International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 20 (2017) pp. 10376-10384

Automatic Classification of ECG Arrhythmia Using Morphological Parameters with HMM and SVM

Dinesh D. Patil
Research Scholar, Department of Computer Science and Engineering,
Sri Satya Sai University of Technology and Medical Sciences, Sehore-466001, Madhya Pradesh, India.
Orcid Id: 0000-0001-9588-8513

Dr. R. P. Singh
Professor, Department of Computer Science and Engineering,
Sri Satya Sai University of Technology and Medical Sciences, Sehore-466001 Madhya Pradesh, India.
Orcid Id:0000-0001-6520-1290

Dr. Vilas M. Thakare
Professor and Head, Department of Computer Science, Faculty in Engineering and Technology,
Sant Gadge Baba Amravati University, Amravati-444601, Maharashtra, India.
Orcid Id: 0000-0001-8722-8881

Dr. Avinash K. Gulve
Associate Professor, Department of Master of Computer Applications,
Government College of Engineering, Aurangabad-431005, Maharashtra, India.
Orcid Id:/0000-0003-0578-2403

Abstract

A method for the automatic classification of cardiopathies from an electrocardiogram (ECG) is presented in the paper. This treatment is based on an analysis of certain morphological parameters for the recognition of 4cardiopathies. The Hidden Markov Model (HMM) was used for parameter analysis and recognition of cardiac arrhythmias. The morphological parameters were divided into homogeneous groups (amplitude, surface, interval and slope). These parameters are calculated for beats with 4 types of abnormalities (RBBB, APC, PVC and LBBB) from ECG records retrieved from the MIT-BIH arrhythmia database. Further SVM is applied as classifier for automatic detection of heart disease. Analysis of the different groups shows the overall recognition performance was 98.43%. The worst is 96.75% for the RBBB class.

Keywords: ECG, HMM, MIT-BIH, SVM.

INTRODUCTION

The Wavelet Transform in recent years has become a technique that has been the object of study by researchers in the analysis of signals from a wide variety of areas of Science, Engineering and Medicine.

The Wavelet Transform analysis is being useful to an extensive range of biomedical signals comprising electromyography (EMG) signals, electroencephalographic (EEG) signals, clinical sounds, respiratory patterns, blood pressure trends, and deoxyribonucleic acid (DNA) sequences, along with the signals object of this project the electrocardiographic (ECG).

At present there are a large number of applications for the processing of signals of physiological origin, mainly due to the complexity in the extraction of rules and specific characteristics for the implementation of algorithms that unequivocally reflect the medical knowledge derived from the interpretation of the biological / biomedical signals treated.

A bibliographical review of the methods and techniques used in the pre-processing, feature extraction, segmentation, as well as the sources related to the reduction of characteristics related to the stages of the ECG signal analysis process is discussed.

In summary, it is a question of elaborating the framework in which the present project fits, reviewing the investigations carried out in the study and analysis of the biomedical signals, the electrocardiographic signals in particular, situating, in this way, the context in the will locate the present study.

To this end, and as a summary of the bibliographical references studied regarding the work on ECG signals, it is observed that a scheme commonly accepted in this process is shown in Figure 1, structured in several stages.

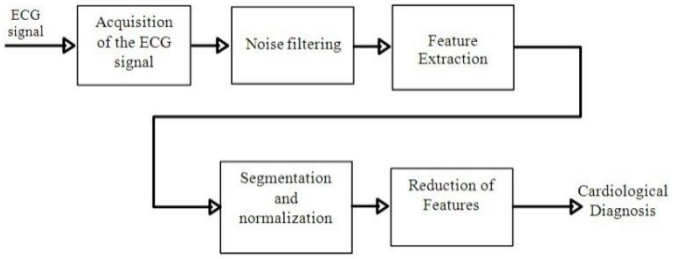

Figure 1: General process developed in the analysis of the beats of a Holter ECG signal

In the following diagram the set of techniques and algorithms more relevant within the specialized literature, as well as its bibliographic reference, in the process of study and analysis of the electrocardiographic-ECG signals are exposed

ECG Signal { ECG Signal Preprocessing; Feature Extraction; Segmentation

A. ECG Signal Preprocessing { Noise in the ECG Signal; Network interference; Baseline Variations

i.Noise in the ECG Signal

Classic Filtering Technique	{ [1],[2]
Mobile Media Filters	{ [2],[3]
Average heart rate	{[4],[5]
Approach through functions	{[6],[7]
Signal to Noise Ratio − SNR	{[7]
Wavelet Transform	{[8],[9],[10]
Artificial Neural Networks − ANN	{[11],[12],[13]

ii. Network Interference { [2], [14]

iii. Baseline Variations { Classic Filters { [1]; Adaptive Filters { [15],[16]

B. Feature extraction

Approximation Problem	{[7]
Calculation of error	{[7]
Definition of threshold	{[7]
Polygonal approximation process	{[7]

C. Segmentation { Detection of QRS complex; Detection of P and T waves; Definition of R − wave peaks

PROPOSED METHOD

The method we propose [17] is based on the study of the spectral parameters of QRS complexes and a multiresolution analysis which allows us to consider the frequently content of normal beats and pathological beats.

Estimation of the QRS complexes

To analyze the frequency content of QRS complexes, we selected four types of beats belonging to the most dominant classes of the MITDB database: normal beats (N), ventricular extra systoles (V), right and left branch blocks (RBBB and LBBB). Each class is represented by 200 beats selected in the MITDB database. The QRS complexes are extracted by means of a window of 180 ms length around the annotated positions of the R waves (50 ms before the R peak and 130 ms after the R peak). Each window then contains 64 samples for the sampling frequency 360 Hz.

Each segment containing the QRS complex is weighted by the Blackman window given by equation (1) in order to force the beginning and the end of the segment to zero and thus eliminate the discontinuities due to possible adjacencies of the and waves. The blackman window is also chosen for its much attenuated secondary lobes in the frequency domain (-74 dB).

$$W(k) = 0.42 - 0.5\cos\left(\frac{2\pi}{K-1}\right) + 0.08\cos\frac{2\pi}{K-1}$$

$$0 \le k \le K-1$$

Where is the number of samples of the segment considered The DFT known as a measurement tool for the amplitude and phase spectrum of the deterministic signals cannot be directly applied to these segments of ECG considered as random signals. The frequency representation of the latter uses a statistical description called: autocorrelation function. In this case, spectral analysis becomes an estimation problem. Estimation of the autocorrelation function: The autocorrelation function of a stationary random signal is defined by:

$$R_{xx}(n) \triangleq E[x(k)x(k+n)] = E[x(k-n)x(k)]\, R_{xx}(-n) \quad (2)$$

The quantity R_{xx} can be approximated using estimators. Note that each estimator of a given magnitude, a, is characterized by:

- Its bias: $B_{\hat{a}} = E[\hat{a}] - a$
- Its variance: $V_{\hat{a}} = E[(\hat{a} - E[\hat{a}])^2]$

1. Unbiased Estimator: if we have K values of $x(k)$, then for each n, the average value of $x(k)x(k+n) = u_n(k)$ can only be estimated from of $K-n$ values $u_n(k)$, one obtains then:

$$\hat{R}_{xx}(n) = \frac{1}{K-|n|}\sum_{k=0}^{K-|n|-1} x(k)x(k+n) \quad (3)$$

The bias $B_{\hat{R}_{xx}}$ of this estimator can be calculated easily. We have:

$$E\left[\hat{R}_{xx}(n)\right] = \frac{1}{K-|n|}\sum_{k=0}^{K-|n|-1} E[x(k)x(k+n)] \quad (4)$$

And

$$\frac{1}{K-|n|}\sum_{k=0}^{K-|n|-1} R_{xx}(n) = R_{xx}(n) \quad (5)$$

$B_{\hat{R}_{xx}} = E\left[\hat{R}_{xx}(n)\right] - R_{xx}(n) = 0$, the estimator is therefore unbiased. This estimator is consistent but has an important variance when the mean (Equation 4) is calculated on few terms (n close to$K$).

2. Biased Estimator: Another estimator very often used is given by:

$$\tilde{R}_{xx}(n) = \frac{1}{K}\sum_{k=0}^{K-|n|-1} x(k)x(k+n) \quad (6)$$

This estimator differs from the unbiased estimator by a multiplicative factor.

$$\tilde{R}_{xx}(n) = \frac{K-|n|}{K}\hat{R}_{xx}(n) \quad (7)$$

The biased estimator has a systematic error but its variance is lower than that of the unbiased estimator when n is close to K.

Wiener-Kinchine Theorem: This theorem states that the power spectral density of a random signal $x(t)$ is estimated by that of its autocorrelation function:

$$\hat{P}_x(f) = \int_{-\infty}^{\infty} C_{xx}(\tau)e^{-j2\pi f\tau}d\tau \quad (8)$$

In the case of a discrete signal $x(k)$ of length K, the simple estimator of its DSP, also called a simple periodogram, is given by:

$$\hat{P}_x(f) = \sum_{-K+1}^{K-1} \hat{R}_{xx}(n)e^{-j2\pi fn} \quad (9)$$

QRS decomposition into frequency bands

Multiresolution analysis decomposes the signal into different frequency bands. Thus the high frequencies and the low frequencies can be analyzed separately. The decomposition levels are set according to the sampling frequency and the frequency band sought. During this decomposition, the choice of the analyzing wavelet is very important, however, there is no general rule allowing the selection of a mother wavelet. This choice depends on the nature of the signal to be treated as well as on the application envisaged. In the case of the present application, the Haar wavelet gives better results because it is compact and offers a better temporal location.

Detection Algorithm

1. Decompose the ECG signal $f(t)$ up to 5 resolution levels.

2. Calculate $h = \left|\prod_{j=4}^{5} d_n^j\right|$

3. Location of QRS:

$$\left\{\begin{matrix} if\ h(n) \geq 0.3\max(h)\ then \\ n = QRS\ candidate\ otherwise\ n\ \neq QRS \end{matrix}\right\}$$

4. Suppose n and n' two selected consecutive positions, then:

$$\left\{\begin{matrix} if\ f|n-n'| < 36\ then\ n\ and\ n' = even\ QRS \\ otherwise\ n\ and\ n' \neq even\ QRS \end{matrix}\right\}$$

(36 f_e = 100 ms is the standard duration of a QRS complex)

5. A multiple detection in a 200 ms interval shall be deleted. This constraint has a physiological meaning: the refractory period is clearly greater than 200 ms.

6. Omitted beat search: If no beat is detected in a period equal to 1.5 times the current RR interval, the detection threshold is divided by 2 to check for any omitted QRS.

Segmentation

After detection of the R peaks, each beat is segmented and then stored in a vector on which a calculation of characteristics will be performed. These vectors must imperatively have the same dimension so that the characteristics of the corresponding beats are calculated under the same conditions.

The temporal representation of each segment containing the beat is obtained according to the procedure developed in [19].

Classification using Support Vector Machine (SVM)

Our objective is to evaluate the performance of a binary classifier with rejection; we first retained the annotations of the American Heart Association (AHA) [20] and the practical recommendations of the AAMI standard to form two classes of beats (normal beats and ectopic beats)

- The positive class (P) represents ectopic beats, premature beats and some unknown beats.
- Negative class (N) represents the normal beats (about 70% of the study base) and some abnormal beats LBBB (Left Bundle Branch Block), Right Bundle Branch Block (RBBB).

According to the AAMI (American Association for Medical Instrumentation) guidelines, records (102, 104, 107, 217) containing beats from pacemakers are excluded from this study. Recordings not containing PVC beats (11 records) are also excluded [21]. We then have 33 recordings of interest. Note that no selection of signals is based on their quality.

Learning Base

To constitute a good learning base that would allow us to generalize our classifier and thus obtain a global classifier, we took in a random way 10 records from which we used the beats present in the first 5 minutes, i.e. 1500 beats tagged by cardiologists. The cardiac beats being segmented and

quantified by discriminating parameters are each represented by a characteristic vector.

Test Basis

From each recording, we use the last 25 minutes for the test phase. Thus, the learning base is completely dissociated from the test database. This allows us to evaluate the generalization capacity of our classification algorithm.

Experimental Protocol

The conditions under which the classification tests are carried out must meet certain criteria. Thus all tests must be carried out in the same way.

Standardization of data

The learning and test data are structured as matrices. The number of rows represents the number of beats (or examplesx_i) and the number of columns represents the number of characteristics y_j quantifying each beat. Each line thus represents a vector characterizing a beat.

These characteristics can be very different from one example to another. Their standardization is therefore necessary. This operation consists in bringing each characteristic (each column of the data matrix) to follow a reduced Gaussian distribution, that is to say of zero mean and of standard deviation equal to unity. The normalization formula is given by:

$$\check{y}_{ij} = tansig\left(\frac{y_{ij}-\bar{y}_j}{\sigma_{yj}}\right) \quad (10)$$

Where y_{ij} is the characteristic component of the j^{th} characteristic. $\bar{y}_j$ and σ_{yj} are the mean and standard deviation of the j^{th} component of the feature vector and

$$tansig(n) = \frac{2}{[1+\exp(-2n)]-1}.$$

Choice of kernel

In kernel methods, such as SVMs, observations xi are implicitly projected into a larger dimensional feature space in which the problem becomes linearly separable. Thus, each scalar product (x_i, x_j) is replaced by its correspondent $k(x_i, x_j)$ in the new space, k is the kernel function.

This property shows both the importance and the difficulty of the choice of the kernel in the SVMs because there are no methods allowing an a priori choice. Generally, several nuclei are tested to select one that would minimize the cost of classification. For the present application, we have retained the Gaussian kernel.

Selection of hyper-parameters by cross-validation

The performance of an SVM technique is to a great extent subject to the determination of penalization parameters and kernel parameters. The optimal parameters are called hyper-parameters. In our case, the parameters to be optimized are: the penalty parameter C and the width of the Kernel expressed by σ. They are determined by a cross validation technique. Cross-validation is available in several sub-methods. The most common is the "k-Fold" method with typically $k \in [4,10]$. It consists in sharing a validation set composed of n examples in k equal parts. We then proceed to k learning experiments from examples. The remaining examples are used to test the rules learned and to evaluate the result. A cross-validation is applied according to the following algorithm:

1. Divide the validation set (EV) into k under balanced sets (same number of positive examples and even number of negative examples).
2. Solve the optimization problem (Learning) on the union of k-1 subsets.
3. Test on the k^{th}subassembly and calculate the mean squared error.
4. Repeat k times steps 3 and 4.
5. The division of EV generating the minimum error must be tested on a grid of values C and σ.

The pair (C, σ) corresponding to the minimum error is retained as a pair of hyper-parameters.

Feature Reduction

The step of reducing or selecting the features of an ECG signal turns out to be the most important step in the overall clustering process. The aim is to extract from the data the least number of features that provide the most information and allow them to discriminate correctly. If features with poor discrimination capacity are selected, the classification results will be confusing. On the other hand, if one characterizes the data while retaining its differential features, the subsequent classification process is greatly simplified and the results are improved. In the specialized literature there are a large number of transformations to be applied to the electrocardiographic data [22]. Namely, the Karhunen-Love transform (KLT), the Wavelet Transform [10], the Fourier transform, multivariate statistical analysis: Principal Component Analysis (PCA), multivariate analysis of variance. Another commonly used method for data characterization is the Hidden Markov (HMM) models.

Hidden Markov Models (HMM)

The use of Hidden Markov Hidden Markov Models (HMM) for the grouping of beats in an ECG signal raises the dilemma of their identification in some of the stages that make up the global clustering process [7]. On the one hand, the modelling of objects can be considered as a reduction of their characteristics to a probabilistic index generated by the model in question: a beat is identified with a model if the probability of the sequence of segments on the model exceeds a certain Threshold and the vector of characteristics of the object is reduced to a vector of probabilities of length equal to the number of models identified.

On the other hand, to complete the clustering with HMM, it is necessary to generate a model for each of the different morphologies presented in the ECG, so that the calculation of the similarity matrix depends on the successful selection of centroids and a correct training of the initial models.

That is, the set of HMMs provide the similarity matrix on which to apply the clustering algorithm, but this set of models is not known until properly identifying the centroids, which is necessary to apply some clustering algorithm.

It is proposed in [7] to consider the models as a way to reduce the characteristics, and using a stochastic shape recognizer (HMM) as the basis for the recognition and classification of ECG Holter beats.

An HMM can be defined as a finite-state and stochastic state automaton characterized by the following parameters:

1. Number of states of the model(N). The states remain hidden but for many practical applications they can be related to some magnitude or physical characteristic. The sequence of states that make up the model as referred to by $S = (S_1, S_2, \dots, S_N)$, and the state at time t is denoted by q_t.

2. Number of different symbols per state M, or also as number of Gaussian sources that participate in function of joint probability density.

3. Matrix of transition probability between states$\mathrm{A} = \{\mathrm{a_{i,j}}\}$, size $\mathrm{N} \times \mathrm{N}$, which defines the probability that the robot is in state i at time t, to be in the state j in $\mathrm{t}+1$.

$$\mathrm{a_{i,j}} = \mathrm{P}\left(\mathrm{q_{t+1}} = \mathrm{S_j} | \mathrm{q_t} = \mathrm{S_i}\right) \quad 1 \le \mathrm{i,j} \le \mathrm{N}\right)$$

4. Initial probabilities of the states $\pi = \{\pi_i\}$, where the probabilities of any of the states at the initial moment are determined, where:

$$\pi_i = \mathrm{P}(\mathrm{q_t} = \mathrm{S_i}) \quad 1 \le \mathrm{i,j} \le \mathrm{N}$$

(12)

5. Emission probabilities that can be divided into two categories depending on whether the observed sequence is discrete or continuous:

   a. Discrete emission probabilities $B = \{b_j(k)\}$. If M are the different symbols observed in each state (the size of the alphabet used), we have an alphabet$V = \{v_1, v_2, \dots, v_M\}$. Thus $b_j$would be the probability of observing the symbol if it is in state j:

$$\mathrm{b_j(k)} = \mathrm{P}(\mathrm{v_k \text{ in } t} \,|\mathrm{q_t} = \mathrm{S_i})$$

$$1 \le \mathrm{j} \le \mathrm{N}, 1 \le \mathrm{k} \le \mathrm{M}$$

(13)

   The sequence of observed symbols constitutes the output that is obtained from the system to be modeled. These types of systems are called discrete HMMs. See figure below.

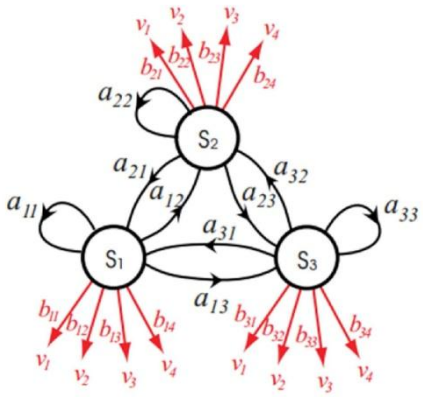

Figure 2: Discrete HMM of Three states with four symbols per state

   b. Probability of continuous emission $\mathrm{B} = \{\mathrm{b_j(O_t)}\}$ where $\mathrm{O} = \mathrm{O_1}, \mathrm{O_2}, \dots, \mathrm{O_T}$. The emission probability density function for each state is defined as a finite mixture of Gaussian sources:

$$b_j(O_t) = \sum_{m=1}^{M} d_{jm} N\left(O_t, \mu_{jm}, C_{jm}\right) \quad 1 \le j \le N$$

(14)

Being O_t the vector of characteristics of the sequence of observations to model, d_{jm} are the coefficients of the m^{th} mixture of state j and N is a Gaussian probability with mean vector μ_{jm} and covariance matrix C_{jm} corresponding to the m^{th} component of the mixture in state j. These models are known as continuous HMM (CHMM).

For the complete specification of an HMM it is necessary to determine the two parameters that determine the geometry of the model $N$and$M$, and the three probabilities A,b,π. All this is done by means of the notation:

$$\lambda = (A, B, \pi) \quad (15)$$

The constraints derived from a stochastic model are another important aspect of this HMM theory. They are expressed as:

$$\begin{cases} \sum_{i=1}^{N} \pi_i = 1 \\ \sum_{j=1}^{N} a_{ij} = 1 \qquad 1 \le i \le N \\ \sum_{j=1}^{M} b_j(k) = 1 \qquad 1 \le i \le N \end{cases} \quad (16)$$

Once the HMM is defined, there are some immediate problems such as learning.

Learning Algorithms in HMM

The unsupervised training methods available in the literature for application to HMM can be classified as shown in following figure

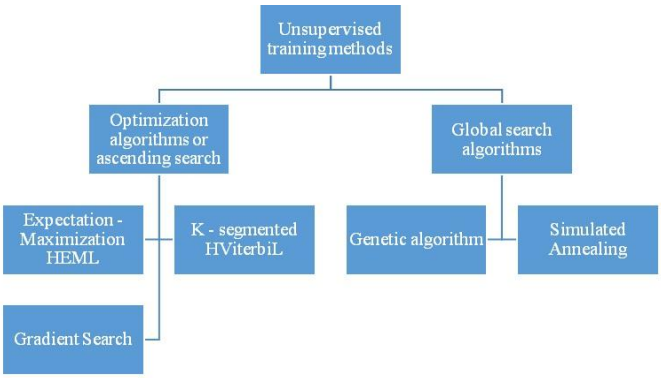

Figure 3 Learning algorithms in HMM

Search algorithms ascending highly dependent on the way in which the model, so that, in practice, if the initial parameters have not been optimal, the search may lead to suboptimal model is initialized. The training algorithms that are considered most relevant are described below.

1. ***Baum-Welch Algorithm:*** The EM algorithm (Expectation-Maximization) is a general method used to estimate the model parameters such that the probability (maximum-likelihood, ML) of a distribution is maximized generated from an incomplete set of data (Data not known for some reason).
2. ***Search for the Gradient:*** Unlike EM, the gradient search algorithm works on-line, directly on the samples obtained converting much faster to the maximum. This is because, while the convergence of the EM did not take the path of highest gradient on the surface of the probability function, requiring a large number of iterations until the maximum, for the search algorithm of the gradient defines a Learning ratio that allows you to converge towards the maximum much faster.
3. ***K-mediated segmented (Viterbi):*** By this method we try to adjust the parameters of the model $\lambda = (A, B\pi)$ to maximize $P(O, I|\lambda)$ where I is the optimum state sequence calculated by the Viterbi algorithm.
4. ***Genetic Algorithm:*** Genetic algorithms (GA) are a stochastic search method capable of optimizing the search space in a global, non-local way.
5. ***Simulated Annealing:*** The simulated annealing algorithm is based on a heuristic approach to probabilistic optimization. The basic idea is that of the heuristic exchange in which, during each iteration, it moves from a possible solution to a more probable one and located in the vicinity of the original. The algorithm stops when it reaches a local optimum.

As indicated above, one of the main applications of hidden Markov (HMM) models is in the field of biomedical signal processing. They have been used for the detection and analysis of sleep stages in EEG signals, to analyze small cardiac intervals (in ECGs) and other records of respiratory activity. They have also been used for the processing of ECG signals both for the detection of the QRS complex and for the identification of the P wave.

SIMULATION RESULTS

Database Explanation

Since 1975, the laboratories of the Beth Israel Hospital in Boston and MIT have developed an MIT / BIH database, which was started to be distributed in 1980. This database contains 48 records extracted from a half hour of ambulatory records with two ECG pathways obtained from 47 subjects studied by the BIH arrhythmia laboratory between 1975 and 1979. Twenty-three registrations were randomly selected from a set of 4000 outpatient 24-hour records (60%) and non-hospitalized patients (40%) at the Beth Israel Hospital in Boston, the remaining 25 records were selected from the same registrations but of rarely observed arrhythmias that have clinical significance. The recordings were sampled at a frequency of f = 360 Hz with an 11-bit resolution over a range of 10 mV. Two or more cardiologists independently annotated each record, about 110,000 annotations were included with the database [23].

Table 1 and Table 2 represents the values for different class for HMM and SVM respectively

Table 1: Table of Positive and Negative for HMM

	Normal	RBBB	APC	PVC	LBBB
True Positive	391	548	178	79	294
False Negative	57	2	3	3	6
False Positive	0	56	12	3	4
True Negative	1109	947	1352	1468	1232

Table 2: Table of Positive and Negative for SVM

	Normal	RBBB	APC	PVC	LBBB
True Positive	412	534	167	81	285
False Negative	37	1	2	2	5
False Positive	0	48	15	4	3
True Negative	1096	927	1279	1386	1207

Table 3 and Table 4 represent the accuracy for each class using HMM and SVM respectively.

Table 3: Result for HMM

Class	**Sensitivity** $=\frac{TP}{TP+FN}$	**Specificity** $=\frac{TN}{TN+FP}$	**Accuracy** $=\frac{TN+TP}{TP+FN+TN+FP}$
Normal	$\frac{391}{391+57}=\mathbf{87.27\%}$	$\frac{1109}{1109+0}=\mathbf{100\%}$	$\frac{391+1109}{1557}=\mathbf{96.33\%}$
RBBB	$\frac{548}{548+2}=\mathbf{99.63\%}$	$\frac{947}{947+56}=\mathbf{94.41\%}$	$\frac{548+947}{1553}=\mathbf{96.26\%}$
APC	$\frac{178}{178+3}=\mathbf{98.34\%}$	$\frac{1352}{1352+12}=\mathbf{99.12\%}$	$\frac{178+1352}{1545}=\mathbf{99.02\%}$
PVC	$\frac{79}{79+3}=\mathbf{95.12\%}$	$\frac{1468}{1468+3}=\mathbf{99.79\%}$	$\frac{79+1468}{1553}=\mathbf{99.61\%}$
LBBB	$\frac{294}{294+6}=\mathbf{98\%}$	$\frac{1232}{1232+4}=\mathbf{99.67\%}$	$\frac{294+1232}{1536}=\mathbf{99.34\%}$

Table 4: Result for SVM

Class	**Sensitivity**=$\frac{TP}{TP+FN}$	**Specificity** $=\frac{TN}{TN+FP}$	**Accuracy** $=\frac{TN+TP}{TP+FN+TN+FP}$
Normal	$\frac{412}{412+37}=\mathbf{91.75\%}$	$\frac{1096}{1096+0}=\mathbf{100\%}$	$\frac{412+1096}{1545}=\mathbf{97.54\%}$
RBBB	$\frac{534}{534+1}=\mathbf{99.81\%}$	$\frac{927}{927+48}=\mathbf{95.07\%}$	$\frac{534+927}{1510}=\mathbf{96.75\%}$
APC	$\frac{167}{167+2}=\mathbf{98.81\%}$	$\frac{1279}{1279+15}=\mathbf{99.21\%}$	$\frac{167+1279}{1463}=\mathbf{98.83\%}$
PVC	$\frac{81}{81+2}=\mathbf{97.59\%}$	$\frac{1386}{1386+4}=\mathbf{99.71\%}$	$\frac{81+1386}{1473}=\mathbf{99.59\%}$
LBBB	$\frac{285}{285+5}=\mathbf{98.27\%}$	$\frac{1207}{1207+3}=\mathbf{99.75\%}$	$\frac{285+1207}{1500}=\mathbf{99.46\%}$

RESULTS

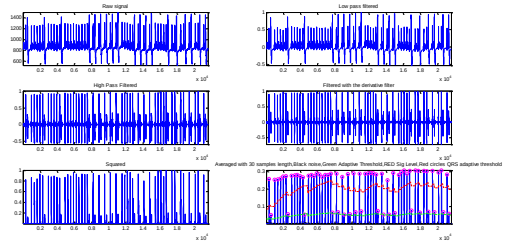

Figure 4: Filtering of original Signal

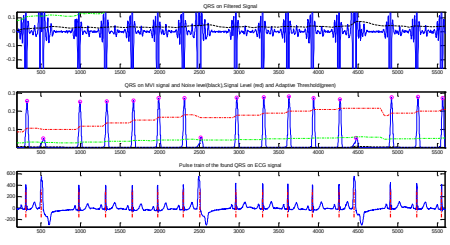

Figure 5: QRS on filtered signal

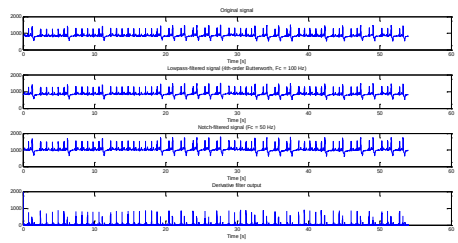

Figure 6: Filter output

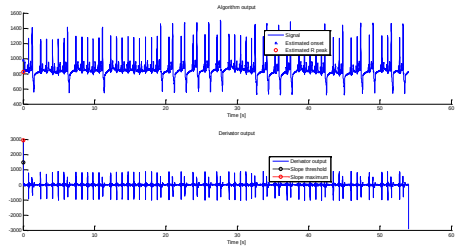

Figure 7: Algorithm output

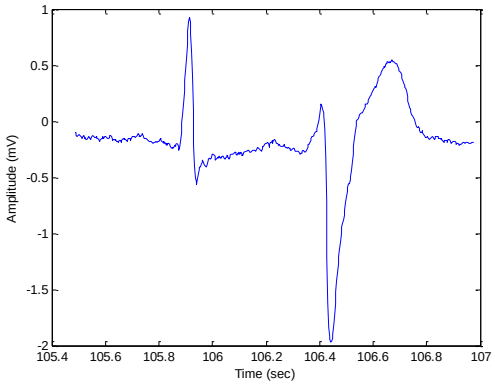

Figure 8: Extracted QRS signal

CONCLUSION AND PERSPECTIVES

The work presented in this paper consists mainly of two domains: signal processing and statistical learning. Signal processing techniques have allowed us to segment and represent each heartbeat by a vector of characteristics. Around these vectors, we constructed a decision rule that allows us to distinguish an ectopic beat from a normal beat. To do this, we chose to work with the SVMs method, a proven algorithm in terms of discrimination performance and generalization power.

The algorithms developed were validated on the American Heart Association (AHA) basis and the MIT-BIH database was severely noisy. A pre-processing step was then unavoidable. These pre-treatments have been performed in the time-scale domain because the signals suffer from non-stationary, particularly during the arrhythmia phases.

In the segmentation of cardiac beats, the presence of arrhythmias and a variety of atypical beats was constraining for the detection of R waves and the localization of QRS complexes. The detected QRS is further used as a feature extraction by using wavelet transform .These features are trained with SVM classifier and accuracy achieved is 99.46%.

We can also conclude that a large number of parameters do not necessarily provide the best result, but it is enough to use the most appropriate parameters, whatever they may be, to have a good performance. It may also be inferred that there is no better type of parameter for the classification of all heart diseases but each group of arrhythmias may have its appropriate parameters.

REFERENCES

[1] Ahlstrom, M.L. and Tompkins, W.J., 1985. Digital filters for real-time ECG signal processing using microprocessors. IEEE Transactions on Biomedical Engineering, (9), pp.708-713.

[2] Luo, S. and Johnston, P., 2010. A review of electrocardiogram filtering. Journal of electrocardiology, 43(6), pp.486-496.

[3] Okada, M., 1979. A digital filter for the ors complex detection. IEEE Transactions on Biomedical Engineering, (12), pp.700-703.

[4] Islam, M.S. and Alajlan, N., 2013. A morphology alignment method for resampled heartbeat signals. Biomedical Signal Processing and Control, 8(3), pp.315-324.

[5] Tamaki, S., Yamada, T., Okuyama, Y., Morita, T., Sanada, S., Tsukamoto, Y., Masuda, M., Okuda, K., Iwasaki, Y., Yasui, T. and Hori, M., 2009. Cardiac iodine-123 metaiodobenzylguanidine imaging predicts sudden cardiac death independently of left ventricular ejection fraction in patients with chronic heart failure and left ventricular systolic dysfunction: results from a comparative study with signal-averaged electrocardiogram, heart rate variability, and QT dispersion. Journal of the American College of Cardiology, 53(5), pp.426-435.

[6] Olmos, S., García, J., Jané, R. and Laguna, P., 1999. Truncated orthogonal expansions of recurrent signals: equivalence to a linear time-variant periodic filter. IEEE transactions on signal processing, 47(11), pp.3164-3172.

[7] Hosseini, H.G., Reynolds, K.J. and Powers, D., 2001. A multi-stage neural network classifier for ECG events. In Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE (Vol. 2, pp. 1672-1675). IEEE.

[8] Donoho, D.L. and Johnstone, I.M., 1995. Adapting to unknown smoothness via wavelet shrinkage. Journal of the american statistical association, 90(432), pp.1200-1224.

[9] Poornachandra, S., 2008. Wavelet-based denoising using subband dependent threshold for ECG signals. Digital signal processing, 18(1), pp.49-55.

[10] Thomas, M., Das, M.K. and Ari, S., 2015. Automatic ECG arrhythmia classification using dual tree complex wavelet based features. AEU-International Journal of Electronics and Communications, 69(4), pp.715-721.

[11] Acharya, U.R., Bhat, P.S., Iyengar, S.S., Rao, A. and Dua, S., 2003. Classification of heart rate data using artificial neural network and fuzzy equivalence relation. Pattern recognition, 36(1), pp.61-68.

[12] Rai, H.M., Trivedi, A. and Shukla, S., 2013. ECG signal processing for abnormalities detection using multi-resolution wavelet transform and Artificial Neural Network classifier. Measurement, 46(9), pp.3238-3246.

[13] Ronzhina, M., Janoušek, O., Kolářová, J., Nováková, M., Honzík, P. and Provazník, I., 2012. Sleep scoring using artificial neural networks. Sleep medicine reviews, 16(3), pp.251-263.

[14] Kocoń, S. and Piskorowski, J., 2012. Digital finite impulse response notch filter with non-zero initial conditions, based on an infinite impulse response prototype filter. Metrology and Measurement Systems, 19(4), pp.767-776.

[15] Jané, R., Laguna, P., Thakor, N.V. and Caminal, P., 1992, October. Adaptive baseline wander removal in the ECG: Comparative analysis with cubic spline technique. In Computers in Cardiology 1992, Proceedings of (pp. 143-146). IEEE.

[16] Meyer, C.R. and Keiser, H.N., 1977. Electrocardiogram baseline noise estimation and removal using cubic splines and state-space computation techniques. Computers and Biomedical Research, 10(5), pp.459-470.

[17] Zidelmal, Z., Amirou, A., Adnane, M. and Belouchrani, A., 2012. QRS detection based on wavelet coefficients. Computer methods and programs in biomedicine, 107(3), pp.490-496.

[18] Zidelmal, Z., Amirou, A. and Belouchrani, A., 2012. Heartbeat classification using support vector machines (SVMs) with an embedded reject option. International Journal of Pattern Recognition and Artificial Intelligence, 26(01), p.1250001.

[19] Manab K. Das. Samit Ari., 2014. Patient-specific ECG beat classification technique.Healthcare Technology Letters. 1(3), pp.98–103.

[20] R. Mark, G. Moody. MIT-BIH Arrhythmia database directory. Massachusette Inst. Technol. (MIT), 1988.

[21] R. Mark, R. Wallen. AAMI-recommended practice: Testing and reporting performance results of ventricular arrhythmia detection algorithms. AAMI, Tech. Rep. ECAR, 1987.

[22] Cuesta-Frau, David, Feature Extraction Methods Applied to the Clustering of Electrocardiographic Signal: a Comparative Study. En Pattern Recognition, 2002. Proceedings. 16th International Conference on. IEEE, 2002. pp. 961-964.

[23] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K. and Stanley, H.E., 2000. Physiobank, physiotoolkit, and physionet. Circulation, 101(23), pp.e215-e220.