Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Engineering 38 (2012) 1333 – 1339

Procedia Engineering

www.elsevier.com/locate/procedia

Modeling and Optimization

ECG Arrhythmia Detection and Classification Using Relevance Vector Machine

S.Gayathri [a], M.Suchetha [b], V.Latha [c] a*

[a] PG Student, Velammal engineering college,Chennai-66,India.
[b]Assistant professor, Velammal Engineering College, Chennai-66, India
[c]Professor, Velammal Engineering College, Chennai-66, India.

Abstract

The Electrocardiogram (ECG) is one of the most effective diagnostic tools to detect cardiac diseases. It is a method to measure and record different electrical potentials of the heart. The electrical potential generated by electrical activity in cardiac tissue is measured on the surface of the human body. This ECG can be classified as normal and abnormal signals. In this paper, a thorough experimental study was conducted to show the superiority of the generalization capability of the Relevance Vector Machine (RVM) in the automatic classification of ECG beats. To achieve the maximum accuracy the RVM classifier design by searching for the best value of the parameters that tune its discriminant function, and upstream by looking for the best subset of features that feed the classifier. The experiments were conducted on the ECG data from the Massachusetts Institute of Technology–Beth Israel Hospital (MIT–BIH) arrhythmia database to classify five kinds of abnormal waveforms and normal beats. The obtained results clearly confirm the superiority of the RVM approach when compared to traditional classifiers.

Keywords: Electrocardiogram, classification, arrhythmia, Relevance Vector Machine.

I Introduction

Heart is a muscular organ that acts like a pump to continuously send blood throughout your body. If disease or injury weakens the heart and body's organs did not receive enough blood to work normally. Electrocardiography is a device that records the electrical activity of the heart over time. It is a Gold standard for diagnosis of cardiac arrhythmias. This paper describes five types of arrhythmias, they are Atrial flutter, Premature Atrial Contraction, Premature Ventricular Contraction, Right Bundle Branch Block, Left Bundle Branch Block. In Atrial Flutter, when the heart rate is sufficiently elevated so that the isoelectric interval between the end of T and beginning of P disappears, the arrhythmia is called atrial flutter. The frequency of these fluctuations is between 220 and 300/min. In Premature Atrial Contraction, Premature beats or extra beats frequently cause irregular heart rhythms. Those that start in the upper chambers are called premature atrial contractions (PACs). It has narrow QRS complex. In Premature Ventricular Contraction, A premature ventricular contraction is one that occurs abnormally early. If its origin is in the atrium, it has a supraventricular origin. The complex produced by this supraventricular arrhythmia lasts less than 0.1 s. If the origin is in the

* *Tel/fax: +91-8903233113*
Email id: sethuramangayathri@yahoo.in

1877-7058 © 2012 Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
doi:10.1016/j.proeng.2012.06.164

ventricular muscle, the QRS-complex has a very abnormal form and lasts longer than 0.1 s. In Right Bundle Branch Block, if the right bundle-branch is defective, the activation reaches the right ventricle by proceeding from the left ventricular. QRS complex is wider than 0.1sec. In Left Bundle Branch Block, when the activation of the left ventricle is delayed, it results in left ventricular contraction.

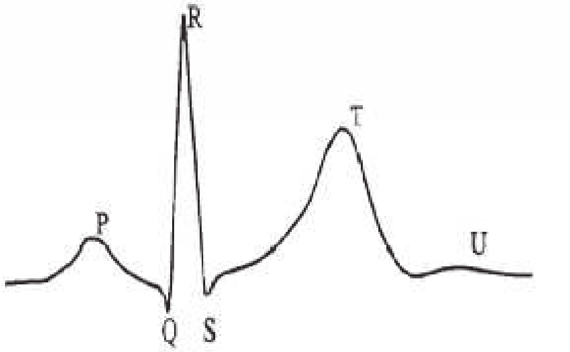


Fig 1. Structure Of ECG Signal

Various works has been carried out by scientists, among them the most recently published works are discussed as follows.

In greater detail, the method presented in [1] is based on a hybrid fuzzy neural network that consists of a fuzzy self-organizing sub-network connected in a cascade with a multilayer perceptron. The authors proposed to use high-order statistics (i.e., cumulants of the second, third, and fourth orders) as input features for feeding their classifier. In [2], a neuro-fuzzy approach for the ECG-based classification of heart rhythms is described. Here, the QRS complex signal is characterized by Hermite polynomials, whose coefficients feed the neuro-fuzzy classifier. In [3], different classification systems based on linear discriminant classifiers are explored, together with different morphological and timing features obtained from single and multiple ECG leads. In [4], L. Khadra, A. S. Al-Fahoum, and S. Binajjaj, proposed a high order spectral analysis technique is suggested for quantitative analysis and classification of cardiac arrhythmias. The algorithm is based upon bispectral analysis techniques. The bispectrum is estimated using an autoregressive model, and the frequency support of the bispectrum is extracted as a quantitative measure to classify atrial and ventricular tachyarrhythmia. Results show a significant difference in the parameter values for different arrhythmias.

In [5], a patient-adapting heartbeat classifier system based on linear discriminants is proposed. The classification system processes an incoming recording with a global-classifier to produce the first set of beat annotations. Then, an expert validates, and, if necessary, corrects a fraction of the beats of the recording. The system then adapts by first training a local classifier using the newly annotated beats, and combines both local and global classifiers to form an adapted classification system. In [6], a rule-based rough-set decision system is presented for the development of an inference engine for disease identification using time-domain features. In [7], the authors present an approach for classifying beats of a large dataset by training a neural network classifier using wavelet and timing features. The authors found that the fourth scale of a dyadic wavelet transform with a quadratic spline wavelet together with the pre/post RR-interval ratio is very effective in distinguishing normal and PVC from other beats. In [8], the author describes the decomposition implementations for two such "all-together" methods and then compared their performance with three methods based on binary classifications: "one-against-all," "one-against-one," and directed acyclic graph SVM (DAGSVM). In [9], the potential of the authors' method for clinical uses and real-time detection was examined using human surface ECGs and intra cardiac electrograms (EGMs). The method achieved high sensitivity and specificity (>0.98) in discrimination of supraventricular rhythms from ventricular ones. The authors also present a hardware implementation of the algorithm on a commercial single-chip CPU. In [10], presents a new solution to the expert system for reliable heartbeat recognition. The recognition system uses the support vector machine (SVM) working in the classification mode. Two different preprocessing methods for generation of features are applied. The combination of classifiers utilizes the least mean square method to optimize the weights of the weighted voting integrating scheme.

In this paper, the approach to ECG beat classification presented a thorough experimental exploration of the RVM capabilities for ECG classification. Further the performances of the RVM approach in terms of classification accuracy are evaluated: 1) by automatically detecting the best discriminating features from the whole considered feature space and 2) by solving the model selection issue. Unlike traditional feature selection methods, where the user has to specify the number of desired features, the proposed system gives a method for extraction of features called as "feature detection". Feature selection and feature detection have the common characteristic of searching for the best discriminative features. Several methods have been proposed for the automatic classification of ECG signals.

II Methodology

2.1 Pre-processing

The ECG signal is first filtered by two median filters. The first median filter is of 200 msec width and removes the QRS complexes and the P waves. The resulting signal is then processed with a second median filter of 600 msec width to remove the T waves. The signal resulting from the second filter operation contains the baseline wanderings and can be subtracted from the original signal. Power line and other high frequency artifacts are then removed from the baseline corrected signal with a FIR filter.

2.2 Feature Extraction

There are different types of arrhythmias which can be divided into two main groups; Morphological and non-morphological. Non-morphological arrhythmias could be diagnosed by having some features of the ECG signal such as amplitude and position of the PQRST peaks. These features can be extracted using signal processing methods either in time or transform domain. To detect morphological arrhythmias it is needed to use methods which enable us to extract the shape and morphological properties of the ECG signal. For a normal person, the heart rate will be 60-90bpm. If the heart rate goes beyond or below the limit, abnormalities occur in human heart. Below are the features of ECG waveform that gives the details about the amplitude and frequency of PQRST wave. Normal person will have PR interval: 0.12-0.20sec, QRS duration: 0.06-0.10sec, QT interval ($QT_{c\,<=}$ 0.40sec), where QTc is the Bazett's formula. QT<=0.38 at 80bpm and QT<=0.42 at 60bpm. Frontal plane QRS axis: +90° to -30° (in adult). In P-wave, it is important to remember that the p-wave represents the sequential activation of the right and left atria, and it is common to see notched or biphasic p-wave of right and left atrial activation. P duration < 0.12 sec, P amplitude < 2.5 mm, Frontal plane P wave axis: 0° to +75°, May see notched P waves in frontal plane. In QRS Complex, the QRS represents simultaneous activation of the right and left ventricles, although most of the QRS waveform is derived from the larger left ventricular musculature. QRS duration $\leq$ 0.10 sec, QRS amplitude is quite variable from lead to lead and from person to person. Two determinates of QRS voltages are: Size of the ventricular chambers (i.e., the larger the chamber, the larger the voltage), Proximity of chest electrodes to ventricular chamber (the closer, the larger the voltage). In ST segment and T wave, ST segment occurs in leads with normal s waves and the normal configuration is concave upward. The normal T wave is usually in the same direction as the QRS except in the right precordial leads. In the normal ECG the T wave is always upright in leads I, II, V3-6, and always inverted in lead a VR. In U Wave, (the most neglected of the ECG waveforms) U wave amplitude is usually < 1/3 T wave amplitude in same lead; U wave direction is the same as T wave direction in that lead, U waves are more prominent at slow heart rates and usually best seen in the right precordial leads, Origin of the U wave is thought to be related to after depolarization which interrupt or follow depolarization.

Temporal classification of ECG describes about the timing interval between the PQRST waves. In RR Interval, the interval between an R wave and the next R wave. Normal resting heart rate is between 60 and 100 bpm. Duration will be 0.6 to 1.2 sec. In PR Interval, the PR interval is measured from the beginning of the P wave to the beginning of the QRS complex. The PR interval reflects the time the electrical impulse takes to travel from the sinus node through the AV node and entering the ventricles. The PR interval is therefore a good estimate of AV node function. Duration will be 120 to 200 ms. In QT Interval, the QT interval is measured from the beginning of the QRS complex to the end of the of T wave. A prolonged QT interval is a risk factor for ventricular tachyarrhythmia and sudden death. It varies with heart rate and for clinical relevance requires a correction for this, giving the QT_c. In ST Interval, the ST interval is measured from the J point to the end of the T wave. Duration will be 320ms.

2.3 Classification of arrhythmia

In this paper, we proposed framework of driving the best classification technique that could build best accurate Classifier for classifying cardiac arrhythmias based on feature extraction that works effectively and efficient.

2.3.1 Relevance Vector Machine

The relevance vector machine (RVM) classifier is a probabilistic extension of the linear regression model, which provides sparse solutions. It is analogous to the SVM, since it computes the decision function using only few of the training examples, which are now called relevance vectors. However training is based on different objectives. The RVM model y(x; w) is output of a linear model with parameters w = (w1, . . . , wN)^T, with application of a sigmoid function for the case of classification:

$$y_{RVM}(x) = \sigma(\; {}_{n=1}^{N} \omega_n \, K(x, x_n)) \qquad (1)$$

where σ(x) = 1/(1+exp(−x)). In the RVM, sparseness is achieved by assuming a suitable prior distribution on the weights, specifically a zero-mean, Gaussian distribution with distinct inverse variance α_n for each weight ω_n:

$$p(\omega|\alpha) = \; {}_{n=1}^{N} N\,(\omega_n|0, \alpha_n^{-1}) \qquad (2)$$

The variance hyper parameters α = (α_1,..., α_N) are assumed to be Gamma distributed random variables:

$$p(\alpha) = \; {}_{n=1}^{N} Gamma\,(\alpha_n|a, b) \qquad (3)$$

The parameters a and b are implicitly fixed and usually they are set to zero (a = b = 0), which provides sparse solutions.

2.3.2 RVM Algorithm

Given a training set $\{x_n, t_n\}_{n=1}^{N}$with t_n {0,1} training in RVM is equivalent to compute the posterior distribution$p(\omega, \alpha|t)$. However, since this computation is intractable, a quadratic approximation log $p(\omega, \alpha|t)$ $(\omega - \mu)^T \; {}^{-1}(\omega - \mu)$is assumed and computed matrix Σ and vector μ as:

$$= (\varphi^T B \varphi + A)^{-1} \qquad (4)$$

$$\mu = \; B\,\varphi^T \hat{t} \qquad (5)$$

with the N × N matrix Φ described as [Φ]ij = K(xi, xj), A = diag(α1, . . . , αN), B = diag (β1, . . . , βN), βn = yRVM(xn)[1−yRVM(xn)] and ˆt = Φμ+B¡1(t−y). The parameters α are set to the values αMP that maximize the logarithm of the following marginal likelihood

$$L(\alpha) = \log p(\alpha|t) = -\frac{1}{2}[\text{Nlog}\, 2\pi + \log|c| + t^T c^{-1} t \qquad (6)$$

With C = B^{-1} + ΦA^{-1}Φ^T. This, gives the following update4

formula:

$$\alpha_n = \frac{1-\alpha_n \; {}_{nn}}{\mu_n^2} \qquad (7)$$

The RVM learning algorithm iteratively evaluates formulas. After training, the value of yRVM(x) = y(x; μ) can be used to estimate the reliability of the classification decision for input x. Values close to 0.5 are near the decision boundary and consequently are unreliable classifications, while values near 0 and near 1 should correspond to reliable classifications. In this experiment, the reliability measure is used

$$RE_{RVM} = |2_{yRVM}(x) - 1| \qquad (8)$$

which uses values near 0 for unreliable classifications and near 1 for reliable classifications.

III Simulation Result

The input to the classifier is the set of vectors xi, representing the ECG beats of different patients, representing different types of arrhythmia. The experiment conducted on the basis of ECG data from Physiobank, the MIT–BIH arrhythmia database. In particular, the considered beats refer to the following classes: normal sinus rhythm (N), atrial premature beat (A), ventricular premature beat (V), right bundle branch block (RB), Left bundle branch block (LB), Atrial flutter. Each beat of the normal human heart originates in the SA node. Because many parts of the heart possess an inherent rhythmicity any part under abnormal conditions can become the dominant cardiac pacemaker, any part under abnormal conditions can become the dominant cardiac pacemaker. The beats were selected from the recordings of 20 patients, which correspond to the following files: 100, 102, 104, 105, 106, 107, 118, 119, 200, 201, 202, 203, 205, 208, 209, 212, 213, 214, 215, and 217. In order to feed the classification process, in this paper, the two following kinds of features are adopted: 1) ECG morphology features and 2) three ECG temporal features, i.e., the QRS complex duration, the RR interval (the time span between two consecutive R points representing the distance between the QRS peaks of the present and previous beats), and the RR interval averaged over the ten last beats. Simulation results shown below.

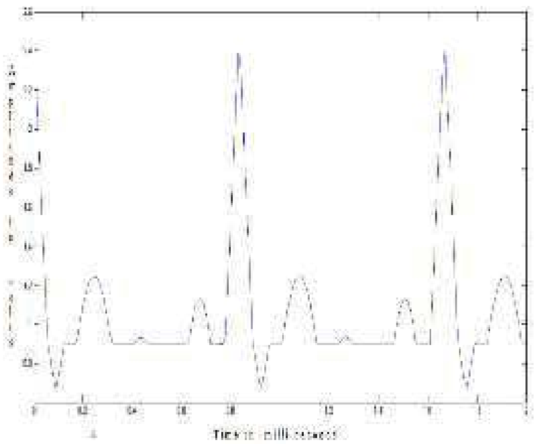
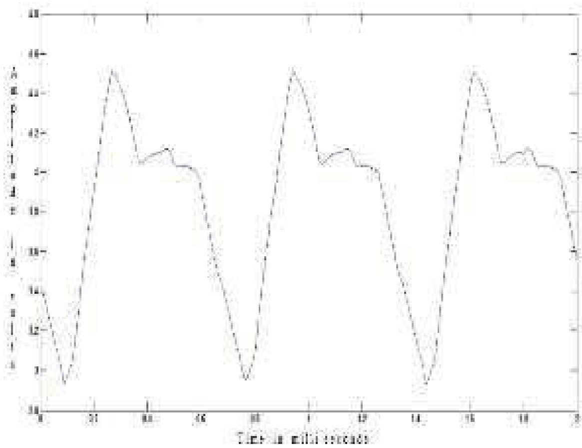
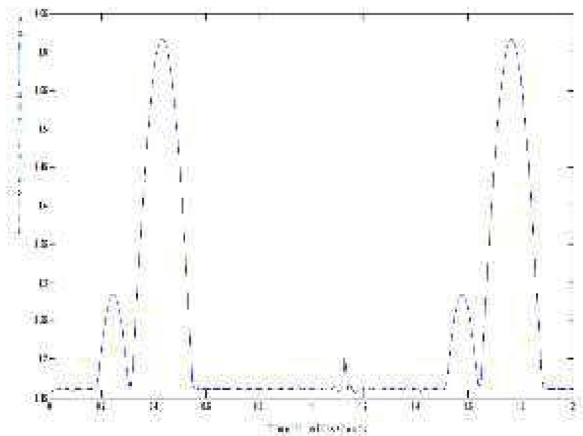


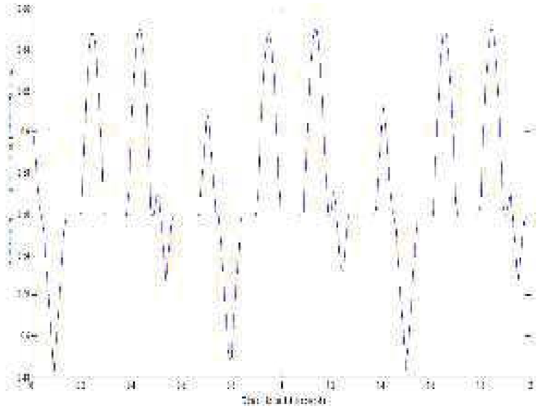
Fig 2.a) Normal Heart Beat



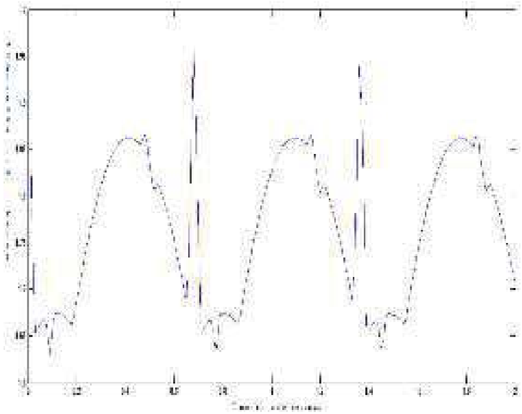
b) Atrial Flutter



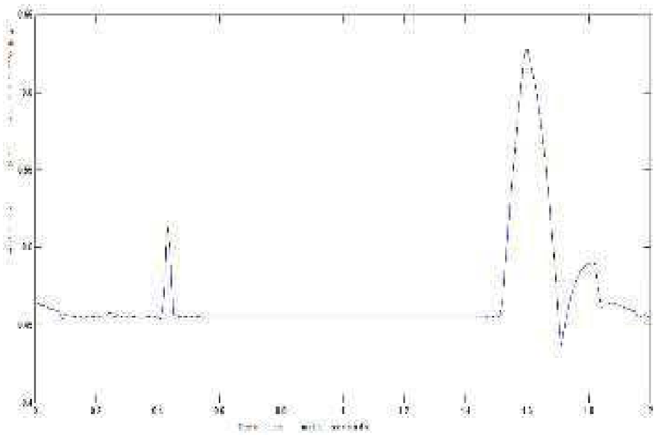
c) Premature Atrial Contraction



d) Premature Ventricular Contraction



e) Right Bundle Branch Block



f) Left Bundle Branch Block

IV Conclusion

In this paper, a novel ECG beat classification system using RVM is proposed and applied to MIT/BIH data base. From the obtained experimental results, it can be strongly recommended that the use of the RVM approach for classifying ECG signals on account of their superior generalization capability.This capability generally provides them with higher classification accuracies and a lower sensitivity to the curse of dimensionality. The results confirm that the RVM classification system substantially boosts the generalization capability. Another advantage of the RVM approach can be found in its high sparseness, which is explained by the fact that the adopted optimization criterion is based on minimizing the number of SVs. It can also be seen that RVM accomplishes better and more balanced classification for individual categories as well in very less training time. In future some advanced techniques can be used to train the RVM classifier and it may enhance the classification accuracy of the ECG and reduce the training time.

V References

[1] S., Osowski and T. H., Linh, 2001. ECG beat recognition using fuzzy hybrid neural network. IEEE Trans. Biomed. Eng., (11), 1265–1271.

[2] T. H., Linh, S., Osowski, and M. L., Stodoloski, 2003. On-line heart beat recognition using Hermite polynomials and neuron-fuzzy network. IEEE Trans. Instrum. Meas., (4), 1224–1231.

[3] F., de Chazal, M., O'Dwyer, and R. B., Reilly, 2004. Automatic classification of ECG heartbeats using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng.,(7), 1196–1206.

[4] L., Khadra, A. S., Al-Fahoum, and S., Binajjaj, 2005.A quantitative analysis approach for cardiac arrhythmia classification using higher order spectral techniques. IEEE Trans. Biomed. Eng., (11), 1840–1845.

[5] F., de Chazal and R. B., Reilly, 2006.A patient adapting heart beat classifier using ECG morphology and heartbeat interval features. IEEE Trans. Biomed. Eng., (12), 2535–2543.

[6] S., Mitra, M., Mitra, and B. B., Chaudhuri, 2006.A rough set-based inference engine for ECG classification.. IEEE Trans. Instrum. Meas., (6), 2198–2206.

[7] T., Inan, L., Giovangrandi, and J. T. A., Kovacs,2006. Robust neural network based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Trans. Biomed. Eng., (12), 2507–2515.

[8] C.-W., Hsu and C.-J., Lin, 2002.A comparison of methods for multiclass support vector machines. IEEE Trans. Neural Netw., (2), 415–425.

[9] Minami, K.., Nakajima, H., Toyoshima, T., 1999. Real-time discrimination of ventricular tachyarrhytmia with Fouriertransform neural network. IEEE Trans. Biomed. Eng., (46), 179–185.

[10] S., Osowski, T. H., Linh, and T., Markiewicz, 2004. Support vector machine based expert system for reliable heart beat recognition. *IEEE Trans. Biomed. Eng.*, (4), 582–589.