

Electrocardiogram signals classification using discrete wavelet transform and support vector machine classifier

Youssef Toulni¹, Taoufiq Belhoussine Drissi², Benayad Nsiri³

^{1,2}Laboratory Industrial Engineering, Information Processing, and Logistics (GITIL), Faculty of Science Ain Chock, University Hassan II-Casablanca, Morocco

^{1,3}Research Center STIS, M2CS, National School of Arts and Crafts of Rabat (ENSAM), Mohammed V University in Rabat, Morocco

Article Info

Article history:

Received Nov 29, 2020

Revised Sep 10, 2021

Accepted Sep 28, 2021

Keywords:

Cardiovascular disease
Discrete wavelet transform
Electrocardiogram signal
Support vector machine

ABSTRACT

The electrocardiography allowed us to make a diagnosis of several cardiovascular diseases by representing the electrical activity of the heart over time; this representation is called the electrocardiogram (ECG) signal. In this study we have proposed a model based on the processing of the ECG signal by the wavelet decomposition using discrete wavelet transform (DWT). This decomposition firstly makes it possible to denoise the signal then to extract the statistical features from the approximation coefficients of the denoised signal and finally to classify the data obtained in a support vector machine (SVM) classifier with cross validation for more credibility. After having tested this model with different mother wavelets at different scales, the accuracies at the fourth scale are high and the best accuracy obtained is 87.50%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Youssef Toulni

Laboratory Industrial Engineering, Information Processing, and Logistics (GITIL)

Faculty of Science Ain Chock

Km 8 Route d'El Jadida, B.P 5366 Maarif Casablanca 20100 Morocco

Email: youssef.toulni@gmail.com

1. INTRODUCTION

Cardiovascular disease is a collection of irregularities affecting the heart; it is considered one of the most important causes of death in the world. According to the World Health Organization, there were approximately 17.9 million deaths in 2016; this big number represents 31% of deaths worldwide [1]. The lives of people with cardiovascular diseases are in constant danger, quick and effective diagnosis of these diseases can save a lot of lives. Several techniques in the medical field are used to diagnose cardiovascular disease, such as blood tests, coronary angiography, cardiac MRI, X-ray and electrocardiography. However, most of these techniques require medical assistance from experienced people, which is not always the case if we knew that almost 30% of cases with these diseases come from poor countries. Electrocardiography is a non-invasive detection technique based on recording the electrical activity of the heart over time [2], the signal obtained during this recording is called an electrocardiogram (ECG). The ECG signal is considered among the most widely used biomedical signals to detect heart problems, ECG signal contains a large number of information that can be of great interest in the detection and diagnosis of many heart diseases [3] which appears in some distortions of the signal shape. Despite the advantages of the use of ECG signals, there are many limitations of this technique; the difficulty of interpreting the signal and the lack of experienced personnel are among the constraints most encountered during identification of ECG signals; also, ECG signal contains various unwanted noises that prevent the correct extraction of useful and necessary

information for classification [2], [4]. Hence, finding solutions to solve these problems has become a necessity. The field of medical engineering try the challenge by developing models capable of processing ECG signals and identifying any abnormalities present in the signal [5], thus, signal processing is required; signal processing techniques analyze efficiently various kinds of signals, especially ECG signals. The aim of ECG signal processing is to extract features to distinguish between normal signals and those representing abnormalities.

Several techniques based on signal processing have been developed over years for processing ECG signals. Ahlstrom and Tompkins have used digital filters for real time ECG signal processing [6] to denoise the signal and detect the QRS complex. Hargittai [7] used Savitzky-Golay Least-Squares Polynomial filters to preserve the details of the signal, Francisco *et al.* [8] process ECG signal using principal component analysis (PCA), Haque *et al.* [9] use adaptive filtering algorithms, while Ahmed *et al.* [10] use a method based on cross correlation theory. Gustavo *et al.* [11] propose a comparison between several methods to remove baseline wander. Also the choice of features and the method used for its extraction affect directly the quality of signal characterization, some techniques are based on the extraction of morphological features such as the detection of the QRS complex proposed by Jiapu [12], the calculation of R-R intervals and peaks detection mentioned in the works of Shanti Chandra *et al.* [13] and Priyanka [14], other methods are based on the extraction of the statistical ones or even a mixture of DWT and statistical features, Abdullah *et al.* [15], or morphological and statistical features, Sahoo *et al.* [16].

The discrete wavelet transform (DWT) is a very powerful tool in the field of signal processing [17]-[20], this technique gives satisfying results in the processing of the noise which affects the signal that can allows us to reconstruct a denoised signal, also the DWT allows to extract different features that characterize the signal. The interest of the discrete wavelet transforms (DWT) pushes us to better exploit this technique in the processing of ECG signal [21], [22]. Also, machine learning techniques are constantly evolving; this evolution is reflected in the use of these techniques in several areas [23], more precisely in the classification and identification of signals. The combination of signal processing techniques and machine learning models gives us promising results and the performance of each model changes depending on the chosen algorithm

The purpose of this modest work is to establish a characterization model of ECG signals able to differentiate between normal and abnormal signals, the adopted model is based on the extraction of statistical features from the approximation coefficients obtained by the wavelet decomposition of the signal and classify these features using an SVM classifier; this model was approved after the choice of the best wavelet and the best scale increasing its accuracy, the recordings of the used ECG signals are taken from the MIT-BIH arrhythmia database [24], [25]. It's composed of 48 recordings of different patients; these patients are classified into two categories: A healthy and a sick one. After the processing of the ECG signals and the extraction of features a classification of patients to healthy or sick ones is needed. To do this we have chosen a support vector machine (SVM) classifier who seems most appropriate to do this task [26]-[28].

2. PRELIMINARIES

2.1. Wavelet analysis

Wavelet analysis is one of the powerful tools in signal processing, it is considered as a technique which aims to solve the problem of non-stationary signals. This notion was introduced in the 20th century by Haar who constructed the simplest wavelet, and then developed in the 1980s through the research work of Mallat [18], [19], [29]. One of the powerful things of this approach is that it allows a signal to be analyzed in time and frequency, which makes it very useful in extracting the various information contained in that signal.

The wavelet transform decomposes a signal $x(t)$ using a series of wavelets $\Psi_{a,b}$ which derive from a mother wavelet by a displacement in time carried out by a translation and a dilation or compression by scaling of the mother wavelet, this series of wavelets is defined as [30]:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (1)$$

where a and b are respectively the scaling and the translation coefficients ψ being the mother wavelet which must verify the following condition [17]:

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0 \quad (2)$$

The wavelet transform exists in two principal forms the CWT and the discrete wavelet transform (DWT), thus we define the continuous wavelet transform (CWT) of a signal $x(t)$ as [18]:

$$c_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-b}{a}\right) dt \quad (3)$$

ψ^* represents the complex conjugate of the wavelet $\psi(t)$.

The computation of this continuous wavelet transform seems sometimes difficult and time consuming. The discrete wavelet transform give a solution to this problem. To do this we use an algorithm (proposed by Mallat) [18] to calculate the discrete wavelet coefficients, this decomposition consists in passing the signal through a succession of complementary low-pass $h[n]$ and high-pass $g[n]$ filters (in stages), the low-frequency components resulting from the low-pass filters $h[n]$ represent approximation coefficients and those from the high frequency $g[n]$ filter are called the detail coefficients [31] as shown in Figure 1.

2.2. Features extraction

The choice of the nature, number and method of features extraction is considered one of the decisive steps in the characterization of a signal, since each of these parameters contain various information which influence on the efficiency of classification. In the literature many are the features and the methods used to characterize a signal especially for the ECG signals, there are the morphological features such as the PR, RR and QRS intervals, which are determined basically by the detection of the QRS complex [12]-[14]. There are also statistical features such as maximum, minimum, mean, variance and standard deviation [15], [16], [32], [33].

In this related work we will focus on the following features:

The mean value (4):

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \tag{4}$$

The root mean square (5):

$$rms = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \tag{5}$$

The variance (6):

$$v = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \tag{6}$$

Standard deviation (7):

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \tag{7}$$

Kurtosis (8):

$$\varphi = \frac{1}{N} \frac{\sum_{i=1}^N (x_i - \mu)^3}{\sigma^3} \tag{8}$$

Wavelet entropy (9):

$$H = -\frac{1}{N} \sum_{i=1}^N x_i^2 \log(x_i^2) \tag{9}$$

where N is the length of the signal x

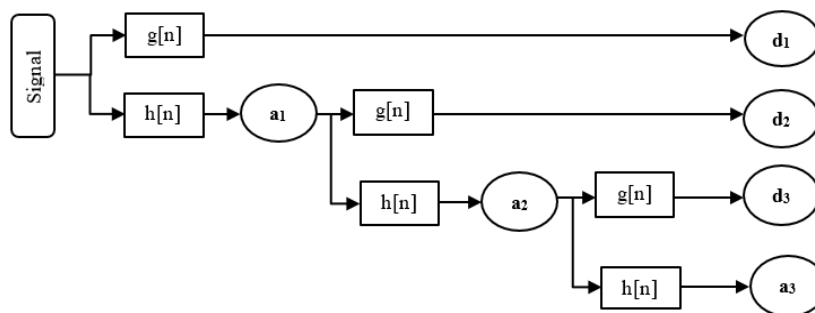


Figure 1. Signal decomposition using DWT at 3rd level of scale

2.3. Classification

The chosen classification technique is support vector machine (SVM), SVM is a machine learning technique that allows binary classifications. The purpose of this method is to separate the features of the signal into different categories [27], [28]. The purpose of this method is to separate the features of the signal into different categories, it's based on the construction of a separation area between the different classes of the learning set $\{x_i, y_i\}$, with $y_i \in \{-1, 1\}$ and $i = 1, 2, \dots, n$ the lines that delimit this area are called a hyperplane which is defined by [28], [29]:

$$y = wx + b \quad (10)$$

w represents a normal vector on the hyperplane and b is the bias.

$$\text{Thus, the area where we have: } y_i = -1 \text{ implies that } wx_i + b \leq -1 \quad (11)$$

$$\text{Similarly, for } y_i = +1 \text{ we have } wx_i + b \geq 1 \quad (12)$$

$$\text{In all cases we have } y_i(wx_i + b) \geq 1 \quad (13)$$

$$\text{Good separation aims at increasing the width of the margin between hyperplanes, this is equivalent to minimizing } \frac{\|w\|^2}{2} \text{ knowing that } y_i(wx_i + b) - 1 \geq 0. \quad (14)$$

If we can't separate linearly these classes, then the optimization problem becomes [26]:

$$\begin{cases} \min\left(\frac{\|w\|^2}{2} + C \sum_{i=1}^n \xi_i\right) \\ \text{avec } y_i(wx_i + b) - \xi_i \geq 1 \end{cases} \quad (15)$$

3. METHODOLOGY

In this work we want to establish a simple model that allows:

- ECG signal processing.
- Extraction of the features that characterize the signal, these features will be calculated by applying (4)-(9).
- The classification of signals processed into normal (healthy patients) and other abnormal signals (the sick ones) using a linear SVM classifier.

Processing techniques and extraction methods are based on discrete wavelet analysis; Figure 2 shows the process of identifying ECG signals.

3.1. Database

The ECG signals are taken from the MIT BIH data base of physionet [24], [25], this database contains 48 recordings each recording is half hour length, the recordings are sampled at 360 Hz, in this study we analyzed duration of one minute for each signal.

3.2. ECG signal processing

The recordings of ECG are not protected from noises; these noises cause significant perturbations during classification and diagnosis. For ECG signals, there are different sources of noises, noises of technical origins such powerline interference (50 or 60 Hz noise from mains supply) or bad wiring and others of physical origins generated by the physical activity of the body such electromyographic (EMG) noise, baseline wander (low frequencies) which are considered among the most important [34]. The steps of signal denoising are shown in Figure 3.

In this stage of processing, we tried to eliminate these noises using the discrete wavelet decomposition. After the decomposition of the signal into eight levels we notice that the noise (high frequencies) is located in the detail coefficients d_1 and d_2 . For the baseline wander we notice that the approximation coefficient A_8 corresponds to the frequency interval of these fluctuations 0-0.5 Hz [20]; Figure 4 shows detail and approximation coefficients resulting from the wavelet decomposition of recording 121 m, and Figure 5 shows part of the same recording before, as shown in Figure 5(a) and after, as shown in Figure 5(b) denoising, the mother wavelet used in the decomposition is Symlet [35].

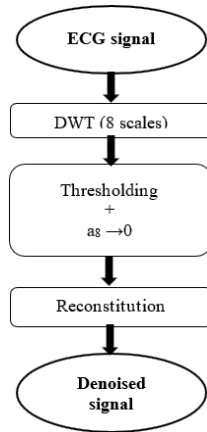


Figure 2. Signal denoising process

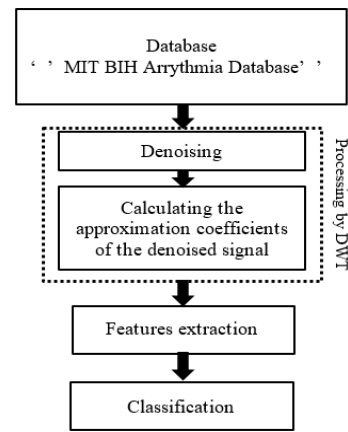


Figure 3. Diagram of the proposed method

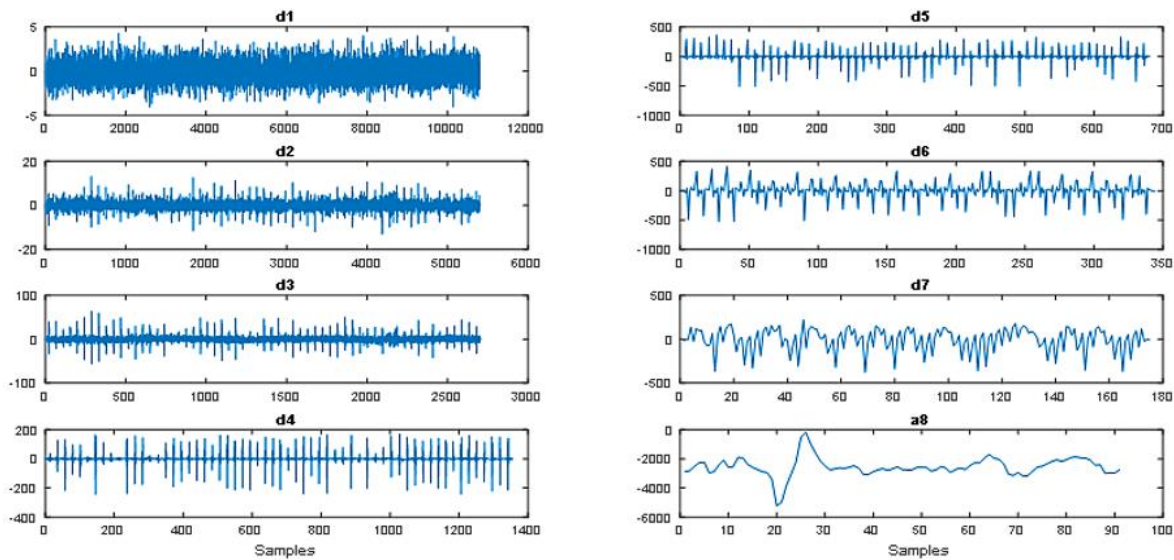


Figure 4. Wavelet decomposition of ECG signal (record No 121) d1, d2, d3, d4, d5, d6, d7 and a8

3.3. Features extraction

After denoising the ECG signal, we decompose the denoised signal using DWT by using different types of wavelets to analyze the signal, and then the approximation coefficients are calculated up to scale 8 as shown in Figure 6. After the extraction of the approximation coefficients, the features mentioned above (4)-(8) are calculated from the approximation coefficients $a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8$, thus we create for each level a data set containing the features for all the records.

3.4. Classification

The resulting dataset is introduced in a SVM classifier and it's divided into two subsets; the first will be served for training the model, while the second will be designed to test the performance of the model. The choice of the training and test subsets is crucial and it can causes many problems which influence the effectiveness of the model such as overfitting and underfitting [36]; to limit this kinds of problems we use k fold cross validation to select this subsets [37], [38], this method consists in dividing the data set into parts of number of k , one of these parts is chosen as a subset of test and the other $k - 1$ parts as a training subset, this process is repeated k times and each time a different part is taken to do the test. The precision of the model is obtained by averaging the precision of each iteration [39]-[41]. In this work we took $k = 8$, that is to say that $7/8$ of the data set is taken for training and $1/8$ is considered as a test subset. The next step will be devoted to measuring the performance of the model for the different types of mother wavelets at different scales (up to

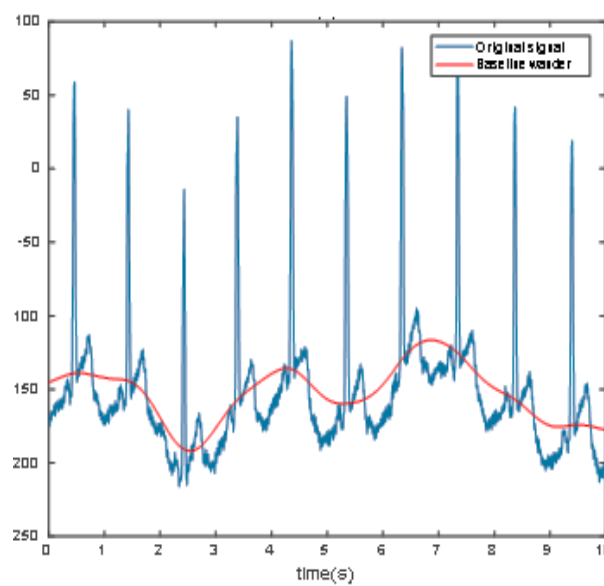
the 8th scale); the accuracy (Acc), the sensitivity (Sen) and the specificity (Spe) are the quantities that inform us about the performance of each model with [29]:

$$Acc = \frac{TN+TP}{TP+TN+FP+FN} \times 100 \quad (16)$$

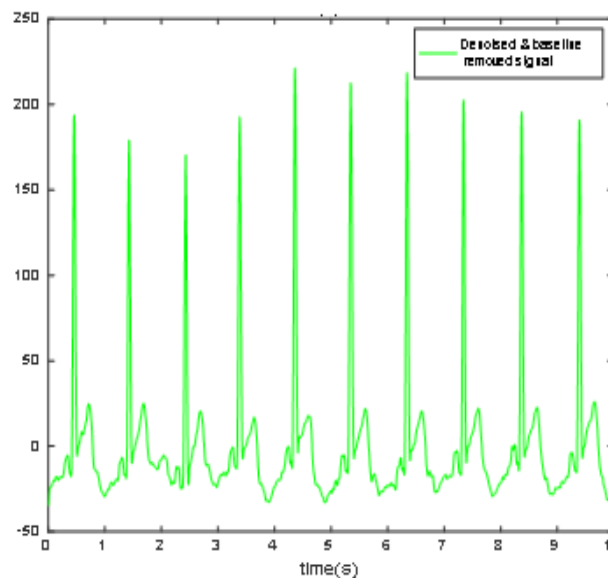
$$Sen = \frac{TP}{TP+FN} \times 100 \quad (17)$$

$$Spe = \frac{TN}{TN+FP} \times 100 \quad (18)$$

TP : True positive (represents normal signal who were correctly classified)
 TN : True negative (represents abnormal signal who were correctly classified)
 FP : False positive (represents normal signal who were incorrectly classified)
 FN : False negative (represents abnormal signal who were incorrectly classified)



(a)



(b)

Figure 5. Record no 121 before and after denoising; (a) original signal and baseline wander, (b) signal after removing noise and baseline wander

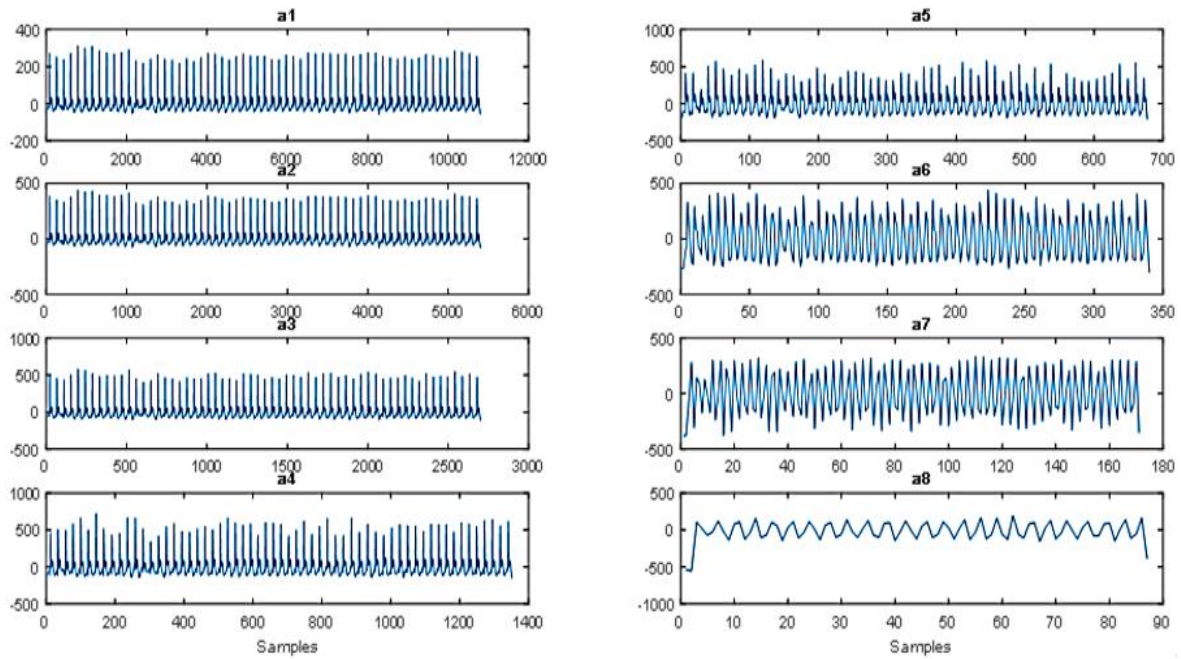


Figure 6. Approximation coefficients a1, a2, a3, a4, a5, a6, a7, a8

4. RESULTS

Table 1 as shown in appendix, summarizes the results obtained for different wavelet families at the first eight scales of the approximation coefficients. As we can see, the best accuracy is generally obtained at the fourth scale. The study carried out by Siti [42] which is based on the calculation of statistical features from the detail coefficients at the 4th, 5th and 6th and the KNN as a classification method reaches an accuracy of 71% using the sym7 wavelet; while in this work we had reached an accuracy of 81.67% using only the statistical features extracted from the approximation coefficient a4 for the same wavelet and an SVM classifier; ;also in the study of Siti [42] we reached an accuracy of 85% as the best accuracy by calculating the MFCC coefficients, while in this work we found 87.50% as better accuracy by adopting only the coif5 wavelet as the mother wavelet as shown in Table 2.

Table 2. Comparison results

		Method		Classifier	Accuracy
Siti [42]	DWT	D4, D5, D6	Haar	KNN	68%
			Sym 7		71%
		MFCC	85%		
Proposed method	DWT	A4	Sym7	SVM	81,67%
			Coif 5		87,50%

5. CONCLUSION

In this article we focused on the classification of ECG signals, we were then able to establish a model that allows to process and classify these signals. ECG signals were processed using DWT discrete wavelet analysis, the discrete wavelet transform allows the calculation of approximation coefficients which are used to extract the features for different scales, the dataset which consists of these parameters is fed into an SVM classifier with cross validation in order to distinguish between a normal and an abnormal signal. This model was tested with different mother wavelets and at different scales to reach an accuracy of 87.50%, which was the best. To conclude, the choice of the mother wavelet, the scale of the decomposition as well as the size of the training and test sets have a considerable influence on the accuracy of the model.

APPENDIX

Table 1. Accuracy, sensitivity and specificity of the different wavelets at the first eight scales (*continue*)

Wavelet	Classifier performance	Approximation coefficients scale level							
		1	2	3	4	5	6	7	8
db2	Accuracy	72.08	71.67	70.83	79.58	70.42	68.33	60.42	56.84
	Sensitivity	75.83	76.67	77.50	90.00	81.67	75.42	68.33	64.50
	Specificity	68.33	66.67	64.17	69.17	59.17	60.00	52.50	50.00
db3	Accuracy	71.25	72.92	72.92	77.92	70.42	65.42	58.75	53.51
	Sensitivity	74.17	76.67	79.17	86.67	79.17	74.58	69.17	53.83
	Specificity	68.33	69.17	66.67	69.17	61.67	53.33	48.33	53.33
db4	Accuracy	72.08	71.25	74.17	75.83	68.33	64.17	58.75	51.97
	Sensitivity	75.00	75.83	80.00	86.67	78.33	72.50	64.17	52.54
	Specificity	69.17	66.67	68.33	65.00	58.33	54.17	53.33	51.67
db5	Accuracy	72.92	70.83	71.25	77.08	68.33	65.00	60.00	39.99
	Sensitivity	79.17	76.67	79.17	89.17	80.00	74.17	69.17	55.53
	Specificity	66.67	65.00	63.33	65.00	56.67	55.00	50.83	29.17
db6	Accuracy	73.33	72.50	72.92	76.25	65.00	65.83	59.58	41.08
	Sensitivity	80.00	78.33	80.00	87.50	74.17	75.42	73.33	68.19
	Specificity	66.67	66.67	65.83	65.00	55.83	53.33	45.83	18.33
db7	Accuracy	70.83	73.33	77.50	72.08	67.08	63.33	60.00	41.68
	Sensitivity	78.33	79.17	85.00	81.67	75.83	75.00	74.17	69.26
	Specificity	63.33	67.50	70.00	62.50	58.33	49.17	45.83	18.33
db8	Accuracy	73.33	74.58	73.75	75.83	67.50	63.75	59.58	49.58
	Sensitivity	79.17	81.67	80.83	86.67	73.33	75.42	67.50	68.59
	Specificity	67.50	67.50	66.67	65.00	61.67	48.33	51.67	30.00
sym2	Accuracy	72.92	71.67	72.92	77.50	70.00	67.92	60.83	54.49
	Sensitivity	76.67	75.00	77.50	85.00	80.00	76.25	71.67	65.10
	Specificity	69.17	68.33	68.33	70.00	60.00	57.50	50.00	45.83
sym3	Accuracy	71.67	71.25	72.08	77.08	70.83	64.17	60.83	56.30
	Sensitivity	76.67	75.00	77.50	89.17	80.00	72.92	72.50	58.28
	Specificity	66.67	67.50	66.67	65.00	61.67	52.50	49.17	55.00
sym4	Accuracy	72.92	71.25	71.67	81.67	67.50	72.92	65.83	45.58
	Sensitivity	75.83	74.17	77.50	95.00	80.83	79.58	70.83	64.00
	Specificity	70.00	68.33	65.83	68.33	54.17	64.17	60.83	30.83
sym5	Accuracy	70.00	71.67	72.50	82.08	71.25	72.08	63.33	44.36
	Sensitivity	75.83	75.00	75.83	93.33	85.83	81.25	68.33	59.91
	Specificity	64.17	68.33	69.17	70.83	56.67	60.83	58.33	31.67
sym6	Accuracy	71.67	74.17	71.25	79.17	71.67	72.50	61.67	46.51
	Sensitivity	76.67	77.50	76.67	91.67	85.00	81.25	69.17	61.18
	Specificity	66.67	70.83	65.83	66.67	58.33	61.67	54.17	35.00
sym7	Accuracy	71.67	72.50	72.50	81.67	72.08	70.83	58.75	42.00
	Sensitivity	76.67	76.67	75.83	91.67	85.83	81.25	66.67	58.77
	Specificity	66.67	68.33	69.17	71.67	58.33	57.50	50.83	27.50
sym8	Accuracy	73.33	72.92	73.75	79.17	68.75	67.50	60.00	42.98
	Sensitivity	78.33	75.83	77.50	90.00	80.00	77.50	69.17	63.21
	Specificity	68.33	70.00	70.00	68.33	57.50	55.83	50.83	25.83
coif1	Accuracy	72.08	72.50	71.67	79.17	71.67	72.92	66.67	51.72
	Sensitivity	75.00	75.83	75.00	88.33	82.50	80.00	70.83	65.32
	Specificity	69.17	69.17	68.33	70.00	60.83	64.17	62.50	40.00
coif2	Accuracy	72.08	72.50	71.67	79.17	71.67	72.92	66.67	51.72
	Sensitivity	75.00	75.83	75.00	88.33	82.50	80.00	70.83	65.32
	Specificity	69.17	69.17	68.33	70.00	60.83	64.17	62.50	40.00
coif3	Accuracy	72.08	72.50	71.67	79.17	71.67	72.92	66.67	51.72
	Sensitivity	75.00	75.83	75.00	88.33	82.50	80.00	70.83	65.32
	Specificity	69.17	69.17	68.33	70.00	60.83	64.17	62.50	40.00
coif4	Accuracy	74.48	73.96	75.00	85.42	71.88	72.40	60.42	49.40
	Sensitivity	81.25	81.25	83.33	97.92	83.33	79.69	72.92	63.96
	Specificity	67.71	66.67	66.67	72.92	60.42	62.50	47.92	38.54
coif5	Accuracy	75.00	73.96	75.00	87.50	73.96	73.96	55.21	42.80
	Sensitivity	82.29	80.21	82.29	100.00	84.38	82.29	65.63	55.31
	Specificity	67.71	67.71	67.71	75.00	63.54	63.54	44.79	33.33
bior1.1	Accuracy	63.75	65.83	67.50	70.42	69.17	64.58	62.92	61.70
	Sensitivity	68.33	69.17	72.50	79.17	85.00	72.92	75.83	73.63
	Specificity	59.17	62.50	62.50	61.67	53.33	53.33	50.00	50.83
bior1.3	Accuracy	71.67	71.25	73.33	79.17	69.58	70.42	60.83	47.44
	Sensitivity	75.00	75.00	76.67	89.17	83.33	76.25	70.00	62.57
	Specificity	68.33	67.50	70.00	69.17	55.83	63.33	51.67	35.00
bior1.5	Accuracy	70.42	71.25	71.25	78.75	72.50	74.17	60.42	40.99
	Sensitivity	74.17	75.00	75.83	90.00	85.83	84.17	66.67	61.32
	Specificity	66.67	67.50	66.67	67.50	59.17	61.67	54.17	22.50

Table 1. Accuracy, sensitivity and specificity of the different wavelets at the first eight scales

Wavelet	Classifier performance	Approximation coefficients scale level							
		1	1	1	1	1	1	1	1
bior2.2	Accuracy	72.08	72.08	69.58	72.50	77.08	69.58	66.67	54.28
	Sensitivity	75.00	75.00	74.17	80.00	89.17	75.83	64.17	61.56
	Specificity	69.17	69.17	65.00	65.00	65.00	61.67	69.17	47.50
bior2.4	Accuracy	71.67	70.83	70.42	77.50	70.42	72.08	60.83	49.27
	Sensitivity	76.67	74.17	74.17	85.83	80.83	76.25	69.17	62.49
	Specificity	66.67	67.50	66.67	69.17	60.00	66.67	52.50	38.33
bior2.6	Accuracy	70.00	71.25	70.42	79.17	72.50	74.17	61.25	46.13
	Sensitivity	75.00	75.83	75.83	88.33	83.33	80.83	67.50	57.49
	Specificity	65.00	66.67	65.00	70.00	61.67	65.83	55.00	37.50
bior2.8	Accuracy	70.42	71.67	70.42	78.33	73.75	72.92	62.08	43.62
	Sensitivity	75.00	75.83	72.50	86.67	85.00	81.25	70.00	64.06
	Specificity	65.83	67.50	68.33	70.00	62.50	62.50	54.17	26.67
rbio1.1	Accuracy	63.33	64.17	65.83	69.17	70.42	62.50	64.58	60.58
	Sensitivity	67.50	67.50	69.17	78.33	85.00	72.50	78.33	76.17
	Specificity	59.17	60.83	62.50	60.00	55.83	49.17	50.83	46.67
rbio1.3	Accuracy	70.00	72.50	72.92	77.08	72.92	70.42	66.67	60.36
	Sensitivity	75.00	75.00	78.33	85.00	82.50	77.50	80.83	75.53
	Specificity	65.00	70.00	67.50	69.17	63.33	61.67	52.50	48.33
rbio1.5	Accuracy	69.17	71.67	72.92	78.33	74.17	72.08	64.17	52.65
	Sensitivity	73.33	75.83	78.33	87.50	81.67	80.42	75.83	70.90
	Specificity	65.00	67.50	67.50	69.17	66.67	61.67	52.50	37.50
rbio2.2	Accuracy	72.92	72.50	75.83	76.67	70.42	70.00	57.08	52.33
	Sensitivity	75.83	77.50	84.17	85.83	81.67	80.83	62.50	54.11
	Specificity	70.00	67.50	67.50	67.50	59.17	56.67	51.67	50.83
rbio2.4	Accuracy	71.67	72.08	75.83	78.33	68.33	68.75	56.67	44.21
	Sensitivity	76.67	77.50	83.33	87.50	78.33	73.33	66.67	56.94
	Specificity	66.67	66.67	68.33	69.17	58.33	62.50	46.67	34.17
rbio2.6	Accuracy	70.42	71.25	77.50	75.42	70.00	68.75	54.17	40.63
	Sensitivity	75.83	76.67	85.83	83.33	79.17	75.83	65.83	56.03
	Specificity	65.00	65.83	69.17	67.50	60.83	60.83	42.50	28.33
rbio2.8	Accuracy	72.92	74.17	78.33	79.17	72.92	65.83	50.83	45.07
	Sensitivity	77.50	81.67	85.83	90.83	81.67	73.33	65.83	61.33
	Specificity	68.33	66.67	70.83	67.50	64.17	55.83	35.83	30.83

REFERENCES

- [1] "WHO," 2017. [Online]. Available: [https://www.who.int/en/newsroom/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/en/newsroom/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- [2] R. Kher, "Signal processing techniques for removing noise from ECG signals," *Journal of Biomedical Engineering*, vol. 1 pp. 1-9, 2019, doi: 10.17303/jber.2019.3.101.
- [3] S. M. Anwar, M. Gul, M. Majid, and M. Alnowami, "Arrhythmia classification of ECG signals using hybrid features," *Comp. and Mathematical Methods in Medicine*, vol. 2018, pp. 1-8, 2018, doi: 10.1155/2018/1380348
- [4] X.-K. Wan, H. Wu, F. Qiao, F.-C. Li, Y. Li, Y.-W. Yan, and J.-X. Wei, "Electrocardiogram baseline wander suppression based on the combination of morphological and wavelet transformation based filtering," *Computational and Mathematical Methods in Medicine*, vol. 2019, pp. 1-7, 2019, Art. no. 7196156, doi: 10.1155/2019/7196156
- [5] A. Rahman, N. Athilah, and J. Asral, "Biomedical health monitoring system design and analysis," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 13, no. 3, pp. 1056-1064, 2019, doi: 10.11591/ijeecs.v13.i3.pp1056-1064.
- [6] M. L. Ahlstrom and W. J. Tompkins, "Digital filters for real-time ECG signal processing using microprocessors," *IEEE Transactions on Biomedical Engineering*, vol. 32, no. 9, pp. 708-713, 1985, doi: 10.1109/TBME.1985.325589.
- [7] S. Hargittai, "Savitzky-Golay least-squares polynomial filters in ECG signal processing," *Computers in Cardiology*, vol. 32, pp. 763-766, 2005, doi: 10.1109/CIC.2005.1588216.
- [8] F. Castells, P. Laguna, L. Sörnmo, B. Andreas, and J. M. Roig, "Principal component analysis in ecg signal processing," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, no. 1, 2007, Art. No. 074580, doi: 10.1155/2007/74580
- [9] Z.-ul-Haque, R. Qureshi, M. Nawaz, F. Y. Khuhawar, N. Tunio, and M. Uzair, "Analysis of ECG signal processing and filtering algorithms," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 10, no. 3, 2019, doi: 10.14569/IJACSA.2019.0100370
- [10] S. Ahmed, A. H.-Alnaqbi, M. Al Hemaury, and M. Al Ahmad, "ECG abnormality detection algorithm," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 8, pp. 215-219, 2018, doi: 10.14569/IJACSA.2018.090827
- [11] G. Lenis, N. Pilia, A. Loewe, W. H. Walther Schulze, and O. Dössel, "Comparison of baseline wander removal techniques considering the preservation of ST changes in the Ischemic ECG: A simulation study," *Computational and Mathematical Methods in Medicine*, vol. 2017, pp. 1-13, 2017, Art. No. 9295029, doi: 10.1155/2017/9295029

- [12] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230-236, 1985, doi: 10.1109/TBME.1985.325532.
- [13] S. Chandra, A. Sharma, and G. K. Singh, "Feature extraction of ECG signal," *Journal of Medical Engineering and Technology*, vol. 42, no. 4, pp. 306-316, 2018, doi: 10.1080/03091902.2018.1492039
- [14] M. Priyanka, "Detection and processing of the R Peak," *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering (IJREEICE)*, vol. 6, no. 11, pp. 36-44, 2018, doi: 10.17148/IJREEICE.2018.6116.
- [15] A. Chashmi and M. Amirani, "An efficient and automatic ECG arrhythmia diagnosis system using DWT and HOS features and entropy- based feature selection procedure," *Journal of Electrical Bioimpedance*, vol. 10, no. 1, pp. 47-54, 2017, doi: 10.2478/joeb-2019-0007
- [16] S. Sahoo, M. Mohanty, S. Behera, and S. K. Sabut, "ECG beat classification using empirical mode decomposition and mixture of features," *Journal of Medical Engineering & Technology*, vol. 41, no. 8, pp. 652-661, 2017, doi: 10.1080/03091902.2017.1394386
- [17] H. Feng, J. Wang, Y. Li, and J. Chen, "Wavelet theory and application summarizing," *International Conference on Information Computing and Applications ICICA 2011*, vol 7030, 2011, pp: 337-343, doi: 10.1007/978-3-642-25255-6_43
- [18] S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, 1989, doi: 10.1109/34.192463.
- [19] S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 617-643, 1992, doi: 10.1109/18.119727.
- [20] O. Hicham, S. Eddarouich, A. Bourouhou, and M. Timouyas, "Comparison between SVM and KNN classifiers for iris recognition using a new unsupervised neural approach in segmentation," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no 3, pp. 429-438, 2020, doi: 10.11591/ijai.v9.i3.pp429-438.
- [21] I. Houamed, L. Saidi, and F. Srairi, "ECG signal denoising by fractional wavelet transform thresholding," *Research on Biomedical Engineering*, vol. 36, pp. 349-360, 2020, doi: 10.1007/s42600-020-00075-7
- [22] O. El B'charri, R. Latif, K. Elmansouri, A. Abenaou, and W. Jenkalm "ECG signal performance de-noising assessment based on threshold tuning of dual-tree wavelet transform," *Bio-Medical Engineering OnLine*, vol. 16, no. 1, 2017, Art. no. 26, doi: 10.1186/s12938-017-0315-1
- [23] A. Aliyu, M. Musa, and S. Usman, "Machine learning for plant disease detection: An investigative comparison between support vector machine and deep learning," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, pp. 670-683, 2020, doi: 10.11591/ijai.v9.i4.pp670-683.
- [24] B. G. Moody and G. R. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45-50, 2001, doi: 10.1109/51.932724.
- [25] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, C. P. Ivanov, R. Mark, and E. H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215-e220, 2000, doi: 10.1161/01.cir.101.23.e215.
- [26] S. Ibrahim, N. A. Zulkifli, N. Sabri, and A. Amilah, M. R. Mohd Noordin, "Rice grain classification using multi-class support vector machine (SVM)," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 8, no. 3, pp. 215-220, 2019, doi: 10.11591/ijai.v8.i3.pp215-220
- [27] A. Shmilovici, O. Maimon, and L. Rokach, "Data mining and knowledge discovery handbook," *Springer, Boston, MA*, 2009, doi: 10.1007/978-0-387-09823-4_12.
- [28] K. S. Parikh and T. P. Shah, "Support vector machine-a large margin classifier to diagnose skin illnesses," *Procedia Technology*, vol. 23, pp. 369-375, 2016, doi: 10.1016/j.protcy.2016.03.039.
- [29] T. B. Drissi, Z. Soumaya, N. Benayad, and A. Abdelkrim, "Diagnosis of Parkinson's disease based on wavelet transform and mel frequency cepstral coefficients," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 3, 2019, doi: 10.14569/IJACSA.2019.0100315.
- [30] I. Hammami, L. Salhi, and S. Labidi, "Voice pathologies classification and detection using EMD-DWT analysis based on higher order statistic features," *IRBM*, vol. 41, no. 3, pp. 161-171, 2020, doi: 10.1016/j.irbm.2019.11.004.
- [31] K. Surekha and B. P. Patil, "ECG Signal compression using the high frequency components of wavelet transform," *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 3, pp. 311-315, 2016, doi: 10.14569/IJACSA.2016.070344.
- [32] A. Darmawahyuni *et al.*, "Delineation of electrocardiogram morphologies by using discrete wavelet transforms," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 22, no. 1, pp. 159-167, doi: 10.11591/ijeecs.v22.i1.pp159-167.
- [33] E. A. Nidhal and F. Zahraa, "Detection and recognition of brain tumor based on DWT, PCA and ANN," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 18, no. 1, pp. 56-63, 2020, doi: 10.11591/ijeecs.v18.i1.pp56-63.
- [34] A. Subasi, "Chapter 4-feature extraction and dimension reduction, practical guide for biomedical signals analysis using machine learning techniques," *Academic Press*, 2019, pp. 193-275, doi: 10.1016/B978-0-12-817444-9.00004-0.
- [35] S. Seung Won, K. S. Kim, C. G. Song, J. W. Lee, J. H. Kim, and G. W. Jeung, "Removal of baseline wandering in ECG signal by improved detrending method," *Bio-medical Materials and Engineering*, vol. 26, no. 1, pp. S1087-S1093, 2015, doi: 10.3233/BME-151405.
- [36] B. Krishna, "Research of machine learning algorithms using K-Fold cross validation," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6S, pp. 215-218, 2021.

- [37] G. Garg, V. Singh, J. R. P. Gupta, and A. P. Mittal, "Optimal algorithm for ECG denoising using discrete wavelet transforms," *2010 IEEE International Conference on Computational Intelligence and Computing Research*, Coimbatore, 2010, pp. 1-4, doi: 10.1109/ICCC.2010.5705839.
- [38] W. N. W. Md Adnan, N. Y. Dahlan, and I. Musirin, "Development of option c measurement and verification model using hybrid artificial neural network-cross validation technique to quantify saving," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 1, pp. 25-32, 2020, doi: 10.11591/ijai.v9.i1.pp25-32.
- [39] N. Seman, and N. A. Razmi, "Machine learning-based technique for big data sentiments extraction," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 9, no. 3, pp. 473-479, 2020, doi: 10.11591/ijai.v9.i3.pp473-479.
- [40] X. Zheng, C. Yuxin, L. Zhonghao, H. Ming, and H. Jianjun, "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation," *Computational Materials Science*, vol. 171, 2020, Art. No. 109203, doi: 10.1016/j.commatsci.2019.109203.
- [41] S. Yadav and S. Shukla, "Analysis of k-Fold Cross-validation over hold-out validation on colossal datasets for quality classification," *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, Bhimavaram, 2016, pp. 78-83, doi: 10.1109/IACC.2016.25.
- [42] S. A. Alodia Yusuf and R. Hidayat, "MFCC feature extraction and KNN classification in ECG signals," *2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, Semarang, Indonesia, 2019, pp. 1-5, doi: 10.1109/ICITACEE.2019.8904285.

BIOGRAPHIES OF AUTHORS



Youssef Toulmi was born in Casablanca, Morocco in 1983. Graduated with a Masters in computer science and scientific instrumentation in high energy physics from the Ain Chok Faculty of Sciences. Hassan II University - Casablanca, Morocco, in 2019, he is a student-researcher at the Industrial, Computer and Logistics Engineering Research Laboratory (GITIL). Faculty of Sciences Ain Chok, Hassan II University - Casablanca, Morocco. He is interested in the processing of biomedical signals.



Taoufiq Belhoussine Drissi was born in Oujda, Morocco in 1978 received the Ph.D. degree in acoustics in 2009 at the university of le Havne (France) Since 2011 he has been an assistant professor at the sciences faculty of Ain chock university Hassan II, Casablanca. His scientific interest lies in the research of nondestructive testing and the signal treatment.



Benayad Nsiri held MBI degree in computer sciences from Telecom Bretagne, in 2005, and Ph.D. degree in signal processing from Telecom Bretagne, in 2004. He received D.E.A (French equivalent of M.Sc. degree) in electronics from the Occidental Bretagne University, in 2000. Currently, he is a Full Professor at the National School of Arts and Crafts of Rabat (ENSAM), Mohammed V University; a member in Research Center STIS, M2CS, Mohammed V University; and a member associate in Researcher, Industrial Engineering, data processing and logistic Laboratory, Hassan II University. He was a Professor in the Faculty of Sciences Ain Chock, Hassan II University. Benayad NSIRI has advised and co-advised more than 12 Ph.D. theses, contributed to more than 80 articles in regional and international conferences and journals. His research interests include but not restricted to computer science, telecommunication, signal and image processing, adaptive techniques, blind deconvolution, MCMC methods, seismic data and higher order statistics.