS complex. Instead, we would align the ECG’s time-series in a way that we would not loose data points in QRS complex neither in T-wave.

Another thing to think about, would be time-scaling. In signals which QRS complex have the same shape but its duration varies (in order of 50-100 ms or higher), this implementation might be useful to improve the performance of similarity measurements.

Lastly, as we know biomedical signal’s characteristics varies cycle to cycle, even if it is just a little, where similarity 1 is literally impossible to achieve, so it would be interesting in replacing the scale [0, 1] into a more realistic one. For instance, we could measure the template’s cardiac cycle one with another in ideal conditions (comparing different cardiac signals from the template), and our real maximum would be the worst measurement.

REFERENCES

[1] World Health Organization, “WHO - Diabetes mellitus,” 2017. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs138/en/. [Accessed 03 09 2017].

[2] World Health Organization - Regional Office for Europe, “WHO/Europe - Diabetes,” 2017. [Online]. Available: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/diabetes . [Accessed 03 09 2017].

[3] World Health Organization, “WHO - 10 facts on diabetes,” 2017. [Online]. Available: http://www.who.int/features/factfiles/diabetes/en/ . [Accessed 03 09 2017].

[4] T.-c. Fu, “A review on time series data mining,” *Engineering Applications of Artificial Intelligence,* vol. 24, pp. 164-81, 2011.

[5] U.S Department of Health & Human Services, “Your Heart's Electrical System,” 17 11 2011. [Online]. Available: https://www.nhlbi.nih.gov/health/health-topics/topics/hhw/electrical. [Accessed 23 09 2017].

[6] R. M. Rangayyan, BIOMEDICAL SIGNAL ANALISYS A Case-Study Approach, Calgary, Alberta, Canada: John Wiley & Sons,Inc, 2002.

[7] “ECG timeline - History of electrocardiography,” 2009. [Online]. Available: https://ecglibrary.com/ecghist.html. [Accessed 23 9 2017].

[8] M. A. a. J. Lindsay, “A brief revier: history to understand fundamentals of electrocardiography,” 30 04 2012. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3714093/. [Accessed 23 09 2017].

[9] “Electrocardiogram (EKG or ECG),” [Online]. Available: http://www.webmd.com/heart/ekg-components-and-intervals. [Accessed 30 09 2017].

[10] R. M. a. B. Birchler, “ECG Measurement and Analysis,” p. 17, 24 2 2014.

[11] "CardioSecur - ECG Lead Systems," [Online]. Available: https://www.cardiosecur.com/en/your-heart/specialty-articles/ecg-lead-systems/. [Accessed 24 9 2017].

[12] "Frank vectorcardiographic system from standard 12 lead ECG: An effort to enhance cardiovascular diagnosis," 23 12 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26806119. [Accessed 24 09 2017].

[13] "Physionet," [Online]. Available: https://www.physionet.org/physiobank/database/ptbdb/. [Accessed 24 9 2017].

[14] "Vectorcardiographic Lead Systems," [Online]. Available: http://www.bem.fi/book/16/16.htm. [Accessed 24 9 2017].

[15] "Heart Rate variability vs Heart rate," [Online]. Available: https://hrvcourse.com/heart-rate-variability-vs-heart-rate/. [Accessed 24 9 2017].

[16] S. M. H. H. a. A. M. M. Nikhil Narayanaswamy, "Assessment of Risk Factor for Cardiovascular Disease Using Heart Rate Variability in Postmenopausal Women: A Comparative Study between Urban and Rural Indian Women," p. 6, 18 6 2013.

[17] K. B. G. R. d. M. C. A. S. J. W. J. S. M. F. R. R. a. O. M. D. Stefanie Hillebrand, "Heart rate variability and first cardiovascular event in populations without known cardiovascular disease: meta-analysis and dose–response meta-regression," p. 8, 30 1 2013.

[18] P. Esling and C. Agon, "Time-Series Data Mining," *ACM Computing Surveys,* vol. 45, pp. 12-34, 11 2012.

[19] Yi-Leh Wu, D. Agrawal and A. El Abbadi, *A comparison of DFT and DWT based similarity search in time-series databases,* Santa Barbara, California.

[20] A. Kianimajd, M. G. Ruano, P. Carvalho, J. Henriques, T. Rocha and S. Paredes, *Comparison of different methods of measuring similarity in physiologic time series,* Faro, 2017.

[21] C. Cassisi, P. Montalto, M. Aliotta, A. Cannata and A. Pulvirenti, “Similarity Measures and Dimensionality Reduction Techniques for Time Series Data Mining,” Catania, Itália, InTech , 2012, pp. 71-96.

[22] “Pearson's Correlation: Definition,” [Online]. Available: http://www.statisticshowto.com/what-is-the-pearson-correlation-coefficient/. [Accessed 24 09 2017].

[23] T. Rocha, S. . Paredes, P. Carvalho and J. Henriques, “An Efficient Strategy for Evaluating Similarity between Time Series based on Wavelet / Karhunen-Loève Transforms,” *Int. Conf. of the IEEE Engineering in Medicine and Biology,* pp. 6216-6219, 2012.

[24] Engineering Productivity Tools Ltd., “Definition of DFT and Inverse DFT (IDFT),” 1999. [Online]. Available: http://www.engineeringproductivitytools.com/stuff/T0001/PT01.HTM. [Accessed 14 09 2017].

[25] “The Fast Lifting Wavelet Transform,” [Online]. Available: http://www.polyvalens.com/blog/wavelets/theory/#7.+The+scaling+function+%5B7%5D . [Accessed 14 09 2017].

[26] “Mathworks,” [Online]. Available: http://www.mathworks.com/help/wavelet/ug/wavelet-packets.html?refresh=true&s_tid=gn_loc_drop. [Accessed 24 9 2017].

[27] “Mathworks,” [Online]. Available: https://www.mathworks.com/products/matlab.html. [Accessed 24 9 2017].

[28] M. Shlomo Stern and M. Samuel Sclarowsky, “The ECG in Diabetes Mellitus,” *Clinician Update,* p. 5, 20 10 2009.

[29] M. M. Reardon M, “Changes in heart rate variability with age.,” 19 11 1996. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/8945057. [Accessed 1 09 2017].

APPENDIX

I – Similarity Measurements between the same patient

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0014_L1	0,09	0,22	0,53	0,51	0,64	0,94	0,06	0,70		**s0014_VY**	0,74	**s0014_V1**	0,21
s0014_L2	0,07	0,11	0,48	0,41	0,46	0,86	0,07	0,60		**s0014_aVR**	0,80	**s0014_VY**	0,22
s0014_L3	0,16	0,25	0,61	0,78	0,41	0,94	0,18	0,78		**s0014_V6**	0,81	**s0014_V6**	0,27
s0014_V1	0,08	0,01	0,27	0,62	0,21	0,86	0,07	0,59		**s0014_L2**	0,86	**s0014_V2**	0,32
s0014_V2	0,11	0,06	0,43	0,65	0,32	0,88	0,12	0,65		**s0014_V1**	0,86	**s0014_VZ**	0,33
s0014_V3	0,27	0,68	0,61	0,83	0,70	0,93	0,17	0,82		**s0014_VZ**	0,86	**s0014_V5**	0,35
s0014_V4	0,34	0,85	0,60	0,84	0,76	0,92	0,27	0,87		**s0014_V2**	0,88	**s0014_L3**	0,41
s0014_V5	0,09	0,09	0,50	0,56	0,35	0,90	0,09	0,65		**s0014_VX**	0,89	**s0014_L2**	0,46
s0014_V6	0,05	0,00	0,43	0,51	0,27	0,81	0,06	0,53		**s0014_V5**	0,90	**s0014_aVR**	0,57
s0014_VX	0,22	0,50	0,54	0,75	0,53	0,89	0,17	0,73		**s0014_aVF**	0,90	**s0014_aVL**	0,63
s0014_VY	0,00	0,00	0,23	0,21	0,22	0,74	0,00	0,40		**s0014_V4**	0,92	**s0014_L1**	0,64
s0014_VZ	0,03	0,02	0,42	0,27	0,33	0,86	0,03	0,54		**s0014_V3**	0,93	**s0014_V3**	0,70
s0014_aVF	0,21	0,33	0,59	0,79	0,86	0,90	0,14	0,85		**s0014_L3**	0,94	**s0014_V4**	0,76
s0014_aVL	0,22	0,36	0,58	0,84	0,44	0,96	0,20	0,79		**s0014_L1**	0,94	**s0014_aVF**	0,86
s0014_aVR	0,02	0,01	0,32	0,07	0,57	0,80	0,01	0,56	0,67	**s0014_aVL**	0,96	**s0014_VX**	0,87
					0,47	0,88					0,93		0,74

II – Similarity Measurements between different patients with the same diagnosis - I

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0088_L1	0,04	0,00	0,28	0,28	0,78	0,72	0,05	0,60		**s0088_L2**	0,00	**s0088_V4**	0,16
s0088_L2	0,02	0,00	0,23	0,12	0,47	0,00	0,02	0,27		**s0088_V1**	0,00	**s0088_V3**	0,17
s0088_L3	0,11	0,13	0,46	0,70	0,56	0,84	0,19	0,70		**s0088_V6**	0,13	**s0088_VY**	0,25
s0088_V1	0,00	0,00	0,08	0,10	0,41	0,00	0,00	0,20		**s0088_aVR**	0,27	**s0088_V2**	0,26
s0088_V2	0,01	0,00	0,13	0,20	0,26	0,28	0,01	0,25		**s0088_V2**	0,28	**s0088_VZ**	0,38
s0088_V3	0,05	0,01	0,20	0,37	0,17	0,51	0,07	0,36		**s0088_VY**	0,45	**s0088_V5**	0,39
s0088_V4	0,10	0,20	0,29	0,45	0,16	0,58	0,12	0,44		**s0088_VX**	0,46	**s0088_V1**	0,41
s0088_V5	0,05	0,06	0,38	0,43	0,39	0,55	0,04	0,46		**s0088_V3**	0,51	**s0088_V6**	0,43
s0088_V6	0,01	0,00	0,27	0,20	0,43	0,13	0,01	0,30		**s0088_V5**	0,55	**s0088_L2**	0,47
s0088_VX	0,03	0,00	0,23	0,19	0,59	0,46	0,03	0,43		**s0088_V4**	0,58	**s0088_aVF**	0,55
s0088_VY	0,02	0,00	0,25	0,39	0,25	0,45	0,02	0,36		**s0088_VZ**	0,65	**s0088_L3**	0,56
s0088_VZ	0,06	0,24	0,31	0,42	0,38	0,65	0,04	0,48		**s0088_L1**	0,72	**s0088_VX**	0,59
s0088_aVF	0,09	0,35	0,46	0,58	0,55	0,74	0,14	0,62		**s0088_aVF**	0,74	**s0088_aVL**	0,65
s0088_aVL	0,07	0,01	0,34	0,58	0,65	0,84	0,12	0,69	0,44	**s0088_aVL**	0,84	**s0088_L1**	0,78
s0088_aVR	0,02	0,00	0,23	0,07	0,79	0,27	0,02	0,43		**s0088_L3**	0,84	**s0088_aVR**	0,79
					0,46	0,47					0,73		0,65

III – Similarity Measurements between different patients with the same diagnosis - II

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0004_L1	0,04	0,24	0,24	0,31	0,27	0,59	0,06	0,39		**s0004_aVF**	0,19	**s0004_L3**	0,07
s0004_L2	0,04	0,02	0,24	0,23	0,78	0,33	0,03	0,45		**s0004_V1**	0,21	**s0004_aVF**	0,16
s0004_L3	0,03	0,11	0,14	0,38	0,07	0,29	0,04	0,27		**s0004_V6**	0,22	**s0004_L1**	0,27
s0004_V1	0,02	0,01	0,12	0,34	0,35	0,21	0,03	0,30		**s0004_L3**	0,29	**s0004_VY**	0,30
s0004_V2	0,19	0,33	0,53	0,78	0,78	0,88	0,11	0,81		**s0004_VY**	0,33	**s0004_aVL**	0,30
s0004_V3	0,27	0,72	0,63	0,83	0,77	0,91	0,23	0,84		**s0004_L2**	0,33	**s0004_V1**	0,35
s0004_V4	0,17	0,51	0,48	0,63	0,87	0,78	0,14	0,76		**s0004_V5**	0,37	**s0004_V5**	0,54
s0004_V5	0,03	0,02	0,29	0,32	0,54	0,37	0,03	0,41		**s0004_aVR**	0,52	**s0004_aVR**	0,58
s0004_V6	0,02	0,03	0,29	0,29	0,61	0,22	0,02	0,40		**s0004_VX**	0,54	**s0004_V6**	0,61
s0004_VX	0,05	0,08	0,24	0,31	0,64	0,54	0,04	0,49		**s0004_VZ**	0,58	**s0004_VX**	0,64
s0004_VY	0,01	0,00	0,13	0,35	0,30	0,33	0,02	0,33		**s0004_L1**	0,59	**s0004_V3**	0,77
s0004_VZ	0,06	0,07	0,35	0,42	0,80	0,58	0,08	0,60		**s0004_aVL**	0,63	**s0004_L2**	0,78
s0004_aVF	0,03	0,02	0,14	0,30	0,16	0,19	0,05	0,22		**s0004_V4**	0,78	**s0004_V2**	0,78
s0004_aVL	0,05	0,40	0,27	0,54	0,30	0,63	0,05	0,49		**s0004_V2**	0,88	**s0004_VZ**	0,80
s0004_aVR	0,03	0,00	0,27	0,15	0,58	0,52	0,03	0,46	0,48	**s0004_V3**	0,91	**s0004_V4**	0,87
					0,52	0,49					0,73		0,77

IV – Similarity Measurements between different Healthy Controls - I

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0303_L1	0,01	0,00	0,32	0,00	0,29	0,77	0,01	0,46		**s0303_aVL**	0,00	**s0303_aVR**	0,18
s0303_L2	0,12	0,25	0,45	0,21	0,31	0,81	0,11	0,53		**s0303_V3**	0,60	**s0303_VY**	0,21
s0303_L3	0,07	0,10	0,44	0,09	0,44	0,62	0,06	0,50		**s0303_L3**	0,62	**s0303_L1**	0,29
s0303_V1	0,27	0,60	0,58	0,67	0,37	0,82	0,18	0,69		**s0303_V2**	0,72	**s0303_L2**	0,31
s0303_V2	0,14	0,07	0,43	0,54	0,42	0,72	0,17	0,57		**s0303_V4**	0,76	**s0303_VZ**	0,35
s0303_V3	0,03	0,00	0,28	0,23	0,57	0,60	0,04	0,48		**s0303_VY**	0,76	**s0303_V1**	0,37
s0303_V4	0,08	0,03	0,32	0,50	0,44	0,76	0,08	0,57		**s0303_aVF**	0,77	**s0303_aVL**	0,39
s0303_V5	0,08	0,17	0,40	0,51	0,66	0,86	0,09	0,68		**s0303_L1**	0,77	**s0303_V2**	0,42
s0303_V6	0,14	0,34	0,46	0,64	0,50	0,87	0,14	0,67		**s0303_L2**	0,81	**s0303_L3**	0,44
s0303_VX	0,06	0,13	0,36	0,48	0,61	0,84	0,05	0,64		**s0303_V1**	0,82	**s0303_V4**	0,44
s0303_VY	0,16	0,07	0,43	0,41	0,21	0,76	0,19	0,53		**s0303_aVR**	0,82	**s0303_V6**	0,50
s0303_VZ	0,21	0,36	0,49	0,73	0,35	0,94	0,18	0,72		**s0303_VX**	0,84	**s0303_V3**	0,57
s0303_aVF	0,21	0,57	0,54	0,49	0,80	0,77	0,18	0,70		**s0303_V5**	0,86	**s0303_VX**	0,61
s0303_aVL	0,00	0,00	0,24	0,00	0,39	0,00	0,01	0,21		**s0303_V6**	0,87	**s0303_V5**	0,66
s0303_aVR	0,03	0,01	0,43	0,01	0,18	0,82	0,03	0,48	0,56	**s0303_VZ**	0,94	**s0303_aVF**	0,80
					0,44	0,73					0,86		0,60

V – Similarity Measurements between different Healthy Controls - II

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0311_L1	0,15	0,13	0,32	0,17	0,39	0,83	0,15	0,51		**s0311_L3**	0,29	**s0311_aVL**	0,01
s0311_L2	0,32	0,68	0,65	0,66	0,93	0,80	0,33	0,74		**s0311_aVF**	0,50	**s0311_V6**	0,32
s0311_L3	0,05	0,01	0,25	0,03	0,76	0,29	0,05	0,44		**s0311_V3**	0,76	**s0311_aVF**	0,33
s0311_V1	0,19	0,52	0,55	0,56	0,83	0,81	0,17	0,73		**s0311_aVL**	0,78	**s0311_L1**	0,39
s0311_V2	0,22	0,33	0,59	0,70	0,68	0,81	0,19	0,73		**s0311_V6**	0,79	**s0311_VX**	0,40
s0311_V3	0,08	0,01	0,38	0,53	0,86	0,76	0,06	0,72		**s0311_L2**	0,80	**s0311_VY**	0,51
s0311_V4	0,15	0,11	0,45	0,70	0,89	0,83	0,11	0,81		**s0311_VY**	0,80	**s0311_aVR**	0,58
s0311_V5	0,19	0,28	0,50	0,76	0,86	0,85	0,15	0,83		**s0311_V2**	0,81	**s0311_VZ**	0,65
s0311_V6	0,02	0,05	0,38	0,15	0,32	0,79	0,01	0,49		**s0311_V1**	0,81	**s0311_V2**	0,68
s0311_VX	0,08	0,20	0,51	0,55	0,40	0,85	0,07	0,63		**s0311_VZ**	0,82	**s0311_L3**	0,76
s0311_VY	0,25	0,59	0,66	0,62	0,51	0,80	0,20	0,69		**s0311_L1**	0,83	**s0311_V1**	0,83
s0311_VZ	0,15	0,09	0,52	0,60	0,65	0,82	0,18	0,69		**s0311_aVR**	0,83	**s0311_V3**	0,86
s0311_aVF	0,16	0,15	0,50	0,35	0,33	0,50	0,15	0,45		**s0311_V4**	0,83	**s0311_V5**	0,86
s0311_aVL	0,07	0,01	0,19	0,00	0,01	0,78	0,08	0,35		**s0311_VX**	0,85	**s0311_V4**	0,89
s0311_aVR	0,25	0,51	0,52	0,48	0,58	0,83	0,22	0,64	0,63	**s0311_V5**	0,85	**s0311_L2**	0,93
					0,60	0,76					0,83		0,86

VI - Similarity Measurements between different patients with different diagnosis - I

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0052_L1	0,09	0,02	0,35	0,47	0,92	0,64	0,08	0,68		**s0052_L2**	0,00	**s0052_L3**	0,25
s0052_L2	0,02	0,00	0,15	0,11	0,91	0,00	0,02	0,39		**s0052_V6**	0,00	**s0052_VY**	0,28
s0052_L3	0,03	0,00	0,31	0,44	0,25	0,49	0,04	0,42		**s0052_VY**	0,00	**s0052_aVF**	0,28
s0052_V1	0,02	0,02	0,18	0,31	0,33	0,25	0,03	0,30		**s0052_aVF**	0,00	**s0052_V1**	0,33
s0052_V2	0,06	0,35	0,39	0,48	0,68	0,70	0,06	0,62		**s0052_V5**	0,11	**s0052_aVL**	0,36
s0052_V3	0,07	0,26	0,41	0,48	0,93	0,71	0,06	0,71		**s0052_V1**	0,25	**s0052_VZ**	0,49
s0052_V4	0,05	0,01	0,25	0,28	0,94	0,60	0,05	0,61		**s0052_aVR**	0,33	**s0052_aVR**	0,67
s0052_V5	0,01	0,00	0,11	0,09	0,97	0,11	0,01	0,39		**s0052_VZ**	0,46	**s0052_V2**	0,68
s0052_V6	0,00	0,00	0,09	0,06	0,95	0,00	0,00	0,37		**s0052_VX**	0,49	**s0052_VX**	0,87
s0052_VX	0,04	0,00	0,25	0,29	0,87	0,49	0,04	0,55		**s0052_L3**	0,49	**s0052_L2**	0,91
s0052_VY	0,01	0,00	0,14	0,24	0,28	0,00	0,01	0,22		**s0052_V4**	0,60	**s0052_L1**	0,92
s0052_VZ	0,04	0,02	0,29	0,37	0,49	0,46	0,08	0,44		**s0052_L1**	0,64	**s0052_V3**	0,93
s0052_aVF	0,02	0,00	0,18	0,21	0,28	0,00	0,02	0,22		**s0052_aVL**	0,65	**s0052_V4**	0,94
s0052_aVL	0,06	0,00	0,35	0,54	0,36	0,65	0,07	0,51		**s0052_V2**	0,70	**s0052_V6**	0,95
s0052_aVR	0,05	0,00	0,24	0,17	0,67	0,33	0,05	0,41	0,45	**s0052_V3**	0,71	**s0052_V5**	0,97
					0,65	0,36					0,63		0,94

VII - Similarity Measurements between different patients with different diagnosis - II

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0045_L1	0,11	0,12	0,34	0,56	0,58	0,72	0,12	0,62		**s0045_VY**	0,00	**s0045_L3**	0,25
s0045_L2	0,05	0,20	0,36	0,32	0,73	0,19	0,06	0,47		**s0045_V1**	0,07	**s0045_VY**	0,26
s0045_L3	0,05	0,12	0,39	0,52	0,25	0,60	0,07	0,51		**s0045_V2**	0,11	**s0045_V1**	0,26
s0045_V1	0,01	0,00	0,16	0,22	0,26	0,07	0,01	0,22		**s0045_V6**	0,14	**s0045_aVF**	0,30
s0045_V2	0,01	0,00	0,16	0,17	0,37	0,11	0,01	0,23		**s0045_L2**	0,19	**s0045_V2**	0,37
s0045_V3	0,04	0,05	0,27	0,34	0,48	0,37	0,04	0,40		**s0045_aVF**	0,36	**s0045_aVL**	0,47
s0045_V4	0,18	0,73	0,52	0,65	0,64	0,76	0,14	0,71		**s0045_V3**	0,37	**s0045_V3**	0,48
s0045_V5	0,05	0,09	0,36	0,43	0,69	0,49	0,04	0,54		**s0045_V5**	0,49	**s0045_VZ**	0,55
s0045_V6	0,01	0,00	0,28	0,18	0,68	0,14	0,01	0,38		**s0045_aVR**	0,52	**s0045_L1**	0,58
s0045_VX	0,09	0,18	0,37	0,49	0,89	0,62	0,09	0,67		**s0045_L3**	0,60	**s0045_V4**	0,64
s0045_VY	0,01	0,00	0,19	0,28	0,26	0,00	0,02	0,24		**s0045_VX**	0,62	**s0045_V6**	0,68
s0045_VZ	0,03	0,02	0,34	0,28	0,55	0,64	0,04	0,49		**s0045_VZ**	0,64	**s0045_V5**	0,69
s0045_aVF	0,05	0,14	0,38	0,41	0,30	0,36	0,06	0,38		**s0045_aVL**	0,70	**s0045_aVR**	0,73
s0045_aVL	0,07	0,03	0,36	0,58	0,47	0,70	0,11	0,58		**s0045_L1**	0,72	**s0045_L2**	0,73
s0045_aVR	0,09	0,09	0,36	0,39	0,73	0,52	0,10	0,55	0,46	**s0045_V4**	0,76	**s0045_VX**	0,89
					0,52	0,42					0,67		0,73

VIII - Similarity Measurements between different patients with different diagnosis - III

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0227_L1	0,17	0,25	0,42	0,63	0,32	0,77	0,17	0,606		**s0227_V1**	0,30	**s0227_VY**	0,13
s0227_L2	0,06	0,04	0,36	0,23	0,32	0,55	0,06	0,303		**s0227_VY**	0,43	**s0227_aVF**	0,18
s0227_L3	0,07	0,10	0,26	0,48	0,20	0,67	0,07	0,471		**s0227_V6**	0,44	**s0227_L3**	0,20
s0227_V1	0,03	0,00	0,15	0,39	0,22	0,30	0,04	0,304		**s0227_L2**	0,55	**s0227_aVL**	0,22
s0227_V2	0,08	0,02	0,32	0,47	0,32	0,60	0,07	0,466		**s0227_V2**	0,60	**s0227_V1**	0,22
s0227_V3	0,13	0,09	0,34	0,59	0,35	0,74	0,12	0,563		**s0227_aVF**	0,61	**s0227_VX**	0,27
s0227_V4	0,13	0,19	0,46	0,54	0,39	0,83	0,12	0,610		**s0227_L3**	0,67	**s0227_V6**	0,27
s0227_V5	0,07	0,03	0,26	0,47	0,35	0,80	0,07	0,541		**s0227_aVR**	0,70	**s0227_L2**	0,32
s0227_V6	0,05	0,01	0,25	0,35	0,27	0,44	0,05	0,355		**s0227_VX**	0,70	**s0227_V2**	0,32
s0227_VX	0,07	0,06	0,27	0,43	0,27	0,70	0,08	0,464		**s0227_aVL**	0,73	**s0227_L1**	0,32
s0227_VY	0,04	0,00	0,31	0,35	0,13	0,43	0,05	0,365		**s0227_V3**	0,74	**s0227_V3**	0,35
s0227_VZ	0,16	0,10	0,46	0,69	0,49	0,84	0,15	0,673		**s0227_L1**	0,77	**s0227_V5**	0,35
s0227_aVF	0,06	0,02	0,28	0,35	0,18	0,61	0,07	0,416		**s0227_V5**	0,80	**s0227_V4**	0,39
s0227_aVL	0,10	0,14	0,32	0,57	0,22	0,73	0,10	0,541		**s0227_V4**	0,83	**s0227_VZ**	0,49
s0227_aVR	0,09	0,10	0,39	0,35	0,61	0,70	0,10	0,567	0,483	**s0227_VZ**	0,84	**s0227_aVR**	0,61
					0,31	0,65					0,79		0,42

IX - Similarity Measurements between a patient and a healthy control - I

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0462_L1	0,12	0,11	0,34	0,56	0,28	0,68	0,11	0,51		**s0462_L2**	0,00	**s0462_L3**	0,10
s0462_L2	0,02	0,00	0,24	0,11	0,56	0,00	0,02	0,30		**s0462_L3**	0,00	**s0462_aVL**	0,12
s0462_L3	0,00	0,00	0,10	0,14	0,10	0,00	0,01	0,11		**s0462_V6**	0,00	**s0462_VY**	0,17
s0462_V1	0,02	0,00	0,17	0,34	0,28	0,24	0,03	0,28		**s0462_VY**	0,00	**s0462_aVF**	0,17
s0462_V2	0,09	0,30	0,45	0,57	0,52	0,70	0,08	0,60		**s0462_aVF**	0,00	**s0462_V1**	0,28
s0462_V3	0,09	0,19	0,48	0,52	0,74	0,81	0,06	0,69		**s0462_V5**	0,10	**s0462_L1**	0,28
s0462_V4	0,05	0,01	0,35	0,29	0,99	0,65	0,05	0,66		**s0462_V1**	0,24	**s0462_aVR**	0,47
s0462_V5	0,01	0,00	0,17	0,10	0,97	0,10	0,01	0,41		**s0462_aVR**	0,32	**s0462_V2**	0,52
s0462_V6	0,00	0,00	0,15	0,08	0,89	0,00	0,00	0,37		**s0462_aVL**	0,49	**s0462_L2**	0,56
s0462_VX	0,05	0,01	0,31	0,32	0,84	0,52	0,04	0,56		**s0462_VX**	0,52	**s0462_VZ**	0,57
s0462_VY	0,00	0,00	0,14	0,18	0,17	0,00	0,01	0,16		**s0462_VZ**	0,59	**s0462_V3**	0,74
s0462_VZ	0,07	0,04	0,41	0,47	0,57	0,59	0,08	0,54		**s0462_V4**	0,64	**s0462_VX**	0,84
s0462_aVF	0,01	0,00	0,13	0,10	0,17	0,00	0,01	0,13		**s0462_L1**	0,68	**s0462_V6**	0,89
s0462_aVL	0,03	0,00	0,25	0,41	0,12	0,49	0,04	0,38		**s0462_V2**	0,70	**s0462_V5**	0,97
s0462_aVR	0,07	0,03	0,38	0,29	0,47	0,32	0,08	0,39	0,41	**s0462_V3**	0,81	**s0462_V4**	0,99
					0,51	0,34					0,66		0,83

X - Similarity Measurements between a patient and a healthy control - II

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0303_L1	0,07	0,19	0,33	0,41	0,38	0,57	0,07	0,45		**s0303_L2**	0,00	**s0303_L3**	0,07
s0303_L2	0,01	0,00	0,16	0,04	0,89	0,00	0,01	0,36		**s0303_L3**	0,00	**s0303_aVF**	0,10
s0303_L3	0,01	0,01	0,14	0,15	0,07	0,00	0,02	0,12		**s0303_VY**	0,00	**s0303_VY**	0,13
s0303_V1	0,02	0,00	0,23	0,25	0,21	0,26	0,02	0,25		**s0303_aVF**	0,00	**s0303_aVL**	0,17
s0303_V2	0,02	0,01	0,30	0,26	0,25	0,32	0,03	0,29		**s0303_V6**	0,11	**s0303_V1**	0,21
s0303_V3	0,04	0,17	0,35	0,36	0,33	0,60	0,05	0,44		**s0303_aVR**	0,23	**s0303_V2**	0,25
s0303_V4	0,07	0,12	0,44	0,37	0,43	0,76	0,06	0,54		**s0303_V1**	0,26	**s0303_V3**	0,33
s0303_V5	0,01	0,00	0,18	0,06	0,76	0,37	0,00	0,34		**s0303_V2**	0,32	**s0303_VZ**	0,37
s0303_V6	0,00	0,00	0,18	0,06	0,50	0,11	0,00	0,25		**s0303_V5**	0,37	**s0303_L1**	0,38
s0303_VX	0,08	0,19	0,34	0,46	0,73	0,57	0,08	0,59		**s0303_aVL**	0,47	**s0303_V4**	0,43
s0303_VY	0,00	0,00	0,20	0,19	0,13	0,00	0,01	0,17		**s0303_L1**	0,57	**s0303_V6**	0,50
s0303_VZ	0,05	0,03	0,36	0,42	0,37	0,76	0,06	0,52		**s0303_VX**	0,57	**s0303_VX**	0,73
s0303_aVF	0,00	0,01	0,13	0,08	0,10	0,00	0,01	0,10		**s0303_V3**	0,60	**s0303_V5**	0,76
s0303_aVL	0,02	0,03	0,28	0,35	0,17	0,47	0,02	0,37		**s0303_V4**	0,76	**s0303_aVR**	0,87
s0303_aVR	0,03	0,00	0,33	0,11	0,87	0,23	0,03	0,48	0,35	**s0303_VZ**	0,76	**s0303_L2**	0,89
					0,41	0,33					0,62		0,70

XI - Similarity Measurements between a patient and a healthy control - III

	ED	DTW	Mi	Mah	WT	CC	DCT	Média (3 melhores métodos)	Média total	CC		WT	
s0311_L1	0,10	0,22	0,36	0,50	0,84	0,67	0,11	0,67		**s0311_VY**	0,00	**s0311_VY**	0,16
s0311_L2	0,03	0,05	0,22	0,13	0,72	0,07	0,03	0,36		**s0311_L2**	0,07	**s0311_aVF**	0,24
s0311_L3	0,06	0,04	0,35	0,53	0,30	0,66	0,07	0,51		**s0311_V6**	0,08	**s0311_V1**	0,29
s0311_V1	0,03	0,00	0,18	0,36	0,29	0,39	0,04	0,34		**s0311_V4**	0,21	**s0311_L3**	0,30
s0311_V2	0,11	0,04	0,49	0,62	0,73	0,75	0,10	0,70		**s0311_V5**	0,37	**s0311_V6**	0,33
s0311_V3	0,19	0,27	0,59	0,74	0,92	0,86	0,12	0,84		**s0311_V1**	0,39	**s0311_VZ**	0,50
s0311_V4	0,21	0,65	0,52	0,72	0,95	0,21	0,16	0,77		**s0311_aVF**	0,43	**s0311_VX**	0,52
s0311_V5	0,04	0,01	0,24	0,32	0,95	0,37	0,03	0,55		**s0311_aVR**	0,47	**s0311_aVL**	0,52
s0311_V6	0,01	0,00	0,20	0,13	0,33	0,08	0,01	0,22		**s0311_VX**	0,56	**s0311_L2**	0,72
s0311_VX	0,02	0,00	0,20	0,16	0,52	0,56	0,02	0,42		**s0311_L3**	0,66	**s0311_V2**	0,73
s0311_VY	0,00	0,00	0,12	0,16	0,16	0,00	0,01	0,15		**s0311_L1**	0,67	**s0311_L1**	0,84
s0311_VZ	0,16	0,15	0,45	0,71	0,50	0,86	0,12	0,69		**s0311_aVL**	0,68	**s0311_V3**	0,92
s0311_aVF	0,03	0,01	0,27	0,28	0,24	0,43	0,04	0,33		**s0311_V2**	0,75	**s0311_aVR**	0,95
s0311_aVL	0,07	0,14	0,37	0,57	0,52	0,68	0,08	0,59		**s0311_V3**	0,86	**s0311_V4**	0,95
s0311_aVR	0,07	0,05	0,31	0,25	0,95	0,47	0,07	0,57	0,51	**s0311_VZ**	0,86	**s0311_V5**	0,95
					0,59	0,47					0,75		0,89