ECG beat classification using a cost sensitive classifier

Z. Zidelmal [a], A. Amirou [b], D. Ould-Abdeslam [b,*], J. Merckle [b]

[a] Laboratoire LAMPA, Université Mouloud Mammeri, Tizi-Ouzou, Algeria
[b] Laboratoire MIPS, Université de Haute Alsace, France

A R T I C L E   I N F O

A B S T R A C T

In this paper, we introduce a new system for ECG beat classification using Support Vector Machines (SVMs) classifier with rejection. After ECG preprocessing, the QRS complexes are detected and segmented. A set of features including frequency information, RR intervals, QRS morphology and AC power of QRS detail coefficients is exploited to characterize each beat.

An SVM follows to classify the feature vectors. Our decision rule uses dynamic reject thresholds following the cost of misclassifying a sample and the cost of rejecting a sample. Significant performance enhancement is observed when the proposed approach is tested with the MIT-BIH arrhythmia database. The achieved results are represented by the average accuracy of 97.2% with no rejection and 98.8% for the minimal classification cost.

© 2013 Elsevier Ireland Ltd. All rights reserved.

1.   Introduction

The electrocardiogram (ECG) is a bio-electric signal that records the electrical activities of the heart. The ECG waveform is characterized by the waves: P, QRS and T. The most important wave is the QRS complex characterizing the ventricular contractions. Distances between them (RR intervals) define the rhythm, which is influenced by the emotions and physical activity.

The normal electrical conduction in the heart allows the impulse generated by the sino-atrial (SA) node to be rapidly propagated to the myocardium. This stimulated myocardium allows efficient contraction of the heart. In heart conduction disorders, ventricular excitation may not originate in the SA node but from other ectopic centers in the myocardium leading to a premature ventricular contraction (PVC) also called extra-systole or ectopic beat. Typically, PVCs are recognized by irregular RR intervals. The normally generated contractions produce normally shaped QRS complexes, while ectopic heartbeats generate a variety of QRS waveforms, quite differing from the normal ones.

Counting the occurrence of ectopic beats is of particular interest to support the detection of ventricular tachycardia and to evaluate the regularity of the depolarization of the ventricles. For example, the risk of ventricular fibrillation (VF) for patients with a structural heart disease is higher with an increased occurrence of PVCs. The automatic detection and classification of the ventricular contractions as normal or premature is a subject of long-term studies. This is the basis of the rhythm analysis, which is usually applied to continuous 24-h ECG recordings (Holter systems) to identify rhythm disorders.

Several algorithms with signal processing techniques have been proposed for detection and classification of heartbeats. Classical techniques extract heuristic ECG descriptors, such as QRS morphology [1–3] and interbeat R–R intervals [1,2,4,6].
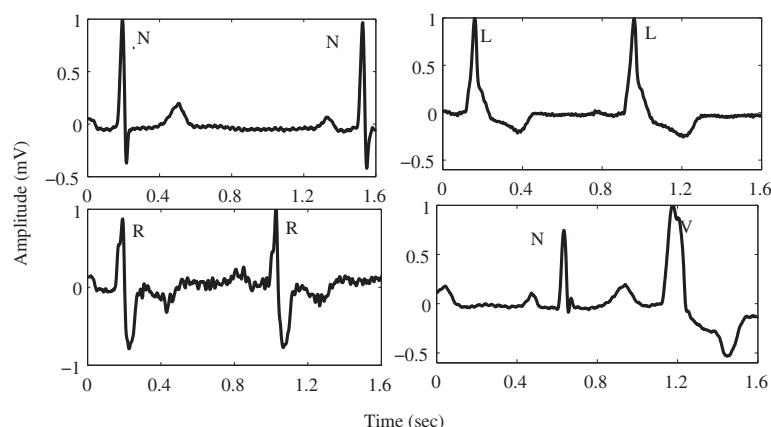

Fig. 1 – Different formes of ECG beats.

Other ECG descriptors rely on QRS frequency components [3–6] or matching pursuits for extraction of time–frequency beat descriptor [3]. Some methods apply QRS template matching procedures [7], wavelet and principal component analysis [1,6,8] and higher order statistics [9,10].

Several discriminative techniques have been developed for ECG beat classification such as artificial neural networks [1,2] or probabilistic neural networks [6]. Other work used the Kth nearest-neighbor rule [3,9] and genetic algorithms [4]. Among all these methods, Support Vector Machines (SVMs) known as excellent tool for classification and regression problems has shown success in this application field e.g. [4,8,10].

Even though the performance of all these techniques, misclassifications cannot be completely eliminated and, thus, can produce severe penalties. This motivates us to introduce a cost sensitive classifier to heartbeat recognition. This option has a particular interest in medical field where the consequences of misclassification can be substantial. In such situations, rejection may even be more preferable than risking misclassification.

This paper is organized as follows. Section 2 presents data signal and preprocessing techniques. Section 3 provides a detailed description of the feature extraction methodology. The classifier model proposed in this study is developed in Section 4. It recalls Bayes rule for binary classification with weighted errors and presents the SVMs in this framework and the learning criterion dedicated for the problem at hand. This proposed method is tested empirically in Section 5. Finally, Section 6 briefly concludes the paper.

2. Signal processing

2.1. ECG data

The annotated ECG records from the MIT-BIH arrhythmia data base [11] has been used in this study. This data base has 48 records. Each record is of length 30 min with 360 Hz sampling frequency. The data were recorded in two channels (modified limb lead II and modified limb lead VI) of surface ECG from long term Holter recorders. They present a variety of waveforms, artifacts, complex ventricular, junctional and supraventricular

arrhythmias and conduction abnormalities. Each record is accompanied by an annotation file in which each ECG beat has been identified by expert cardiologists. These labels referred to as 'truth' annotation and are used to develop the classifier and to evaluate its performance in the testing phase.

Since this study is to evaluate the performance of a binary classifier with a reject option, we followed the American Heart Association (AHA) records equivalent annotation [11] and the AAMI recommended practice [12] to form two classes of heartbeats: (i) the positive class representing the ventricular ectopic beats (V) including premature ventricular contractions, fusion of ventricular and normal beats and unknown beats; (ii) the negative class representing the normal beats (N), including all normal heartbeats (approximately 70% of the database) and some of the abnormal beats (left or right bundle branch block, aberrated atrial premature beat, nodal or atrial premature beat, nodal escape beat, atrial ectopic beat and nodal ectopic beat). We further restricted the N class to contain only heartbeats, which are representative for the predominant rhythm of the patient: normal beats, left bundle branch block beats (L) and right bundle branch block beats (R). In agreement with the AAMI recommended practice, records containing paced beats (102, 104, 107, and 217) were excluded. From this study, records with no PVC beats (11 records) were also excluded leaving 33 records of interest. The forms of predominant heartbeats are displayed in Fig. 1.

2.2. ECG filtering

The objective of this paper is to classify heartbeats. Before performing this task, several pre-processing steps were performed on the raw data. In fact the electrocardiogram (ECG) from body electrodes is corrupted by noise. Usually, two principal sources of ECG noise can be distinguished: the first one caused by the physical parameters of the recording equipment and the second one representing the bioelectrical activity also called background activity or baseline wander.

Several noise removal techniques were recently developed. ECG is a non-stationary biosignal corrupted by additive noise. An efficient tool for denoising such signal is the wavelet transform. The latter is a time-scale representation which describes a signal by using the correlation with translation
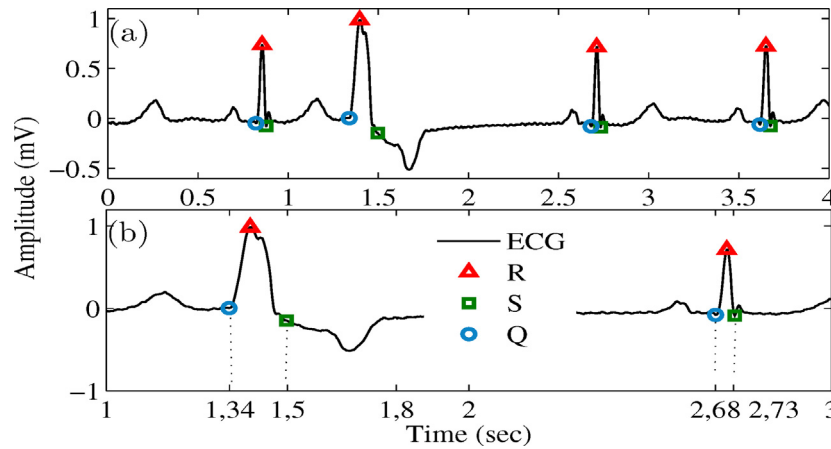
Fig. 2 – (a) Example of R peak detection and QRS segmentation. (b) Zoom on QS segments of PVC beat (left) and normal beat (right).

and dilatation of a function called mother wavelet. Basic theory can be found in many papers, i.e. [13]. The objective of wavelet based denoising process is to estimate the signal of interest $s(i)$ from the composite one $f(i)$ by discarding the noise $e(i)$ assuming that $f(i) = s(i) + e(i)$.

The noisy signal is decomposed into six levels by discrete wavelet transform (DWT) using the Daubechies wavelet $db4$. Passing the obtained coefficients through a threshold δ, certain values are set to zero. The denoised signal is recovered by taking the inverse discret wavelet transform (IDWT) of the resulting coefficients. For each decomposition level j, the threshold δ_j is computed according to the noise level. $\delta_j = \sigma_j \sqrt{2 \log(N)}$ where N is the number of samples and σ_j is the noise standard deviation which can be estimated from the median of its detail coefficients $\sigma_j = \mathrm{MAD}(d_i^j)/0.6745$ where MAD is the median absolute deviation of the corresponding level. More details can be found in [14].

The baseline drift is removed by zeroing the scaling coefficients of the DWT at level 6. This is equivalent to filtering the ECG signal with a H-Pass filter having a cutoff frequency of 2.8 Hz. The ECG signal sampled at $f_s = 360$ Hz is used as the initial scaling coefficients of DWT. They lie in the bandwidth $[0, f_s/2]$. The decomposition with the analysis filter bank is repeated over the scaling coefficients. Each new iteration of the filter bank divides the bandwidth of the scaling coefficients by two. After 6 iterations the resulting scaling coefficients correspond to the low frequency baseline drift of the ECG signal. Now, if these coefficients are set to zero, the reconstructed ECG will have all the details of the original ECG, but with low frequency baseline drift removed. Level 6 is also fixed referring to [15–17]. In these papers, it was shown that most of the QRS power lies in [5–22 Hz] therefore, this baseline removal do not affect the region of interest (QRS window).

2.3. ECG beat segmentation

In this study, the R peaks were detected using a robust method based on wavelet coefficients that was introduced in [15]. The peaks Q and S are detected using simple peak detection method. The peak Q is assumed to lie within 50 ms before R

peak, while S peak is assumed to lie within 100 ms after R peak. An example of QRS detection and QS segmentation is displayed in Fig. 2.

3. Feature extraction

The QRS complex in ECG signal varies with origination and conduction path of the activation pulse in the heart. When the activation pulse originates in the SA node and travels through the normal conduction path, the normal QRS complex has a sharp and narrow deflection and the spectrum contains high frequency components. When the activation pulse originates in the ventricle and does not travel through the normal path, the QRS becomes wide and the high frequency components of the spectrum are attenuated. A set of algorithms from signal conditioning to measurements of average wave amplitudes, durations, morphology, and areas is usually adopted to perform a quantitative description of a heartbeat. In this study, we select features that clearly characterize and discriminate the quality of the beats such as the R–R intervals, the QRS morphology and the frequency components.

3.1. RR intervals

RR intervals provide useful information for clinical diagnosis and identification of pro-arrhythmic events associated with heart-rate variation [2]. We studied the specific deviations of the RR intervals surrounding the heartbeat in order to give indications about the type of the heartbeat. For a beat i, the RR interval is $\mathrm{RR}_i = R_i - R_{i-1}$. This value is normalized to make it independent of the heart rate. The normalized RR interval is $\mathrm{RR}_{\mathrm{norm}}(i) = (\mathrm{RR}_i - \overline{\mathrm{RR}})/max(\mathrm{RR})$ where $\overline{\mathrm{RR}}$ and $max(\mathrm{RR})$ are respectively the mean and the maximum of the RR intervals in the same recording. We can see in Fig. 3 the normalized RR intervals obtained with N, L, R and V beats. We observe a unique peak around zero for all QRS complexes with supraventricular origin (N, R, L) considered in the same class according to the AAMI recommended practice [12] in opposite to V beats leading to a high RR variance.
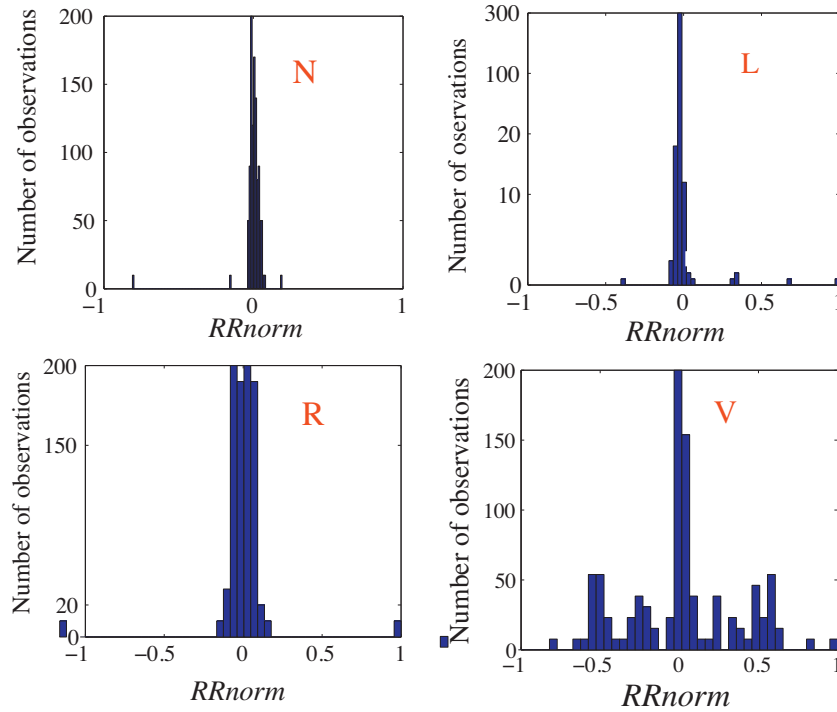
Fig. 3 – RR histogram for N, L, R and V beats.

As in [2], we extracted three features for each beat from the RR sequence: the pre-RR interval is the RR interval between a given beat and the previous one, the post-RR interval is the interval between a given beat and the following one and a local average interval obtained by averaging the ten RR intervals surrounding the considered beat.

3.2. Morphological descriptors

- QRS duration varies with origination and conduction path of the activation pulse in the heart so, it is a fundamental feature used for classifying beats. In this work, the QRS duration is represented by the time interval between the two peaks Q and S.
- The morphology of the beat is captured by a four Linear Predictive Coding (LPC) coefficients. The basic idea of this technique is that future values of a discrete signal are estimated as a linear function of previous samples. The most common representation is

$$\widehat{y}_n = \sum_{k=1}^{p} a_k y_{n-k} \qquad (1)$$

where a_k is the kth linear prediction coefficient, p is the order of the predictor, $\widehat{y}_n$ the present predicted sample and y_{n-k}, the kth precedent sample. Note that a_1 is usually equal to 1.

3.3. Frequency features

In Fig. 4(a), the estimated power spectrum density (PSD) of windowed (N, R, L and V) beats is displayed. Before computing fast Fourier transform (FFT), a 180 ms Blackman window is applied to suppress the discontinuities due to possible adjacent P and T waves. We can see on the same figure (b) that the spectra of V beats distinguish from the spectra of N and R beats at all energy levels because of shifting towards low frequencies. Some confusion can be made between L and V spectra but it is dismissed by RR interval feature which separates perfectly the two types of beats. Accordingly, we extract 6 features, the normalized power levels corresponding to 7.5 Hz, 10 Hz, 12.5 Hz, 15 Hz, 17.5 Hz and 20 Hz to discriminate the beats in the frequency domain.

3.4. QS power

To complete the frequency information, we consider detail coefficients at levels 4 and 5 obtained with discrete wavelet decomposition. At these levels, the power of each QS segment is computed and used as a classification feature. In [15], the authors have clearly showed that d_4 carries the dominant power of normal QRS complexes while d_5 carries the dominant power of ectopic QRS complexes (see Fig. 5). In each subband, signal variance represents the averaged power in that band as:

$$\sigma_x^2 = \frac{1}{N} \sum_{n=1}^{N} [x(n) - \overline{x}]^2 \qquad (2)$$

where $\overline{x}$ is the sample mean of the signal and N is the number of samples in the considered segment. The value of N differs from one beat to another depending on whether it is large (ectopic beat) or narrow (normal beat).
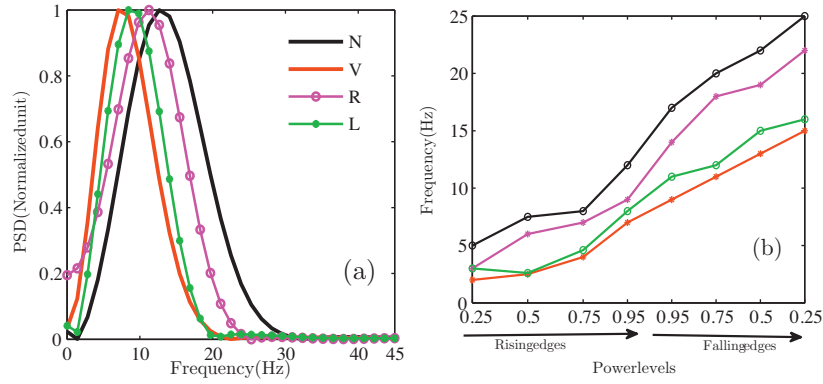
Fig. 4 – Power spectrum of different beats (a) and the relation between frequencies and power levels (b). For better representation, we limited the frequency axis to $f_s/4$.

3.5. *Data normalization*

Using all the observations above, each beat 'i' is represented by a 15 element vector x_i. The quantities of the features may be quite different, so their normalization is necessary to standardize all the features to the same level before training the SVM classifier. The formula of the normalization is defined as:

$$x_{ij'} = tansig\left(\frac{x_{ij} - \overline{x_j}}{\sigma_{x_j}}\right)$$

where x_{ij} is the jth component of the ith feature vector. $\overline{x_j}$ and σ_{x_j} are the mean and the standard deviation of the jth component of the feature vectors, and $tansig(.)$ the hyperbolic tangent sigmoid transfer function. The objective is to make each jth component to be normal distributed with zero-mean and unit standard deviation.
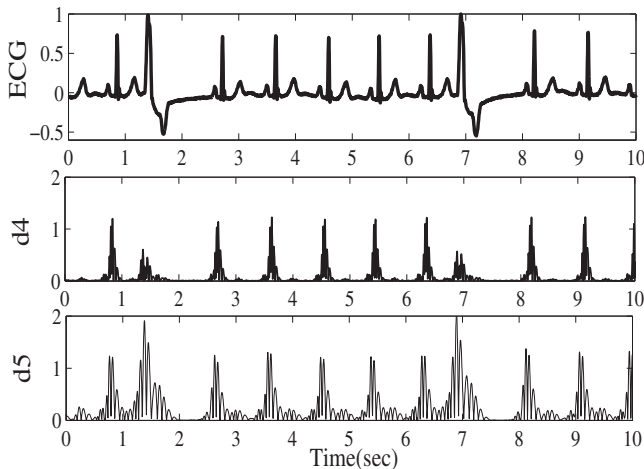
4. Classifier Model

4.1. *Bayes rule*

A discriminant function $f : \mathcal{X} \mapsto \mathbb{R}$ classifies an observation $x \in \mathcal{X}$ into one of two classes, labeled +1 or −1. Viewing $f(x)$ as a proxy value of the conditional probability $P(x) = \mathbb{P}(Y = 1|X = x)$, one is less confident for small values of $|f(x)|$ corresponding for $P(x)$ around 1/2. The strategy used in this work is to report $sgn(f(x)) = +1$ or −1 if $|f(x)|$ exceeds a threshold δ corresponding to $P(x) = P_o$ and no decision otherwise. Assuming that the cost of making a wrong decision is c and that of withholding a decision is r, the smallest risk $\mathcal{R}$ for a decision d is:

$$\mathcal{R}(d) = \min\{cP(x), c(1 - P(x)), r\}. \tag{3}$$

From Fig. 6, one can see that rejecting a pattern is a viable option if and only if $r < (c/2)$.

Fig. 5 – Decomposition of ECG signal using 'Haar' wavelet. Segment of MITDB119.
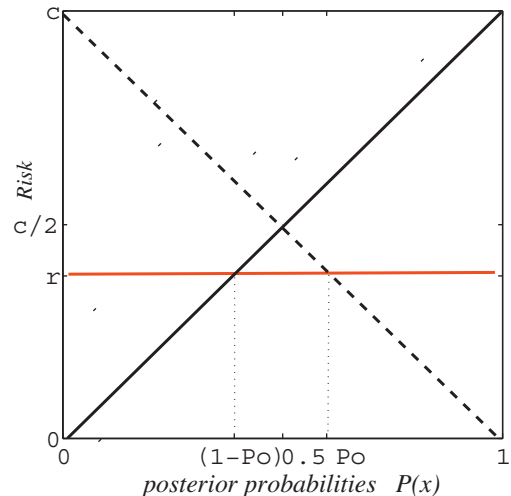

Fig. 6 – Expected losses against posterior probabilities $P(x)$.

Bayes rule can than be stated as:

$$d(x) = \begin{cases} +1 & \text{if } P(Y=1|X=x) > P_o, \\ -1 & \text{if } P(Y=1|X=x) < (1-P_o), \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

To minimize the empirical counterpart of the risk (3) computationally not feasible, one could replace it by surrogate loss functions. The most popular are the hinge loss motivated by [18] leading to sparse solutions [19,20] and the logistic regression model offering ability to estimate the posterior probability $P(Y=1|X=x)$ and then a good choice of the threshold δ. In this study, $P(Y=1|X=x)$ have to be accurate only in the vicinity of of P_o and $1-P_o$ (see Eq. (4)). As in [19], we built a hinge loss function tangent to neg-log-likelihood at $f_+ = \log[P_o/(1-P_o)]$ and at $f_- = \log[(1-P_o)/P_o]$. The decision rule can be expressed in terms of the function f as follows:

$$d^*(x,f) = \begin{cases} +1 & \text{if } f(x) > f_+, \\ -1 & \text{if } f(x) < f_-, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

where the output $f(x)$ of the SVM corresponds to the decision d^*, minimizer of the risk (3).

4.2.   Optimization

Introducing the double hinge loss results in an optimization problem that is similar to the standard SVMs problem. Let C be the penalization parameter to optimize in the training phase, $D = C(1-P_o)$. The optimization problem is:

$$\begin{cases} \min_{f,b,\xi,\eta} & \frac{1}{2}\|f\|^2 + C\sum_{i=1}^{n}\xi_i + D\sum_{i=1}^{n}\eta_i, \\ \text{S.c} & y_i(f(x_i)+b) \geq \tau - \xi_i \quad i=1,\ldots,n, \\ & y_i(f(x_i)+b) \geq -\eta_i \quad i=1,\ldots,n, \\ & \xi_i \geq 0 \quad \eta_i \geq 0 \quad i=1,\ldots,n. \end{cases} \quad (6)$$

where $\tau = (P_o\log(P_o))/(1-P_o) - \log(1-P_o)$. To compute the solution of (6), we use an active set algorithm following a strategy that proved to be efficient for standard SVMs. We refer to [19] for a detailed discussion.

5.   Results and discussion

The input of the classifier is a set of 15 element vectors x_i, representing the ECG beats. To fine tune a global classifier, we have selected 10 records containing the largest number of PVC beats. This is done to provide a relative balance between the positive class and the negative class in the training set. We used the first 5 min of data taken in the selected records. This practice conforms to the AAMI-recommended procedure allowing the usage of at most 5-min section from the beginning of each patient's recording for training [21]. From each segment of 5 min, were selected 150 beats leading to a training set of 1500 vectors. During the test phase, only the last 25 min
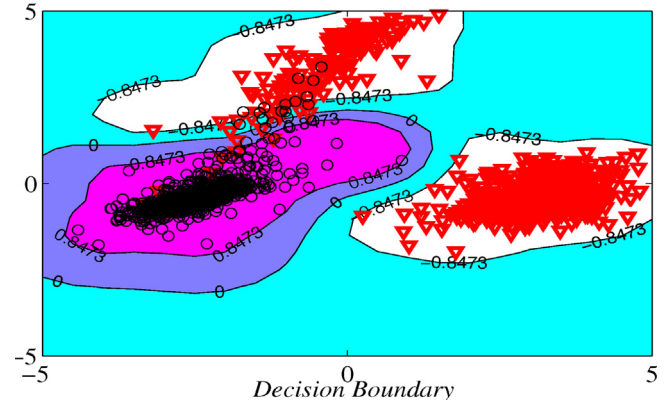

Fig. 7 – Scatter plot showing the reject region induced by the reject thresholds in correspondence to the costs of misclassifying and rejecting samples. Positive cases are represented by black circles and negative cases by red triangles. The pink region corresponds to classification Y = +1, and white regions to classification Y = −1. The blue and the light blue regions correspond to no decision. The lines +0.8473 and −0.8473 correspond respectively to f_+ and f_- and the line 0 corresponds to $f(x) = 0$ or $P(Y = 1 | X = x) = 0.5$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

in each type are used. Hence, the testing data are never part of any training data.

To explain the effectiveness of the classification methodology and the impact of the reject option, we simply apply standard statistical data visualization technique by projecting the high-dimensional data onto a 2D space. Fig. 7 shows an example of the reject region produced by the SVM classifier when our algorithm is applied to 221 and 214 containing together a negative classes of 3574 beats (1543 of normal beats and 2003 of L beats) and 652 of V beats considered as a positive classes. In this figure, we can see the reject thresholds given by Eq. (5); $f_+ = 0.8473$ and $f_- = -0.8473$.

A complete description of a such classifier is given by the error-reject tradeoff (the error rate E against the reject rate R). A typical tradeoff curve is displayed in Fig. 8(a). For small rejection rate ($R < 1.9\%$), the curve shows a high slope. This means that the majority of the rejected patterns would have been erroneously classified. We point out that the advantage of classifying with rejection can be judged by the error-reject tradeoff. Fig. 8(b) indicates the variation of the classification cost given by

$$C_c = \frac{[c(\text{FN} + \text{FP}) + r R_{rej}]}{N_{tot}} \quad (7)$$

where FP (False Positive) is the number of examples labeled −1 and categorized in the positive class, FN (False Negative) the number of examples labeled +1 and categorized in the negative class, R_{rej} the number of rejected examples and N_{tot}, the number total of examples. This figure shows that the optimal classification cost C_c corresponds to a good error-reject tradeoff.
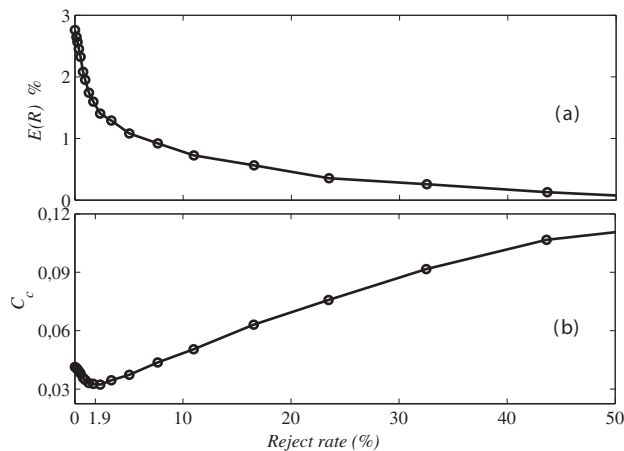
Fig. 8 – Error-reject tradeoff (a) and the classification cost against the reject rate (b).

Table 1 – Comparison of the performances of our classifier with others using artificial neural network (ANN) or hybrid Genetic Algorithm-SVM (SVMGA) in terms of sensitivity.		
Method	Model classifier	S_e(%)
[2]	ANN	94.4
[4]	SVMGA	98.91
[7]	ANN	98.4
Our algorithm with NR	–	98.2
Our algorithm for MC_c	–	99

A statistical parameter is also used to compare detection algorithms. The sensitivity of the classifier is computed by

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

where True Positive (TP) are the samples labeled +1 categorized in the positive class.

A lot of studies dealing with heartbeat classification using the same database are present in the literature. All previous works used a two decision rule (No Rejection). The results we obtained with No Rejection (NR) shown to be competitive to average results reported by other authors separating beats whose origin is from the SA node from premature ventricular contractions (see Table 1). On the same table, we reported our achieved results corresponding to the Minimal Classification Cost (MC_c). Note however that the main advantage of our approach is the ability to optimize the classification cost.

6.   Conclusion

In this paper, we proposed an automated heartbeat classifier, using Support Vector Machines with an embedded reject option. The proposed system accomplishes preprocessing, feature extraction and recognition tasks for ectopic heartbeats detection.

A set of features including frequency information, RR intervals, QRS morphology and AC power of QRS detail coefficients is exploited to characterize each beat. The resultant feature vector representing each ECG heartbeat is used as input of the modified SVM classifier.

For this purpose, a cost-sensitive reject rule for SVMs is set. The main feature of our system is that the decision rule uses dynamic reject thresholds following the cost of rejecting a sample and the cost of misclassifying a sample in order to optimize the classification cost. This approach is important in medical applications when rejection may even be more preferable than risking misclassifications.

Our results shown above illustrate a good error reject trade-off and indicate an error rate of 2.8% with no rejection and less than 1.2% for the minimal classification cost. These results are competitive to other published studies.

REFERENCES

[1] T. Ince, S. Kiranyaz, M. Gabboudj, A generic and robust system for automated patient-specific classification of ECG signals, IEEE Transactions on Biomedical Engineering 56 (5) (2009) 1415–1426.
[2] P. De-Chazal, R. Reilly, A patient adapting heartbeat classifier using ECG morphology and heartbeat interval features, IEEE Transactions on Biomedical Engineering 53 (12) (2006) 2535–2543.
[3] I. Christov, G. Gomez-Herrero, V. Krasteva, I. Jecova, A. Gotchev, Comparative study of morphological and time–frequency ECG descriptors for heartbeat classification, Medical Engineering and Physics (28) (2006) 876–887.
[4] A. Khazaee, A. Ebrahimzadeh, Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features, Biomedical Signal Processing and Control (5) (2010) 252–263.
[5] M.L. Talbi, A. Charef, PVC discrimination using the QRS power spectrum and self-organizing maps, Computer Methods and Programs in Biomedicine (94) (2009) 223–231.
[6] S.N. Yu, Y.H. Chen, Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network, Pattern Recognition Letters (28) (2007) 1142–1150.
[7] V. Krasteva, I. Jecova, QRS template matching for recognition of ventricular ectopic beats, Annals of Biomedical Engineering 5 (12) (2007) 2065–2076.
[8] C.P. Shen, W.C. Kao, Y.Y. Yang, M.C. Hsu, Y.T. Wu, F. Lai, Detection of cardiac arrhythmia in electrocardiograms using adaptive feature extraction and modified support vector machines, Expert System with applications (39) (2012) 7845–7852.
[9] Y. Kutlu, D. Kuntalp, Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients, Computer Methods and Programs in Biomedicine 105 (3) (2012) 257–267.
[10] F. Melgani, Y. Bazi, Classification of electrocardiogram signals with support vector machines and particle swarm optimization, IEEE. Transactions on Information Technology in Biomedicine 12 (5) (2008) 667–677.
[11] R. Mark, G. Moody, MIT-BIH arrhythmia database, http://www.physionet.org/physiobank/database/mitdb/ (accessed 12.07.11).
[12] R. Mark, R. Wallen, AAMI-recommended practice: Testing and reporting performance results of ventricular arrhythmia detection algorithms. AAMI, Tech. Rep. ECAR, 1987.
[13] S. Mallat, A Wavelet Tour of Signal Processing, 3rd ed., Academic Press, Burlington, 2009.
[14] M.I. Johnstone, B.W. Silverman, Wavelet threshold estimators for data with correlated noise. Technical Report, Dept. of Statistics, Stanford University, 1996.
[15] Z. Zidelmal, A. Amirou, M. Adnane, A. Belouchrani, QRS detection using wavelet coefficients, Computer Methods and Programs in Biomedicine 107 (3) (2012) 490–496.

[16] N.V. Thakor, J.G. Wbstor, W.J. Tompkins, Estimation of the QRS complex power spectra for design of a QRS filter, IEEE Transactions on Biomedical Engineering (11) (1984) 702–706.

[17] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Transactions on Biomedical Engineering 3 (32) (1985) 479–482.

[18] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer Series in Statistics, 1995.

[19] Y. Grandvalet, A. Rakotomamonjy, J. Keshet, S. Canu, Support vector machines with a reject option, Advances in Neural Information Processing Systems (21) (2009) 537–544.

[20] M. Wegkamp, M. Yuan, Classification methods with reject option based on convex risk minimization, Journal of Machine Learning Research (11) (2010) 111–130.

[21] American Hearth Association, Recommended Practice for Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms, Assoc. Adv. Med. Instrum., Arlington, 1987.