Premature Ventricular Beat Detection by Using Spectral Clustering Methods

BR Ribeiro[1], AM Marques[1], JH Henriques[1], MA Antunes[2]

[1]University of Coimbra, Coimbra, Portugal
[2]Department of Cardiothoracic Surgery, University Hospital, Coimbra, Portugal

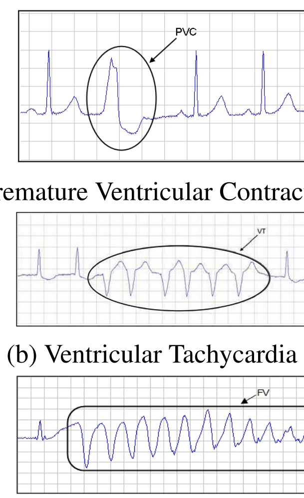
Abstract

In this paper, we the look at the spectral properties of features extracted from segmented ECG signals containing Normal (N) and premature ventricular beats (V) prior to apply classification methods for reliable PVC detection. In a first stage, feature extraction based on signal basic analysis which computes not only intervals and amplitudes on each beat, but also description of wave morphology was performed. Extracted parameters that describe the basic shape of the beat such as: average wave amplitudes, durations and areas have been computed. In a second stage, the eigen decomposition of data allows finding structure in records which is optimal to attain high performance of classification. In a third stage, Support Vector Machines (SVM) which are benchmarked against several techniques have been chosen for PVC detection. By applying SVM Recursive Feature Elimination (SVM RFE) where the weight magnitude is used as ranking criterion we reduced the feature dimension to smaller sets. Then, with newly constructed dimension input features space we combine spectral clustering with SVM classifiers for attaining superior performance.

1. Introduction

Arrhythmias are a disturbance in the rhythm of the heart that can range from a mild skipped beat to a life-threatening failure to pump. Among the ventricular arrhythmias, the premature ventricular contraction (PVC) is of great importance since if its occurrence is higher than normal it increases the risk of sudden death in patients.

Premature ventricular beat (PVB), also known as premature ventricular contraction (PVC) or extrasystole, is a form of irregular heartbeat in which the ventricle contracts prematurely. Ventricular tachycardia (V-tach or VT) is a fast rhythm that originates in one of the ventricles of the heart. This is a potentially life-threatening arrhythmia because it may lead to ventricular fibrillation and sudden death. Ventricular fibrillation (V-fib or VF) is a condition in which there is uncoordinated contraction of the cardiac muscle of the ventricles in the heart. As a result the heart fails to adequately pump blood and hypoxia will occur followed by unconsciousness within 20 - 30 seconds. Detection of above mentioned types of cardiac arrhythmia, illustrated in Figure 1, have been covered within a broad range of techniques ranging from autoregressive models [1], kth nearest neighbour rules [2] to neural fuzzy networks [3] among others. In this paper a new approach combining



(a) Premature Ventricular Contraction (PVC)

(b) Ventricular Tachycardia (VT)

(c) Ventricular Fibrillation (VF)

Figure 1. Ventricular Arrhythmias

spectral clustering with Support Vector Machines (SVM) following an appropriate feature extraction and selection on data provided from MIT-BH Arrhythmia databases is presented. In the following section we describe briefly the Recursive Feature Elimination (RFE) for feature selection, Spectral Clustering which finds structure in data, and the SVM for PVC classification. In the section 3 the experimental setup is described, followed by the feature extraction and feature selection procedures, spectral PVB clustering and SVM results. In the last section conclusions are presented with some lines of future research.

2. Methods

2.1. SVM-recursive feature elimination

The Recursive Feature Selection (RFE) is an iterative method introduced in [4]. The criterium estimates the effect of removing one feature at each iteration on the objective function J. The iterative procedure is an instance of backward feature elimination [5] and can be summarized in three steps: 1) Train the classifier (optimize the weights w_i with respect to J); 2) Compute the ranking criterion for all features ($DJ(i)$ or $(w_i)^2$); 3) Remove the feature with smallest ranking criterion. For computational reasons, several features can be removed at a time, possibly with classification performance degradation. In such a case, the method produces a feature subset ranking, as opposed to a feature ranking. SVM-RFE is an application of RFE using the weight magnitude as ranking criterion by the SVM.

2.2. Spectral clustering methods

Spectral clustering is a simple yet powerful method which finds structure in data using spectral properties of a pairwise matrix. The concept initially inspiring spectral clustering is graph partitioning [6]. The algorithm typically constructs an affinity matrix from the data, takes its eigen-decomposition, and then uses traditional clustering techniques, such as K-means, to a subspace of those eigenvectors. This subspace is found by specifying that there are K clusters, and thus using the first K eigenvectors as the clustering space. The elements of the $N \times N$ affinity matrix are the pairwise similarities of the data points calculated by using a kernel $W(i,j) = e^{-d(x_i,x_j)/2\sigma^2}$ with σ being the kernel width and $d(x_i,x_j) = ||x_i - x_j||^2$. The affinity matrix is then normalized to form a matrix $L = D^{-1/2}WD^{-1/2}$ where $D = diag(\sum_{j=1}^{d} W_{ij})$. L is positive definite with eigenvalues smaller or equal to 1. The first K eigenvectors are computed and arranged as columns in a matrix $\hat{Y}$. The rows of $\hat{Y}$ are then normalized and treated as K dimensional vectors; performing KMeans on these vectors will return the desired clustering.

2.3. Support vector machines

Support Vector Machines (SVM) are powerful kernel-based learning machines which combine essentially two strong ideas: maximum margin classifiers with low capacity and implicit features spaces defined by kernel functions [7]. In other words, they conjoint the following properties: low Vapnik-Chervonenkis (VC) dimension solutions through maximization of the margin and kernel nonlinearity. These properties bring good generalization to the Vapnik learning machine. Given a training data set consisting of input-output pairs $\{\mathbf{x}_n, t_n\}_{n=1}^{N}$ SVM use the convolution of the scalar product to build, in input space, the nonlinear decision functions of the form:

$$f(\mathbf{x}) = \sum_{n=1}^{N} w_n K(\mathbf{x}, \mathbf{x}_n) + w_0 \quad (1)$$

K represents the kernel (or mapping) function, which must be a positive semi-definite matrix:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)\dot{\phi}(\mathbf{x}_j) = \exp(-\beta||\mathbf{x}_i - \mathbf{x}_j||^2) \quad (2)$$

ϕ is the mapping function from input space to feature space, β is given by Gaussian kernel width, and the weights are the non-zero Lagrange multipliers, called support vectors (SV).

3. Results

3.1. Experimental setup

Experimental data used to test the proposed approaches were taken from MIT-BIH Arrhythmia Database[1]. Extensive pre-processing of the data files allowed the construction of the training, test and validation data sets each one consisting of 19391 sample data points [8].

3.2. Performance measures

To evaluate results and to compare classifiers, F1 and AUC metrics were used. F1 can be described as

$$F1 = \frac{2 \times Recall \times Precision}{Precision + Recall}, \quad (3)$$

where Precision and Recall can be computed using the true positives (TP), false positives (FP) and true negatives (TN).

$$Precision = \frac{TP}{TP+FP} \quad Recall = \frac{TP}{TP+FN} \quad (4)$$

F1 measure was preferred to simple accuracy or error measures, since it circumvents misinterpretations, when the datasets are unbalanced, i.e the metric is reliable even for skewed datasets.

3.3. Feature extraction

Table 1 shows 18 extracted features, computed from the output of the QRS complex detector. Figure 2 illustrates graphically the QRS complex where the extracted features, in the previous Table, are also indicated for better clarification. These features are able to distinguish normal from abnormal PVC beats, as will be shown in Section 3. With the feature extraction technique we were able to reduce the size of the descriptive feature vector and, hence, the size of the input space. However, as it will be shown, a step forward will be done for further reduction.

[1] http://www.physionet.org/physiobank/database/html/mitdbdir/

Feature	Description
RRav	RR mean interval
RR0	Last RR interval
SN	Signal/Noise estimation
Ql	Q-wave length
(Qcx, Qcy)	Q-wave mass centre (x,y) coordinates
(Qpx, Qpy)	Q-wave peak (x,y) coordinates
Rl	R-wave length
(Rcx, Rcy)	R-wave mass centre (x,y) coordinates
(Rpx, Rpy)	R-wave peak (x,y) coordinates
Sl	S-wave length
(Scx, Scy)	S-wave mass centre (x,y) coordinates
(Spx, Spy)	S-wave peak (x,y) coordinates

Table 1. Extracted features from ECG signal

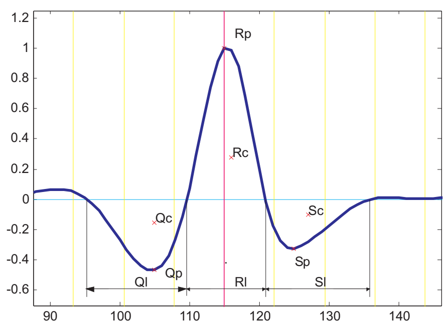


Figure 2. QRS Complex

3.4. Feature selection

In the experiments below we reduce the size of the feature vector and our findings show that the optimal number of selected features is 10 according to the results in Figure 3. The method used is an embedded method RFE-SVM which eliminates, in each iteration, one feature in agreement with a criteria; the smallest ranking value is chosen to eliminate the corresponding feature.

3.5. Spectral PVB clustering

Figure 4 illustrates spectral properties of data which map into (V) and (N) clusters. Two features (see Table 1) are plotted to show desired clusters. Spectral clustering applied on the reduced 10 dimension features set shows their discriminant ability across clusters spread.

3.6. Choosing kernels in SVM

Different kernels were used in the setup of the SVM learning machines and the algorithms run under the implementation of LIBSVM [9] and WEKA [10]. Table 2 shows that Gaussian kernel is the best regarding correct classified

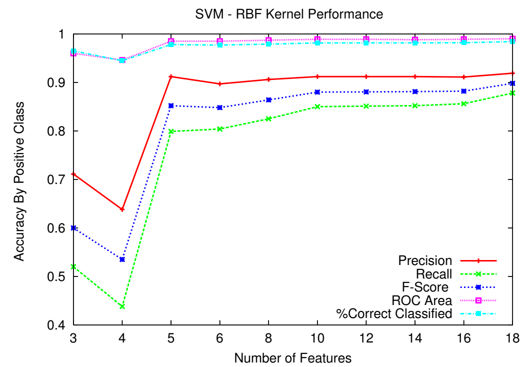


Figure 3. RFE-SVM Feature Selection

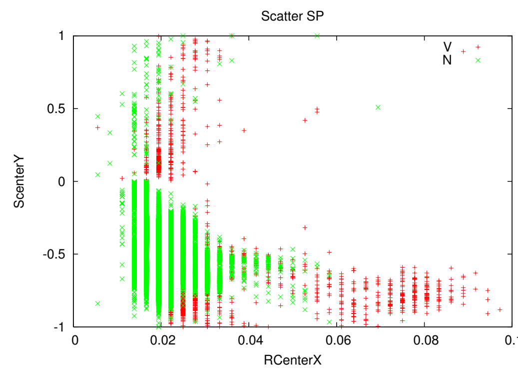


Figure 4. Spectral properties of data

instances (CC), Root Mean Square Error (RMSE), F1 and Area Under Curve (AUC). The results were obtained by 10-fold cross-validation. For a fair comparison, in Table 3,

Table 2. Accuracy, RMSE, F1, AUC.

	Linear	Polynomial	Gaussian	Sigmoid
CC	0.976	0.984	0.990	0.935
RMSE	0.158	0.128	0.098	0.248
F1-Score	0.834	0.892	0.936	0.323
AUC	0.878	0.925	0.959	0.717

we include also the results from neural networks, Multi-layer Perceptron (MLP) and Radial Basis Function (RBF), and Naive Bayes. Percent of correct classified samples are presented for training, testing and validation data sets. We have used 10 runs of 10-fold cross validation resulting in 100 evaluations of the learning machines. The mean and standard deviations for the specified measures are evaluated and compared. After a statistical t-test performed with a confidence level of 0.05, using two data sets (10 Features and all Features), the results (see Figure 5) indicate that the mean value of predicted performances is significantly better for SVM with Gaussian kernel whereas the neural network approaches (MLP and RBF) are worst. For

Table 3. Learning Machines Performance.

	SMO			LIBSVM			MLP			RBF		
	Train	Test	Valid	Train	Test	Valid	Train	Test	Valid	Train	Test	Valid
CC	0.995	0.989	0.992	0.991	0.985	0.989	0.990	0.985	0.986	0.987	0.981	0.984
RMSE	0.071	0.130	0.090	0.097	0.120	0.125	0.089	0.106	0.103	0.098	0.119	0.115
F1	0.968	0.933	0.959	0.939	0.908	0.926	0.935	0.904	0.927	0.915	0.884	0.895
AUC	0.974	0.962	0.969	0.960	0.944	0.956	0.996	0.991	0.992	0.955	0.990	0.943

the sake of a fair comparison we included also the Naive Bayes baseline classifier. We show that reducing the number of features by appropriate selection based on data spectral properties do not affect performance significantly.

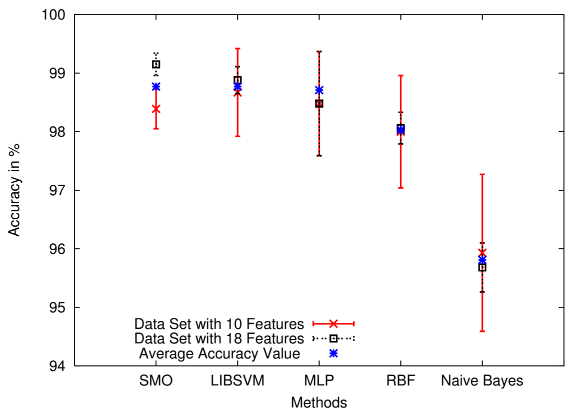


Figure 5. Performance Comparison

4. Discussion and conclusions

We propose 18 input features for arrhythmia detection and, by applying SVM Recursive Feature Elimination (SVM RFE) we reduced the feature dimension to smaller sets. Then, with newly constructed dimension input features space we combine spectral clustering with SVM classifiers for attaining superior performance. The results demonstrate superior ability performance using records of MIT Arrhythmia Database. We obtained high beat detection performance with F1-score of 95.9% and a positive predictability of 99.2%. Moreover, premature ventricular contraction beats were detected using an original approach combining a spectral clustering method followed by a classification strategy. The performances obtained allow us to point out the advantages of our approach according to the state of the art. The results obtained validate our approach for real world application. Future work will study the influence of the spread of clusters on results improvement.

Acknowledgements

MyHeart European Project IST-2002-507816 is gratefully acknowledged for partial financing support.

References

[1] Ge D, Srinivasan N, Krishnan SM. Cardiac arrhythmia classification using autoregressive modeling. Biomed Eng Online 2002;1–5.

[2] Christov I, Jekova I, Bortolan G. Premature ventricular contraction classification by the kth nearest-neighbours rule. Physiol Meas 2005;26:123–130.

[3] Ramirez-Rodriguez CA, Hernandez-Silveira MA. Multithread implementation of a fuzzy neural network for automatic ECG arrhythmia detection. Comp Cardiol 2001; 28:297–300.

[4] Guyon I, Weston J, S SB, V.Vapnik. Gene selection for cancer classification using SVM. Machine Learning 2002; (46):389–422.

[5] Kohavi R, John G. Wrappers for feature subset selection. Artificial Intelligence 1997;273–324.

[6] J. JS, Malik J. Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis 2000;22(8):888–905.

[7] Vapnik VN. The nature of statistical learning theory. New York: Springer-Verlag, 1995.

[8] Ribeiro B, MArques A, Henriques J, Antunes M. Choosing real time predictors for ventricular arrhythmia detection. International Journal of Pattern Recognition and Artificial Intelligence (to appear);.

[9] Chang CC, Lin CJ. LIBSVM: a library for support vector machines. Technical report, Computer Science and Information Engineering Dep, National Taiwan University, 2000.

[10] Witten IH, Frank E. Data Mining Practical Machine Learning Tools and Techniques. Elsevier, 2005.

Address for correspondence:

Bernardete Ribeiro
Department of Infomatics Engineering, DEI-FCTUC, Polo II, 3030-290 Coimbra, Portugal
bribeiro@dei.uc.pt