Computer-aided Morphological Analysis of Holter ECG Recordings Based on Support Vector Learning System

S Jankowski[1], A Oręziak[4], A Skorupski[2], H Kowalski[2], Z Szymański[2], E Piątkowska-Janko[3]

[1] Institute of Electronic Systems, [2] Institute of Computer Science, [3] Institute of Radioelectronics - Warsaw University of Technology, Warsaw, Poland
[4] Chair and Department of Internal Medicine and Cardiology, Central Teaching Hospital, Warsaw, Poland

Abstract

The paper presents a new approach to computer-aided analysis of ECG Holter recordings. In contrast to existing tools it is a learning system: the pertinent features of the signal shape are automatically discovered upon the examples carefully selected and commented by cardiologists.

Mathematical basis of our system is the theory of support vector machines that are applied for two tasks: signal approximation and pattern classification. Numerical procedures implement the algorithm of sequential minimal optimisation. The computer program is developed in Borland C++ Builder environment.

The excellent performances of our approach, high rate of successful pattern recognition and computational efficiency, make use of our tools possible in clinical practice. The system is tested at the Chair and Department of Internal Medicine and Cardiology, Central Teaching Hospital in Warsaw, Poland.

1. Introduction

Automatic recognition of cardiac pathologies from the investigation of Holter ECG recordings is usually based on the analysis of heart rate variability and mainly on the shape of QRS complexes and ST segment.

Objective of this work is to provide a new software package for computer-aided diagnosis of ECG Holter recordings. Our toolbox enables to consider the shape of entire beat including P, QRS and T waves. These details are of importance in automatic discrimination of both the atrial and ventricular depolarization.

We present the complete toolbox for computer-aided diagnosis of Holter recordings. Its mathematical basis is the theory of support vector machines (SVM).

Instead of carefully defined specific parameters of the signal shape for every considered case, our system is based on a set of examples.

Mathematical methods of SVM approximation enables to encode all significant details automatically upon the examples selected and commented by cardiologists.

We claim that practically all pathologies that can be visually recognised by specialists can be automatically detected by our approach. However, there are special requirements for users, as: proper selection of the learning set. The learning set must be representative for a studied pathology with respect to the morphological variability and the sufficient number of examples.

It can be concluded from the Cover theorem that the number of example in each class should be geater than the double number of the parameters of the heartbeat representation [5, 9].

2. Method

Support vector machine (SVM) is the new paradigm of learning system. The classification problem is defined as a quadratic optimisation in order to obtain a border between classes with the maximum margin. The ECG Holter recordings is filtered and segmented into single beats by use of the wavelet analysis to detect R-points. A learning set is defined by assigning labels to selected beats by cardiologists. Then the SVM approximation of all beats is performed using Gaussian kernels. This procedure transforms the digital signal into vectors of Lagrange multipliers. Then the SVM classifier is trained.

Support vector machine performs new ideas of supervised learning from examples. Let:

$X\ R^n$ - the input space, Y - the output domain,

for binary classification $Y = [-1, 1]$,

for function approximation $Y \subseteq R$.

A training set consists of a collection of training examples:

$S = ((x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)) \subseteq (XY)^l$

where l - number of examples, x - examples, y - their labels.

2.1. Support vector classification

Given a linearly separable set S the optimal separating hyperplane (w,b) solves the following optimization problem:

$$\min_{\mathbf{w},b} (\mathbf{w}\cdot\mathbf{w})$$

subject to

$$y_i((\mathbf{w}\cdot\mathbf{x}_i)+b)\geq 1, \quad i=1,\ldots,l$$

It realizes the maximal margin hyperplane with the geometric margin $\gamma=1/\|\mathbf{w}\|_2$. This problem can be solved by introducing Lagrange multipliers $\alpha_i \geq 0$ and a primal Lagrangian L:

$$L(\mathbf{w},b,\boldsymbol{\alpha})=\frac{1}{2}(w\cdot w)-\sum_{i=1}^{l}\alpha_i[(y_i(w\cdot x_i)+b)-1], \quad \alpha_i\geq 0$$

The Lagrangian L has to be minimized with respect to the primal variables w and b and maximized with respect to the dual variables α_i. The dual form is the following:

$$L(\mathbf{w},b,\boldsymbol{\alpha})==\sum_{i=1}^{l}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{l}y_iy_j\alpha_i\alpha_j(\mathbf{x_i}\cdot\mathbf{x_j})$$

The optimal separating hyperplane can be described as a linear combination of the training points:

$$\mathbf{w}^*=\sum_{i=1}^{l}y_i\alpha_i^*\mathbf{x_i}$$

This expansion consists of only a small subset of data from the training set that correspond to non-zero Lagrange multipliers - these points are called the support vectors.

The optimal separating hyperplane can be expressed in the dual representation in terms of this subset of parameters:

$$f(\mathbf{x},\alpha^*,b^*)=\sum_{i=1}^{l}y_i\alpha^*(\mathbf{x_i}\cdot\mathbf{x})+b^*==\sum_{i\in SV}y_i\alpha^*(\mathbf{x_i}\cdot\mathbf{x})+b^*$$

The Lagrange multipliers associated with each point become a dual variables. Points that are not support vectors have no influence.

2.2. Support vector approximation

In order to approximate functions, an -insensitive loss function will be used: The learning phase of the machine corresponds to fitting an elastic tube onto the training points - the data of the problem. If some points lie outside the chosen tube the slack variables ξ_i, ξ_i^* are introduced to deal with this case. The value of C determines how much deviations larger than ε are tolerated with respect to the tube.

The key idea is to construct a Lagrange function from both the objective function and the constraints, by introducing a set of dual variables:

$$\mathrm{L}=\frac{1}{2}\|\mathbf{w}\|^2+C\sum_{i=1}^{m}(\xi_i+\xi_i^*)-\sum_{\mathrm{i}=1}^{\mathrm{m}}\alpha_i(\varepsilon+\xi_i-\mathrm{y_i}+\langle\mathbf{w},\mathbf{x}_\mathrm{i}\rangle+\mathrm{b})-\sum_{i=1}^{m}\alpha_i^*(\varepsilon+\xi_i^*+\mathrm{y_i}-\langle\mathbf{w},\mathbf{x}_\mathrm{i}\rangle-\mathrm{b}$$
$$-\sum_{\mathrm{i}=1}^{\mathrm{m}}(\eta_i\xi_i+\eta_i^*\xi_i^*)$$

The dual variables are α_i, α_i^*, η_i, η_i^* and have to satisfy positive constraints (α_i, α_i^*, η_i, $\eta_i^* \geq 0$).

The primal variables are w,b and ξ_i, ξ_i^* are the slack variables and η_i, η_i^* have been introduced to induce their positivity

We use the algorithm that quickly solves the support vector machine problem - sequential minimal optimisation (SMO) - extreme decomposition of the problem that involves two Lagrange multipliers only at one step [4].

3. Numerical procedures

Database: ECG Holter recordings from the Department of Internal Medicine and Cardiology at the Central Teaching Hospital in Warsaw.

The signals are digitized with 128 Hz sampling rate and filtered with a 514 coefficients digital FIR filter.

The tools developed in this work perform following tasks: segmentation, SVM approximation, parameterization, SVM classification.

3.1. Segmentation

Given a complete filtered ECG Holter recording as input, the segmentation tool can divide it into a series of single heartbeats.

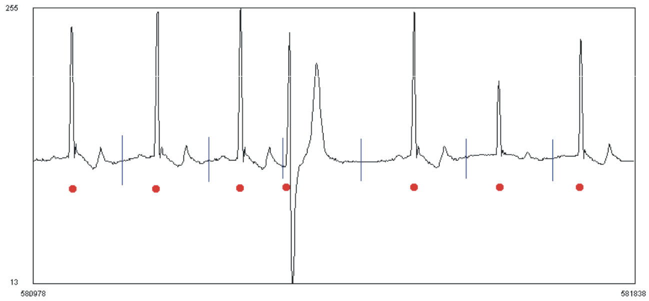

Figure 1. Visual application of the segmentation idea. Heartbeats can be identified by finding their QRS complexes (indicated by red dots). Segmentation has to be made after the T waves, as indicated by the blue lines. Note that the fourth beat is ectopic (ventricular beat without P wave). In this case segmentation has to be made near to its QRS complex

3.2. SVM approximation

The support vector machine provides an adequate approximation of the waveform as a sum of Gaussian functions:

$$k(x_i,x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x_i-x)^2}{2\sigma^2}\right)$$

The Lagrange multipliers in this approximation can substitute the discrete samples representing the signal. This new kind of processing stresses the characteristics of a single beat. The parameters are homogeneous - they all state the height or weight of a Gaussian function at a certain position. This position is inherent to the data as it depends on the position of the parameter within the series. Then the support vector expansion is as follows:

$$f(x) = \sum_{i=1}^{m} (\alpha_i - \alpha_i^*) k(x_i, x) + b$$

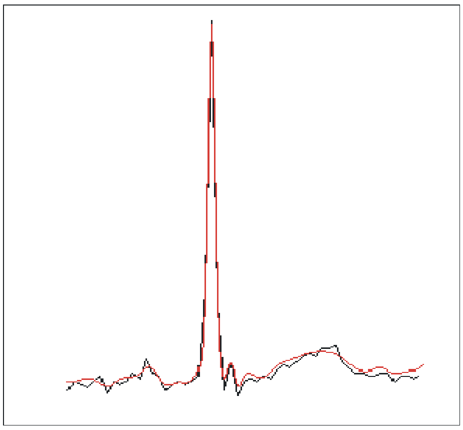

Figure 2. Smooth approximation (in red) of a noisy heartbeat (in black) using σ^2 =0.3 for the QRS complex and σ^2 =1.5 for P and T waves.

3.3. Parameterization

Selection of the multipliers according to their position within the beat. The selection provides a meaningful representation of a heartbeat for the SVM classifier.

Each heartbeat can be represented by 30 parameters:
P wave - 7 (to 20) parameters,
QRS complex - 7 parameters,
ST interval and T wave - 16 parameters.

E.g. for atrial fibrillation we focus our attention on P wave only using 20 parameters. The reduced representation improves the performance of the SVM classification.

3.4. SVM classification

Classification consists of two steps: learning and testing. Our classifier is a learning machine of the supervised type. In the learning phase it receives some patterns as input. These patterns are heartbeats represented by m parameters that can be seen as points in m-dimensional space. Each point has a label assigned by cardiologists that indicates its class corresponding to a type of pathology or at least to some precise feature within the beat shape. Then the machine becomes able to find the labels of new vectors by comparing them with those used in the learning phase.

4. Results

In order to estimate the system functionality we examined one case: the ECG Holter recordings of a patient having sinus rhythm with multifocal ventricular contractions. The heartbeats were assigned into 3 classes corresponding to: normal beats and 2 classes of pathological shapes from 2 focal ventricular contraction, probably LBBB - left bundle branch block and RBBB - right bundle branch block, as shown in fig. 3. Learning set consisted of 1712 heartbeats. We obtained 3 linear classifiers using one-against-all scheme with the following numbers of support vectors: normal against all pathological beats: 46; RBBB-like morphology against all other beats: 36; LBBB-like morphology against all other beats: 20. The test results are listed in Table 1.

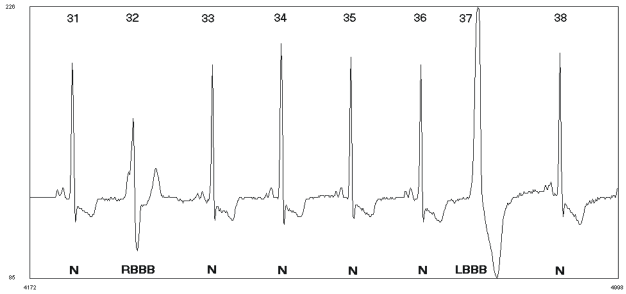

Figure 3. Several heartbeats with class labels assigned by a cardiologist.

Table 1. Classification validation results

	Normal	RBBB like	LBBB like	Total
Number of beats	8584	244	1044	9852
Correctly classified beats	8424	231	1035	9690
Misclassified beats	22	11	2	35
Unclassified	118	2	7	127
Rate of success	98,1	94,7	99,1	98,3

5. Examples of user's friendly interface

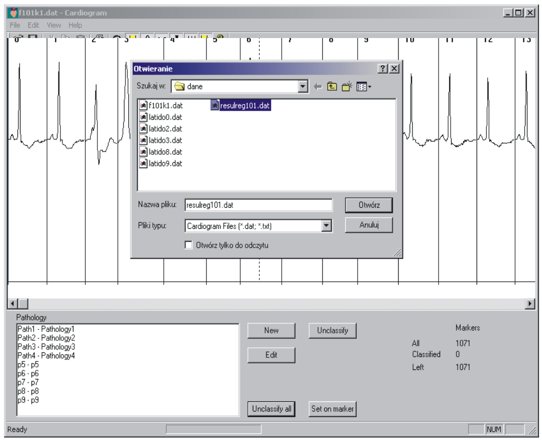

Figure 4. Example window of Cardiogram application - selection of data set.

Figure 5. Cardiogram application - heartbeat selection
.

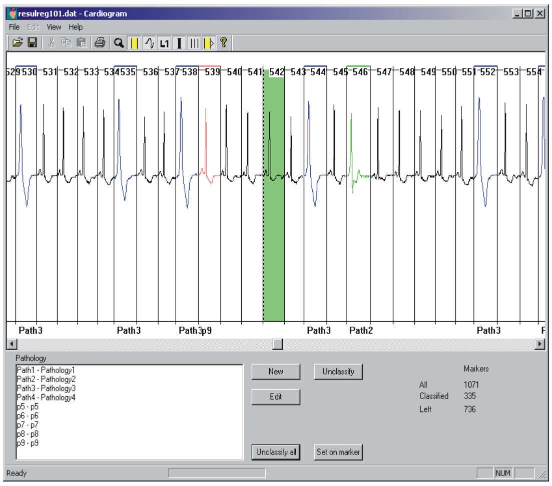

Figure 6. Cardiogram application - assigning class label to selected heartbeat

Figure 7. Cardiogram application - results of classification made by cardiologist.

6. Conclusions

Our approach gives the excellent performances of successful recognition. The learning system for ECG computer-aided analysis is a flexible tool. The cardiologists can define the goals of classification by the choice of learning sets with respect to all details of the electrocardiograms. The system is tested at the Chair and Department of Internal Medicine and Cardiology, Central Teaching Hospital in Warsaw, Poland.

References

[1] Vapnik VN. Statistical Learning Theory. New York: Wiley, 1998.

[2] Schölkopf B, Burges CJC, Smola AJ. Advances in Kernel Methods - Support Vector Learning. Cambridge: MIT Press, 1999.

[3] Cristianini N, Shaw-Taylor J. Support Vector Machines, Cambridge University Press, 2000.

[4] Platt J. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. In: Schölkopf B, Burges CJC, Smola AJ. Advances in Kernel Methods. Cambridge: MIT Press, 1998.

[5] Cover TM. Geometrical and Statistical Properties of Systems of Linear Inequalities with Application to Pattern Recognition. IEEE Trans. Electronic Computers 1965;6:326-334.

[6] Devroye L, Gyorfi L, Lugosi GA. Probabilistic Theory of Pattern Recognition. New York, Berlin Heidelberg: Springer-Verlag, 1996.

[7] Jankowski S, Tijink J, Vumbaca G, Balsi M, Karpinski G. Support vector machine for the recognition of atrial and ventricular depolarization in Holter ECG recordings. IJBEM 2002;4:343-344.

[8] Jankowski S, Tijink J, Vumbaca G, Balsi M, Karpinski G. Morphological analysis of ECG Holter recordings by support vector machines. Lecture Notes in Computer Science 2002;2526:134-143.

[9] Dreyfus G, Martinez JM, Samuelides M, Gordon MB, Badran E, Thiria S, Hérault L. Réseaux de neurones. Paris: Eyrolles, 2002.

Address for correspondence.

Stanisław Jankowski
Institute of Electronic Systems
ul. Nowowiejska 15/19
00-665 Warszawa, Poland
sjank@ise.pw.edu.pl