ELSEVIER

Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal

Babak Mohammadzadeh Asl [a,*], Seyed Kamaledin Setarehdan [b], Maryam Mohebbi [a]

[a] Department of Biomedical Engineering, Tarbiat Modares University, Tehran, Iran
[b] Control and Intelligent Processing Center of Excellence, Faculty of Electrical and Computer Engineering, University of Tehran, P.O. Box 14395/515, Tehran, Iran

Summary

Objective: This paper presents an effective cardiac arrhythmia classification algorithm using the heart rate variability (HRV) signal. The proposed algorithm is based on the generalized discriminant analysis (GDA) feature reduction scheme and the support vector machine (SVM) classifier.
Methodology: Initially 15 different features are extracted from the input HRV signal by means of linear and nonlinear methods. These features are then reduced to only five features by the GDA technique. This not only reduces the number of the input features but also increases the classification accuracy by selecting most discriminating features. Finally, the SVM combined with the one-against-all strategy is used to classify the HRV signals.
Results: The proposed GDA- and SVM-based cardiac arrhythmia classification algorithm is applied to input HRV signals, obtained from the MIT-BIH arrhythmia database, to discriminate six different types of cardiac arrhythmia. In particular, the HRV signals representing the six different types of arrhythmia classes including *normal sinus rhythm, premature ventricular contraction, atrial fibrillation, sick sinus syndrome, ventricular fibrillation* and *$2°$ heart block* are classified with an accuracy of 98.94%, 98.96%, 98.53%, 98.51%, 100% and 100%, respectively, which are better than any other previously reported results.
Conclusion: An effective cardiac arrhythmia classification algorithm is presented. A main advantage of the proposed algorithm, compared to the approaches which use the ECG signal itself is the fact that it is completely based on the HRV (R—R interval) signal which can be extracted from even a very noisy ECG signal with a relatively high

* Corresponding author. Tel.: +98 912 4715235; fax: +98 21 88633029.
  E-mail address: b.mohammadzadeh@modares.ac.ir (B.M. Asl).

accuracy. Moreover, the usage of the HRV signal leads to an effective reduction of the processing time, which provides an online arrhythmia classification system. A main drawback of the proposed algorithm is however that some arrhythmia types such as left bundle branch block and right bundle branch block beats cannot be detected using only the features extracted from the HRV signal.

1. Introduction

Heart diseases are a major cause of mortality in the developed countries. Many different instruments and methods are developed and being daily used to analyze the heart behavior. One of the relatively new methods to assess the heart activity and to discriminate different cardiac abnormalities is to analyse the so-called heart rate variability (HRV) signal. HRV signal, which is generated from electrocardiogram (ECG) by calculating the inter-beat intervals, is a nonlinear and nonstationary signal that represents the autonomic activity of the nervous system and the way it influences the cardiovascular system. Hence, measurement and analysis of the heart rate variations is a non-invasive tool for assessing both the autonomic nervous system and the cardiovascular autonomic regulatory system. Furthermore, it can provide useful information about the current and/or the future heart deficiencies [1]. Therefore, HRV analysis can be considered as an important diagnostic tool in cardiology.

Several methods have been proposed in the literature for automatic cardiac arrhythmia detection and classification. Some examples of the techniques used include: threshold-crossing intervals [2], neural networks [3—10], wavelet transforms [11], wavelet analysis combined with radial basis function neural networks [12], support vector machines [13], Bayesian classifiers [14], fuzzy logic combined with the Markov models [15], fuzzy equivalence relations [16], and the rule-based algorithms [17]. Most of these studies [2—6,11—13] are based on the analysis of the ECG signal itself. In most methods, the various features of the ECG signal including the morphological features are extracted and used for classification of the cardiac arrhythmias. This is a time-consuming procedure and the results are very sensitive to the amount of the noise.

An alternative approach would be to extract the HRV signal from the ECG signal first by recording the R—R time intervals and then processing the HRV signal instead. This is a more robust method since the R—R time intervals are less affected by the noise. Different HRV signal analysis methods for cardiac arrhythmia detection and classification were introduced in the past. Tsipouras and Fotiadis [8] proposed an algorithm based on both time and time—frequency analysis of the HRV signal using a set of neural networks. Their method could only classify the input ECG segments as "normal" or "arrhythmic" segments without the ability to identify the type of the arrhythmia. Acharya et al. [16] employed a multilayer perceptron (MLP) together with a fuzzy classifier for arrhythmia classification using HRV signal. They could classify the input ECG segments into one of the four different arrhythmia classes. In [17], Tsipouras et al. proposed a knowledge-based method for arrhythmia classification into four different categories. The main drawback of their algorithm was the fact that the atrial fibrillation, which is an important life-threatening arrhythmia, was excluded from the ECG database.

In this paper a new arrhythmia classification algorithm is proposed which is able to effectively identify six different and more frequently occurring types of cardiac arrhythmia. These arrhythmias are namely the normal sinus rhythm (NSR), the premature ventricular contraction (PVC), the atrial fibrillation (AF), the sick sinus syndrome (SSS), the ventricular fibrillation (VF) and the $2°$ heart block (BII). The proposed algorithm is based on the two kernel learning machines of the *generalized discriminant analysis* (GDA) and the *support vector machine* (SVM). By cascading SVM with GDA, the input features will be nonlinearly mapped twice by radial basis function (RBF). As a result, a linear optimal separating hyperplane can be found with the largest margin of separation between each pair of arrhythmia classes in the implicit dot product feature space.

GDA is a data transformation technique which was first introduced by Baudat and Anouar [18]. It can be considered as a kind of generalization to the well-known *linear discriminant analysis* (LDA) algorithm and has become a promising feature extraction scheme [19—24] in recent years. The main steps in GDA are to map the input data into a convenient higher dimensional feature space F first and then to perform the LDA algorithm on the F instead of the original input space. By GDA therefore, both dimensionality reduction of the input feature space and selection of the useful discriminating features can be achieved simultaneously.

SVM, which was first proposed by Vapnik [25], has been considered as an effective classification

scheme in many pattern recognition problems recently [22–24,26,27]. It is often reported that SVM provides better classification results than other widely used methods such as the neural network classifiers [28,29]. This is partly because SVM aims to obtain the optimal answer using the available information and in the same time it shows better generalization ability on the unseen data.

In continue the details of the proposed algorithm for cardiac arrhythmia classification from the HRV signal is presented. Section 2 provides the overall block diagram of the proposed algorithm together with the details of each block. The results of the application of the proposed algorithm to the MIT-BIH arrhythmia database are presented in Section 3. Section 4, compares the results obtained by the proposed algorithm to those obtained by the other previously reported techniques. This is followed by a discussion on the results and the methods. Finally, Section 5 concludes the paper.

2. Materials and methods

2.1. Database

The HRV data used in this work is generated from the ECG signals provided by the MIT-BIH Arrhythmia Database [30]. The database was created in 1980 as a reference standard for serving all those who are conducting a research on the cardiac arrhythmia detection and classification problem [31].

The MIT-BIH Arrhythmia Database includes 48 ECG recordings each of length 30 min with a total of about 109,000 R–R intervals. The ECG signals were bandpass-filtered in the frequency range of 0.1–100 Hz and were sampled with a sampling frequency of 360 Hz. Each of the about 109,000 beats was manually annotated by at least two cardiologists working independently. Their annotations were compared, consensus on disagreements was obtained, and the reference annotation files were prepared [31]. The reference annotation files include beat, rhythm, and signal quality annotations. Due to the lack of the VF data in the MIT-BIH arrhythmia database, which is needed in the current study, the Creighton University Ventricular Tachyarrhythmia Database was added to the MIT-BIH data as the VF arrhythmia class after resampling it at a rate of 360 Hz.

Finally, a total number of 1367 ECG segments each with 32 R–R intervals were selected from the above-mentioned database and used in this work, which contains all six different arrhythmia classes considered in this study. The specialists defined rhythm annotations for each segment were also considered along with the segments.

2.2. The proposed algorithm

The block diagram of the proposed algorithm is demonstrated in Fig. 1. As seen, it comprises the four steps of preprocessing, feature extraction, GDA-based feature reduction and SVM-based arrhythmia classification. In continue, each block is described in more details.

2.2.1. Preprocessing

As a first step, it is necessary to extract the HRV signals from the ECG signals within the database. In general, this process can be affected by many interfering signals such as the mains 50 Hz, the interferences from electromyogram (EMG) signals and also the baseline wandering. The interfering signals are effectively eliminated from the input ECG signal using a 5–15 Hz bandpass filter. Furthermore, the cubic splines are used for baseline approximation, which is then subtracted from the signal [32].

Next, the tachograms are extracted from the filtered ECG signals as follows. Initially, using the Hamilton and Tompkins algorithm [33,34], a point within the QRS complex is detected (QRS point). Afterwards, the main wave of the QRS complex, i.e., the R wave, is identified by locating the maximum absolute value of the signal within the time window [QRS–280 ms, QRS + 120 ms]. The HRV signal is then constructed by measuring the time intervals between the successive R waves (R–R intervals). Plotting the R–R intervals against the time indices provides the so-called tachograms. The tachograms are then divided into small segments each containing 32 R–R intervals and characterized using the database rhythm annotation. It must be noted that the resulting tachograms are sequences of unevenly sampled beat-to-beat intervals. Therefore, for the case of the frequency domain analysis in the forthcoming Section 2.2.2, the cubic spline interpolation method is used at a sampling rate of 4 Hz to produce an evenly sampled data. This resampling procedure
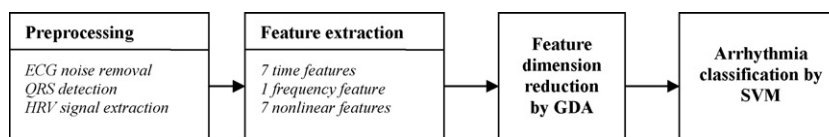

Figure 1 Block diagram of the proposed arrhythmia classification algorithm.

is necessary prior to using the well-known methods of power spectral density (PSD) estimation which are only applicable to the evenly sampled signals.

2.2.2. Feature extraction

The next step in the block diagram is the feature extraction step. In general, the cardiovascular system, hence the HRV signals, demonstrates both linear and nonlinear behavior. Different linear and nonlinear parameters are defined and used for HRV signal description. In this work, a combination of both linear and nonlinear features of the HRV signal is considered. Time and frequency domain features are among the standard linear measures of the HRV signals which are strongly recommended in a special report published by the Task Force of the European Society of Cardiology and North American Society of Pacing and Electrophysiology in 1996 [1]. As in most previous works, these features are used in the current study.

2.2.2.1. Linear analysis: time domain features. Seven commonly used time domain parameters of the HRV signal which are also considered in this work are as follows:

> *Mean*: This refers to the mean value of the 32 R—R intervals within each segment.
> *RMSSD*: This refers to the root mean square successive difference of the 32 R—R intervals in each segment.
> *SDNN*: This refers to the standard deviation of the 32 R—R intervals within each segment.
> *SDSD*: This refers to the standard deviation of differences between the adjacent R—R intervals within each segment.
> *pNN50, pNN10, pNN5*: These refer to the number of successive difference of intervals which differ by more than 50, 10 and 5 ms, respectively, divided by 32, the total number of the R—R intervals within each segment.

2.2.2.2. Linear analysis: frequency domain features. Although the time domain parameters are computationally effective but they lack the ability to discriminate between the sympathetic and parasympathetic contents of the HRV signal [35]. As the most popular linear technique used in the HRV signal analysis, however, the frequency domain analysis of the HRV signal has the ability to discriminate between the two. In fact, it is generally accepted that the spectral power in the high-frequency (HF) band (0.15—0.4 Hz) of the HRV signal reflects the respiratory sinus arrhythmia (RSA) and thus cardiac vagal activity. On the other hand, the low-frequency (LF) band (0.04—0.15 Hz), is related to the baror-

eceptor control and is mediated by both vagal and sympathetic systems [1,7]. In this work, the ratio of the LF and HF bands power (LF/HF) is used as the frequency domain feature of the HRV signal.

2.2.2.3. Nonlinear analysis. Seven different nonlinear parameters of the HRV signal are used in this work, which are listed and described in continue.*SD1/SD2*: Let us consider the HRV signal as a time series of the R—R intervals which is denoted by $RR(i)$. Now, if each interval $RR(n + 1)$ is plotted as a function of the previous interval $RR(n)$, then the resulting plot is known as the Poincaré plot, which is a relatively new tool for HRV signal analysis. A useful feature of this tool is that it does not require the HRV to be considered as a stationary signal [36]. Poincaré plot can be seen as a graphical representation of the correlation between the successive R—R intervals. This plot can be quantitatively analyzed by calculating the standard deviations of the distances of the points $RR(i)$ from the lines $y = x$ and $y = x + 2RR_m$, where RR_m is the mean of all $RR(i)$ values. These standard deviations are denoted by SD_1 and SD_2, respectively. In fact, SD_1 represents the fast beat-to-beat variability, while SD_2 describes the relatively long-term variability in the HRV signal [37]. The ratio SD_1/SD_2 is usually used to describe the relation between the two components.

ApEn: Approximate entropy (ApEn) is a measure of unpredictability of the fluctuations in a time series, and reflects the likelihood that particular patterns of observations will not be followed by additional similar observations. A time series containing many repetitive patterns has a relatively small ApEn, while a more complex (i.e., less predictable) process has a relatively high ApEn [38]. We have used the method proposed in [39] for calculating ApEn for each HRV segment by setting the pattern length $m = 2$ and the measure of similarity $r = 20\%$ of the standard deviation of the segment, as proposed in [40].

SpEn: Similar to ApEn, spectral entropy (SpEn) quantifies the complexity of the input time series (HRV segment) but in the frequency domain [41]. The Shannon's channel entropy is used in this work to obtain an estimate of the spectral entropy of the process as

$$H = -\sum_f p_f \log(p_f) \tag{1}$$

where p_f is the value of the probability density function (PDF) of the process at frequency f [7]. Heuristically, the entropy can be interpreted as a measure of uncertainty about the event at frequency f.

LLE: Lyapunov exponent is a measure of how fast two initially nearby points on a trajectory will

diverge from each other as the system evolves, hence providing useful information about the system's dependency on initial conditions [42]. A positive Lyapunov exponent strongly indicates that the system is a chaotic one [43,44]. Although an *m*-dimensional system has *m* Lyapunov exponents, however, in most applications it is sufficient to obtain only the average largest Lyapunov exponent (LLE) as follows. First, a starting point is selected within the reconstructed phase space of the system and then all those points residing within a neighborhood of a predetermined radius ε from the starting point are determined. Next, the mean distances between the trajectory of the initial point and the trajectories of the neighboring points are calculated as the system evolves. By plotting the logarithm of the above-mentioned mean values against the time index, the slope of the resulting line provides the LLE. To remove the dependency of the calculated values to the starting point, this procedure is repeated for different starting points and the average is taken as the average LLE [45] used as a feature to quantify the chaotic behavior of the HRV signal.

DFA: The detrended fluctuation analysis (DFA) is a useful parameter to quantify the fractal scaling properties of the R—R intervals. This technique is a modification to the root-mean-square analysis of the random walks applied to nonstationary signals [46]. For the detrended fluctuation analysis of the HRV signal, the R—R time series (of total length *N*) is integrated using the following equation first:

$$y(k) = \sum_{i=1}^{k}(RR(i) - RR_m) \qquad (2)$$

where $y(k)$ is the *k*th value of the integrated series, $RR(i)$ is the *i*th inter-beat interval and RR_m is the average inter-beat interval over the entire series. Then, the integrated time series is divided into windows of equal length *n*. In each window of length *n*, a least-square line is fitted to the R—R interval data (representing the trend in that window). The *y* coordinate of the straight line segments are denoted by $y_n(k)$. Next, the integrated time series within each window *n* is detrended. The root-mean-square fluctuation of this integrated and detrended time series is then calculated as

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2} \qquad (3)$$

This computation is repeated over all window sizes (time scales) to obtain the relationship between $F(n)$ and the window size *n* (i.e., the number of beats within the observation window). $F(n)$ is usually plotted against the observation window size *n* on a log—log scale. Typically, $F(n)$ increases with the window size. The fluctuation in small windows can be characterized by a scaling exponent (self similarity factor), α, which represents the slope of the line relating $\log F(n)$ to $\log n$ [35].

Sequential trend analysis: Sequential trend analysis of the HRV signal evaluates not only the sympathetic—parasympathetic balance but also provides the spectral analysis of the signal without the necessity to consider the signal is stationary. To perform the sequential trend analysis, it is necessary to plot $\Delta RR(n)$ against $\Delta RR(n+1)$ and divide the plane into four quadrants. The points located in the +/+ quadrant indicate two consecutive interval increments, which means the heart rate is decreasing and the ones in the −/− quadrant indicate two consecutive interval decrements, which means the heart rate is increasing. In this work the density of the points within the −/− and +/+ quadrants are used as two features that measure the sympathetic and parasympathetic activities, respectively [36].

2.2.3. Feature dimension reduction by GDA

Having defined the above-mentioned linear and nonlinear features, due to the large variations in the HRV patterns of various arrhythmia classes, there is usually a considerable overlap between some of these classes in the feature space. For example, the SSS and NSR classes demonstrate a large overlap with each other making it difficult to distinguish between the two. In this situation, a feature transformation mechanism that can minimize the within-class scatter and maximize the between-class scatter will be very beneficial. GDA [18] is such a transform which is employed in this work.

GDA is a nonlinear extension to the ordinary LDA. The input training data is mapped by a kernel function to a high-dimensional feature space, where different classes of objects are supposed to be linearly separable. The LDA scheme is then applied to the mapped data, where it searches for those vectors that best discriminate among the classes rather than those vectors that best describe the data [47]. In fact, the goal of the LDA is to seek a transformation matrix that maximizes the ratio of the between-class scatter to the within-class scatter. Furthermore, given a number of independent features which describe the data, LDA creates a linear combination of the features that yields the largest mean differences of the desired classes [48]. As a result, if there are *N* classes in the data set, the dimension of feature space can be reduced to $N-1$.

Let us assume that the training data set X contains M feature vectors out of N classes. Let $\mathbf{x}_{pq}$ denotes the qth HRV feature vector in the pth class, n_p is the class size of the pth class, and ϕ is a nonlinear mapping function so that the space X is mapped into a higher dimensional feature space: ϕ: $\mathbf{x}_i \in R^f \rightarrow \phi(\mathbf{x}_i) \in R^F, F \gg f$.

The observations $\phi(\mathbf{x}_i)$ are assumed to be centered in space F. Before projecting the training data set X into a new set Y by means of the GDA, two matrices of the within-class scatter matrix $\mathbf{V}$ and the between-class scatter matrix $\mathbf{B}$ in space F are defined as

$$\mathbf{V} = \frac{1}{M}\sum_{p=1}^{N}\sum_{q=1}^{n_p}\phi(\mathbf{x}_{pq})\phi^T(\mathbf{x}_{pq}) \quad (4)$$

$$\mathbf{B} = \frac{1}{M}\sum_{p=1}^{N}n_p\left(\sum_{q=1}^{n_p}\phi(\mathbf{x}_{pq})\right)\left(\sum_{r=1}^{n_p}\phi(\mathbf{x}_{pr})\right)^T \quad (5)$$

The purpose of the GDA is to find the projection vector $\mathbf{v}$ such that the inter-classes inertia is maximized and the intra-classes inertia is minimized in space F, which is equivalent to solving the following maximization problem:

$$\mathbf{v} = \arg\max_{v}\frac{\mathbf{v}^T\mathbf{B}\mathbf{v}}{\mathbf{v}^T\mathbf{V}\mathbf{v}} \quad (6)$$

The projection vector $\mathbf{v}$ is the eigenvector of the matrix $\mathbf{V}^{-1}\mathbf{B}$ associated with the eigenvalue $\lambda = \mathbf{v}^T\mathbf{B}\mathbf{v}/\mathbf{v}^T\mathbf{V}\mathbf{v}$. All solutions $\mathbf{v}$ lie in the span of $\phi(\mathbf{x})$. So, there exist expansion coefficients a_i such that
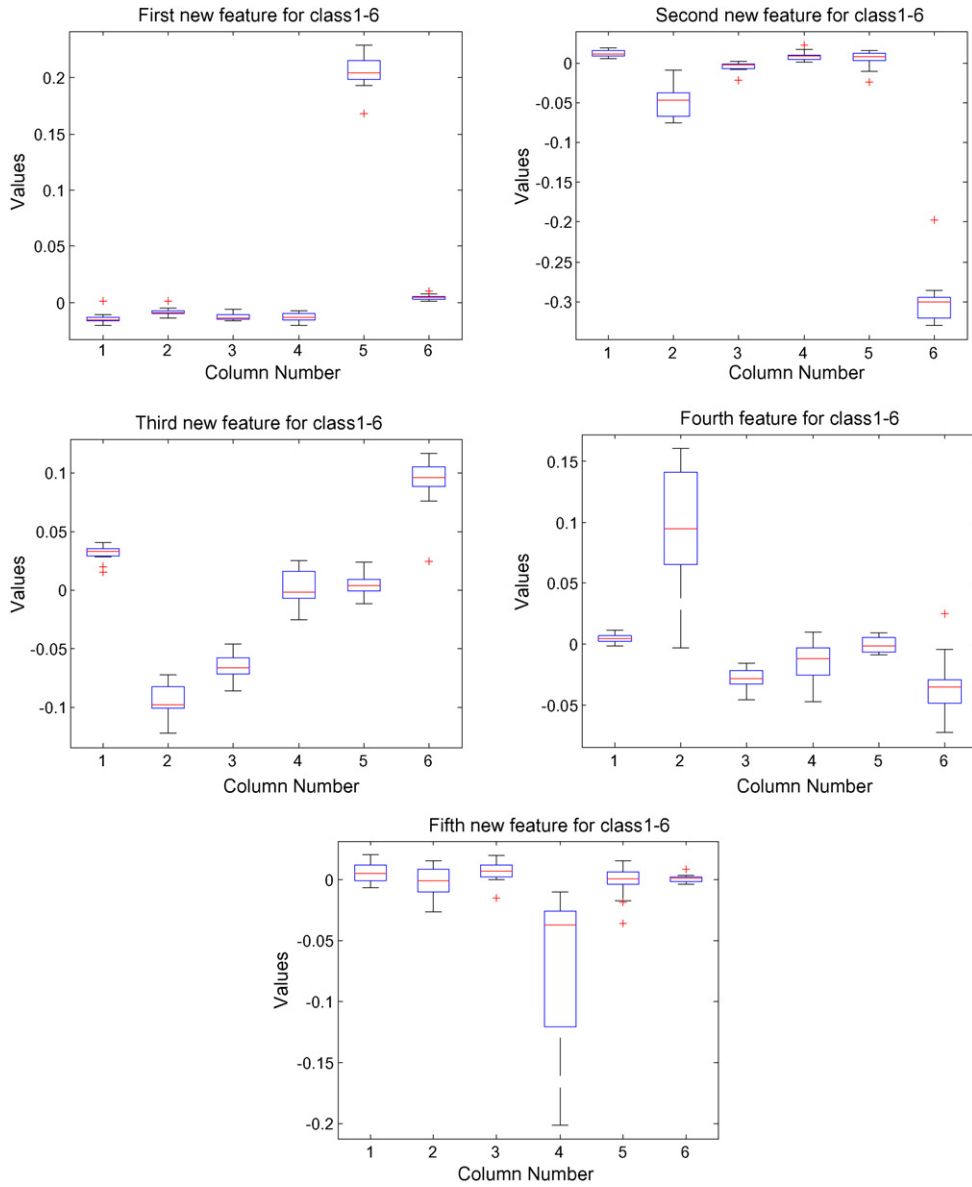

Figure 2 Box-plots of the new five features for different arrhythmia classes (1 = NSR, 2 = PVC, 3 = AF, 4 = SSS, 5 = VF, 6 = BII). For more information see the text.

$$\mathbf{v} = \sum_{i=1}^{M} a_i \phi(\mathbf{x}_i) \qquad (7)$$

By using the kernel function $k(\mathbf{x}_i, \mathbf{x}_j) = k_{ij} = \phi(\mathbf{x}_i)\phi(\mathbf{x}_j)$ and performing the eigenvectors decomposition on the kernel matrix $K = (k_{ij})_{i=1,\ldots,M;j=1,\ldots,M}$, M normalized expansion coefficients for each projection vector $\mathbf{a} = \mathbf{a}/(\mathbf{a}^T K\mathbf{a})^{1/2}$ is obtained.

Now, for a feature vector $\mathbf{x}$ from the test HRV data set, the projection on the ith eigenvector $\mathbf{v}_i$ can be computed as

$$\mathbf{y}^i = \mathbf{v}^{i^T}\phi(\mathbf{x}) = \sum_{j=1}^{M} a^i_j \phi(\mathbf{x}_j)\phi(\mathbf{x}) = \sum_{j=1}^{M} a^i_j k(\mathbf{x}_j, \mathbf{x}) \qquad (8)$$

where a^i_j denotes the jth expansion coefficient of the ith eigenvector. For the purpose of feature dimension reduction, $N-1$ eigenvectors associated with the first largest nonzero $N-1$ eigenvalues are selected to form the transformation matrix $\mathbf{W}^T = (\mathbf{v}^1, \ldots, \mathbf{v}^{N-1})$. So, each HRV feature vector is projected into a new coordinates using the $N-1$ projection vectors.
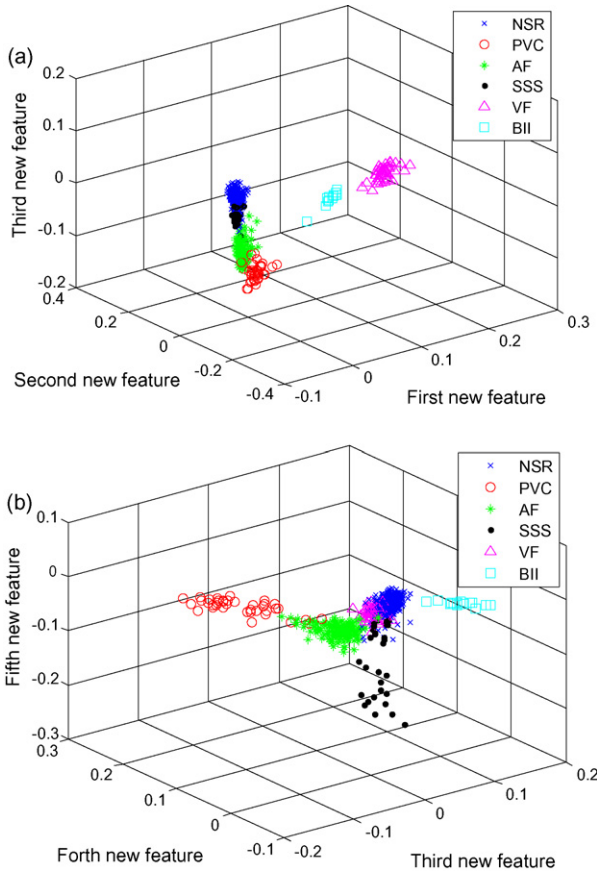
It is worth to notice that the optimal number of eigenvectors for the data transformation is generally equal to $N-1$ [22]. In this paper, assuming the number of classes (different arrhythmia to be identified) is 6, the number of original 15 features was designed to be reduced to 5 by GDA with a guarantee that the performance is comparable to that of the non-reduced feature set.

The box-plots and the feature space plots of the new five features for different arrhythmia classes, which are generated by this process, are presented in Figs. 2 and 3, respectively. As seen the patterns related to the different arrhythmia classes are located close to each other and relatively well separated from the other classes within the feature space. Therefore the new reduced feature set not only increases the classification procedure in the next step but also provides an appropriate tool for a better discrimination of the different arrhythmia classes. To demonstrate the usefulness of the GDA technique, a comparative study is carried out in the next section comparing GDA to the commonly used feature dimension reduction techniques of the principal component analysis (PCA) and the LDA.

2.2.4. SVM-based arrhythmia classifier

The next step in the block diagram is the classification of the HRV segments by considering their reduced features. Different classification methods have been used for cardiac arrhythmia classification in the past [2−17]. In this work the SVM scheme is used for classification. SVM is a machine-learning technique which has established itself as a powerful tool in many classification problems [22−24,26,27]. Simply stated, the SVM identifies the best separating hyperplane (the plane with maximum margins) between the two classes of the training samples within the feature space by focusing on the training cases placed at the edge of the class descriptors. In this way, not only an optimal hyperplane is fitted, but also less training samples are effectively used; thus high classification accuracy is achieved with small training sets [49].

Given a training set $(\mathbf{x}_i, y_i)$, $i = 1, 2, \ldots, l$, where $\mathbf{x}_i \in R^n$ and $y_i \in \{-1, 1\}$, the traditional SVM algorithm is summarized as the following optimization problem:

$$\min_{w,b,\xi} \left\{ \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\left(\sum_{i=1}^{l}\xi_i\right) \right\}$$

subject to : $y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad \zeta_i > 0 \ \forall i$

$$(9)$$

where $\phi(\mathbf{x})$ is a nonlinear function that maps $\mathbf{x}$ into a higher dimensional space [50]. w, b and ξ_i are the

Figure 3 Feature space plots of (a) first, second, and third new features; (b) third, fourth, and fifth new features for different arrhythmia classes (NSR: ✕, PVC: ○, AF: ✳, SSS: ●, VF: △, BII: □).

weight vector, bias and slack variable, respectively. C is a constant and determined a priori. Searching for the optimal hyperplane in (9) is a quadratic programming problem, which can be solved by constructing a Lagrangian and transforming it into a dual maximization problem of the function $Q(\boldsymbol{\alpha})$, defined as follows:

$$\max Q(\boldsymbol{\alpha}) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j)$$

$$\text{subject to}: \sum_{i=1}^{l} \alpha_i y_i = 0;$$

$$0 \le \alpha_i \le C, \quad \text{for } i = 1, 2, \ldots, l \qquad (10)$$

where $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ is the kernel function and, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_l)$ is the vector of nonnegative Lagrange multipliers.

Assuming that the optimum values of the Lagrange multipliers are denoted as $\alpha_{o,i}(i = 1, 2, \ldots, l)$, it is then possible to determine the corresponding optimum value of the linear weight vector $\mathbf{w_o}$ and the optimal hyperplane as in (11) and (12), respectively:

$$\mathbf{w_o} = \sum_{i=1}^{l} \alpha_{o,i} y_i \phi(\mathbf{x}_i) \qquad (11)$$

$$\sum_{i=1}^{l} \alpha_{o,i} y_i K(\mathbf{x}, \mathbf{x}_i) + b = 0. \qquad (12)$$

The decision function can be written as

$$f(\mathbf{x}) = \text{sign}\left(\sum_{i=1}^{l} \alpha_{o,i} y_i K(\mathbf{x}, \mathbf{x}_i) + b\right). \qquad (13)$$

In this work, the radial basis function (RBF) is used as the kernel function and the parameters — kernel width σ and regularization constant C — were experimentally defined to achieve the best classification result.

Although SVM separates the data only into two classes, classification into additional classes is possible by applying either the one against all (OAA) or one against one (OAO) methods. In the OAA method, a set of binary classifiers (k parallel SVMs, where k denotes the number of classes) is trained to be able to separate each class from all others. Then each data object is classified to the class for which the largest decision value has been determined. The OAO method constructs $k(k-1)/2$ parallel SVMs where each SVM is trained on the data from two classes [51]. Then, the voting strategy [51] aggregates the decisions and predicts that each data object is in the class with the largest vote. A comparative study is carried out in the next section comparing the performances of the OAO and OAA decomposition methods in training the proposed multiclass SVMs. In continue the results of the application of the proposed algorithm to the data set are presented.

3. Results

To evaluate the performance of the proposed arrhythmia classification algorithm, a total number of 1367 HRV segments are used which includes 835 NSR segments, 57 PVC segments, 322 AF segments, 50 SSS segments, 78 VF segments, and 25 BII segments. The relatively high percentage of the NSR segments in the data set is not far from reality as ECG recordings usually have a higher percentage of normal beats compared to arrhythmic segments. The HRV signals at each class are randomly divided into the train and test sets in an approximate ratio of 2/3 and 1/3. The exact number of the train and test segments for each class is shown in Table 1.

Table 1 The correct classification results on the test set for each class by the proposed algorithm together with the spread of the erroneous classifications into the other classes (these are the average of 100 train and test procedures)

Total number of train/test segments			Classification (GDA+SVM)					
			NSR	PVC	AF	SSS	VF	BII
555/280	Database annotation	NSR	277.90	0	0	2.1	0	0
38/19		PVC	0	18	1	0	0	0
214/108		AF	0	3.80	102.20	2	0	0
30/20		SSS	2.80	0	0	17.20	0	0
52/26		VF	0	0	0	0	26	0
15/10		BII	0	0	0	0	0	10

Table 2 Performance analysis of the SVM classifier on the original features (ORG) and the reduced features (by GDA) in terms of the average values of the four commonly used measures in % (numbers inside the parenthesis are the standard deviations)

Arrhythmia classes	GDA+SVM				ORG+SVM			
	Sensitivity (%)	Specificity (%)	Positive predictivity (%)	Accuracy (%)	Sensitivity (%)	Specificity (%)	Positive predictivity (%)	Accuracy (%)
NSR	99.25(0.11)	98.47(0.50)	99.00(0.23)	98.94(0.18)	97.86(0.23)	96.72(0.00)	97.86(0.01)	97.41(0.14)
PVC	94.74(0.00)	99.14(0.16)	82.57(2.61)	98.96(0.15)	100(0.00)	98.65(0.13)	76.00(3.06)	98.70(0.12)
AF	94.63(0.65)	99.72(0.00)	99.03(0.21)	98.53(0.15)	92.59(0.53)	99.72(0.12)	99.01(0.53)	98.06(0.15)
SSS	86(1.59)	99.07(0.12)	80.75(0.72)	98.51(0.18)	65(2.11)	98.19(0.14)	68.42(1.48)	96.76(0.15)
VF	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)
BII	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)	100(0.00)
Average	95.77(0.39)	99.40(0.13)	93.56(0.63)	99.16(0.11)	92.57(0.48)	98.88(0.06)	90.21(0.85)	98.49(0.09)

The 15 aforementioned linear and nonlinear arrhythmia features are then calculated for each HRV segment in the train and test sets. Next, the 15 original features are reduced to only 5 new features by means of the GDA algorithm. A radial basis function (RBF) is used as the kernel where a kernel width of 7 was chosen empirically. Afterwards, the SVM classifier is trained using the reduced feature vectors of the training set. To optimize the learning cost and the classification performance, the SVM classifier parameters, the kernel width σ and the regularization constant C, have to be chosen appropriately. For this purpose, the training data set itself is divided into the train and validation sets. The optimum values of the parameters σ and C are then chosen such that the minimum error is achieved on the validation data set. The resulting optimum values were 0.08 and 30 for σ and C, respectively. The test data for each class is then used for performance analysis of the classifier.

The whole procedure including randomly dividing the data set into the train and test sets, training the classifier and testing it by the test data set was repeated 100 times. The average correct classification results over the 100 runs on the test set for each class together with the spread of the erroneous classifications into the other classes are shown in Table 1. As seen, for the NSR, in average only 2.1 segments are misclassified as the SSS (0.75%). For the PVC, in average only one segment is misclassified as the AF (5.26%). For the AF, in average 3.8 segments are misclassified as the PVC (3.52%) and 2 segments as the SSS (1.11%). For the SSS, in average 2.8 segments are misclassified as the NSR (14%). For the VF and the BII, there is no misclassification to any other classes.

For a more-detailed performance analysis of the proposed algorithm, in continue, the four commonly used measures of *sensitivity, specificity, positive predictivity*, and *accuracy* [9,10] are derived for the proposed GDA + SVM-based algorithm. Furthermore, for comparison purposes these parameters are also calculated for the SVM classifier which is trained using the 15 original features (ORG + SVM)

Table 3 Comparisons of the performances of the one-against-one (OAO) and the one-against-all (OAA) SVM classifiers in terms of the average correct classification results obtained for each class

		Classification (OAO/OAA)										
		NSR		PVC		AF		SSS		VF		BII
Database annotation	NSR	277.3	277.90	0	0	0.70	0	2	2.10	0	0	0	0
	PVC	0	0	18	18	1	1	0	0	0	0	0	0
	AF	1.90	0	4.20	3.80	100.80	102.20	1.10	2	0	0	0	0
	SSS	4.40	2.80	0	0	0	0	15.6	17.20	0	0	0	0
	VF	0	0	0	0	0	0	0	0	26	26	0	0
	BII	0	0	0	0	0	0	0	0	0	0	10	10

Table 4 Comparisons of the performances of the one-against-one (OAO) and the one-against-all (OAA) SVM classifiers on the reduced feature space by the GDA in terms of the average values of the four commonly used measures in % (and their respective standard deviations)

Method	Sensitivity (%)	Specificity (%)	Positive predictivity (%)	Accuracy (%)
GDA + OAO SVM	94.18 (0.42)	99.07 (0.06)	93.44 (0.26)	98.90 (0.06)
GDA + OAA SVM	95.77 (0.39)	99.40 (0.13)	93.56 (0.63)	99.16 (0.11)

instead of the 5 reduced ones. Table 2 shows the resulting average values for the performance measuring parameters together with their standard deviations for both GDA + SVM and ORG + SVM algorithms. As seen, the proposed method can discriminate the NSR with an average accuracy of 98.94%, the PVC with 98.96%, the AF with 98.53%, the SSS with 98.51%, the VF with 100%, and the BII with 100%. These results demonstrate the effectiveness of the proposed arrhythmia classification algorithm in discriminating the six different types of arrhythmia. It must be noted that due to the different number of data segments available for each class in the data set, the performance measuring parameters are calculated for each class separately and the average values of these per class measures are used as the overall average classification rates for each method.

As mentioned earlier, a comparative study is carried out on the performances of the OAO and OAA decomposition methods in training the proposed multiclass SVMs. The obtained results are presented in Tables 3 and 4. As seen from the tables, the OAA method is superior to the OAO method in terms of the average correct classification results obtained for each class (Table 3) and in terms of the average correct classification rates represented in % in Table 4. Therefore, the OAA decomposition method is chosen to train the multiclass SVM in this work.

4. Discussion

This section presents the comparative discussions over the performances of the feature reduction techniques, classification techniques, and the whole arrhythmia classification procedures.

4.1. Comparing the feature reduction techniques

As one of the most commonly used dimension reduction techniques, PCA finds the most representative set of projection vectors such that the projected samples retain the most information about the original data samples. LDA, on the other hand, uses the class information and introduces a set of vectors that maximize the between-class scatter while minimizing the within-class scatter. Lastly GDA, similar to LDA maximizes the class separation however this takes place within a different feature space. According to the results presented in Table 5, application of the PCA to the data set prior to the MLP and SVM classifiers has not improved the classification performances of these classifiers. Table 5 also shows that although the combined LDA/MLP and LDA/SVM methods demonstrate better classification performances compared to the PCA based techniques; however the results are not satisfactory yet.

It should be noted that both PCA and LDA techniques are effective when linear projections can better describe the data structure. The arrhythmia classes are however not linearly separable and the HRV patterns demonstrate nonlinearity. Due to its ability in dealing with nonlinear problems, GDA performs better on the HRV data and can improve the classification performances significantly as this can be seen from Table 1. For example, misclassification of the HRV segments that belong to the SSS

Table 5 Comparison of the performances of different classifiers in terms of the average values of the four commonly used measures

Method	Sensitivity (%)	Specificity (%)	Positive predictivity (%)	Accuracy (%)
MLP	90.64	98.51	87.60	98.22
SVM	92.57	98.88	90.21	98.49
PCA + MLP	83.35	97.80	80.32	96.93
PCA + SVM	86.19	98.05	88.95	97.65
LDA + MLP	90.46	98.70	87.44	98.10
LDA + SVM	88.99	98.46	87.03	98.06
GDA + MLP	92.63	98.98	90.00	98.49
GDA + SVM	95.77	99.40	93.56	99.16

Table 6 The classification results produced by the MLP classifier and the proposed algorithm versus the gold standards defined by the experts

		Classification (MLP/GDA+SVM)										
		NSR		PVC		AF		SSS		VF		BII	
Database annotation	NSR	272.5	277.90	0	0	0	0	7.5	2.1	0	0	0	0
	PVC	0	0	16.67	18	2.33	1	0	0	0	0	0	0
	AF	0	0	4.50	3.80	100	102.20	3.33	2	0	0	0.17	0
	SSS	5.83	2.80	0	0	0.50	0	13.50	17.20	0	0	0.17	0
	VF	0	0	0	0	0.17	0	0.17	0	25.66	26	0	0
	BII	0	0	0	0	0	0	0	0	0	0	10	10

Table 7 A summary of different arrhythmia classification algorithms together with their reported results in terms of the four commonly used measures of sensitivity, specificity, positive predictivity, and accuracy

Author	Methods	Dataset	Arrhythmia types	Sensitivity (%)	Specificity (%)	Positive predictivity (%)	Accuracy (%)
Clayton et al. [2]	Threshold crossing intervals	**ECG** (Segment-based classification)	2 classes VF detection	46	72		
	Autocorrelation function			67	38		
	VF filter leakage			77	55		
	Signal spectrum shape			53	93		
Clayton et al. [3]	Neural networks	**ECG** (Segment-based classification)	2 classes VF detection	84	59		
Yang et al. [4]	Neural networks	**ECG** (Segment-based classification)	2 classes AF detection	92	92.3		
S. Kara et al. [10]	Neural networks + Wavelet features	**ECG** (Segment-based classification)	2 classes AF detection	100	100	100	100
Omer T.Inan et al. [9]	Neural networks + Wavelet transform & Timing interval features	**ECG** (Beat classification)	3 classes NSR, PVC, Others				95.16
Tsipouras et al. [8]	Neural networks + Time features	**HRV** (Segment-based classification)	All types included in MIT database (Normal, Arrhythmic)	87.53	89.48		
Tsipouras et al. [8]	Neural networks + Time-frequency features	**HRV** (Segment-based classification)	All types included in MIT database (Normal, Arrhythmic)	89.95	92.91		
Tsipouras et al. [16]	Knowledge-based system	**HRV** (Segment-based classification)	4 classes NSR, PVC, VF, BII				94.26
This work	SVM + Linear & nonlinear analysis	**HRV** (Segment-based classification)	6 classes NSR, PVC, AF, VF, SSS, BII	92.57	98.88	90.21	98.49
	SVM + Feature reduction by GDA			95.77	99.40	93.56	99.16

class to the NSR class is reduced from 7 segments to only 2.8 segments (in average) by applying the GDA technique to the original features prior to classification.

According to Table 2, in general, the resulting sensitivity, specificity, positive predictivity, and accuracy are improved from 92.57%, 98.88%, 90.21% and 98.49% to 95.77%, 99.40%, 93.56%, and 99.16%, respectively, by applying GDA to the original features prior to classification. Therefore, the overall performance of the combined GDA/SVM classifier is better than that of the SVM classifier applied to the original features by more than 3% in sensitivity, 0.5% in specificity, 3.35% in positive predictivity, and 0.65% in accuracy. Furthermore, the necessary time for training the SVM classifiers is significantly reduced when using the reduced number of the input features compared to when the original features are used.

4.2. Comparing performances of the classification techniques

In this section the performance of the SVM classifier, which is adapted in this work, is compared to that of the MLP classifier, which has been widely used in the past for ECG pattern analysis and arrhythmia classification [3—10]. A three layer MLP with 15 inputs (5 inputs for the case of the reduced feature vectors), one hidden layer with 20 neurons and 6 outputs for the six arrhythmia classes each with a real value in the interval [0, 1] was developed. The MLP was trained using the training data set and employing the backpropagation strategy. For each input feature vector, the output with the largest value indicates the appropriate class that the input vector belongs to. The resulting average classification rates are summarized in Tables 5 and 6.

Table 6 indicates that the proposed (SVM + GDA) method outperforms the MLP classifier. In particular, in comparison to the MLP classifier the proposed algorithm could better discriminate the arrhythmia classes of NSR, PVC, AF, and SSS by an average value of 5.40, 1.33, 2.20, and 3.7 segments, respectively. Moreover, Table 5 shows that the sensitivity and positive predictivity of the proposed classifier are 5.13% and 5.94% higher compared to the MLP classifier, respectively. This demonstrates the effectiveness of the proposed GDA + SVM classifier in the arrhythmia classification applications.

4.3. Comparison of the proposed approach with other methods in the literature

Several researchers have addressed the arrhythmia detection and classification problem using the ECG signals directly or by analyzing the heart rate variability signal in the past. A summary of different methods together with their reported results in terms of the four commonly used measures of sensitivity, specificity, positive predictivity, and accuracy is summarized in Table 7. Most papers [2—4,10] have focused on the detection of a single arrhythmia type (mostly the VF and AF) within normal sinus rhythms. In [10] the authors have reported a classification rate of 100% for the AF within their small data set. The authors of [8] have classified the ECG signal segments into the normal or arrhythmic classes via neural networks. In another attempt different heart rhythms were detected and classified into the two or three arrhythmia types using a knowledge-based system [16]. In a recent paper [9] another neural network-based algorithm using the wavelet transform was developed for discriminating the NSR and PVC beats.

Comparing to these papers, an effective HRV-based algorithm is proposed in the current work which provides a better accuracy over a wider range of different types of cardiac arrhythmia (six different classes). Another important advantage of the proposed algorithm is its ability in effectively discriminating the SSS arrhythmia from the NSR rhythm which is a difficult task and none of the previously reported methods could perform such a task.

5. Conclusions

In this paper, an effective HRV-based cardiac arrhythmia classification algorithm was presented. Initially, 15 original features were extracted from the input HRV signals including 8 linear features (7 time domain features and 1 frequency domain feature) and 7 nonlinear features. These features were used for discriminating six different types of cardiac arrhythmia by means of the SVM classifier. In order to reduce the learning time and also to improve the learning efficiency of the classifier, the 15 original features were reduced to only 5 new features by means of the GDA algorithm. The new features were also used for classification of the six arrhythmia classes by the SVM scheme. Comparing the classification results which was carried out in Tables 1 and 2, it was shown that the proposed hybrid GDA + SVM cardiac arrhythmia classification algorithm outperforms the SVM classifier applied to the original features producing the discrimination accuracy of 98.94%, 98.96%, 98.53%, 98.51%, 100%, and 100%, for the arrhythmia classes of NSR, PVC, AF, SSS, VF, and BII, respectively.

Moreover, comparing the performance of the proposed algorithm to those of the previously

reported methods in the literature in Table 7, it was shown that the proposed algorithm is more effective than any of those methods. One of the important advantages of the proposed algorithm when compared to the ECG-based approaches in the literature is that it is completely based on the HRV signal which can be extracted from the initial ECG signal with a high accuracy even for noisy and/or complicated recordings. This is while, most ECG-based methods use the morphological features of the ECG, which is seriously affected by noise. In addition, the exclusive use of the R—R interval duration signal effectively reduces the processing time, compared to the direct ECG-based methods.

On the other hand, one drawback of the proposed HRV-based algorithm is that some of the arrhythmia types such as the left bundle branch block and the right bundle branch block beats cannot be detected using only the heart rate variability features.

As a last point, due to the short processing time and relatively high accuracy of the proposed method, it can be used as a real-time arrhythmia classification system.

References

[1] Task force of the European society of cardiology and the North American society of pacing and electrophysiology. Heart rate variability—standards of measurements, physiological interpretation, and clinical use. Eur Heart J 1996;17(3):354—81.

[2] Clayton RH, Murray A, Campbell RWF. Comparison of four techniques for recognition of ventricular fibrillation from the surface ECG. Med Biol Eng Comput 1993;31(2):111—7.

[3] Clayton RH, Murray A, Campbell RWF. Recognition of ventricular fibrillation using neural networks. Med Biol Eng Comput 1994;32(2):217—20.

[4] Yang TF, Devine B, Macfarlane PW. Artificial neural networks for the diagnosis of atrial fibrillation. Med Biol Eng Comput 1994;32(6):615—9.

[5] Minami K, Nakajima H, Toyoshima T. Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network. IEEE Trans Biomed Eng 1999;46(2):179—85.

[6] Wang Y, Zhu YS, Thakor NV, Xu YH. A short-time multifractal approach for arrhythmia detection based on fuzzy neural network. IEEE Trans Biomed Eng 2001;48(9):989—95.

[7] Acharya RU, Kumar A, Bhat PS, Lim CM, Iyengar SS, Kannathal N, et al. Classification of cardiac abnormalities using heart rate signals. Med Biol Eng Comput 2004;42(3):288—93.

[8] Tsipouras MG, Fotiadis DI. Automatic arrhythmia detection based on time and time—frequency analysis of heart rate variability. Comp Meth Prog Biomed 2004;74(2):95—108.

[9] Inan OT, Giovangrandi L, Kovacs GTA. Robust neural-network-based classification of premature ventricular contractions using wavelet transform and timing interval features. IEEE Trans Biomed Eng 2006;53(12):2507—15.

[10] Kara S, Okandan M. Atrial fibrillation classification with artificial neural networks. Pattern Recogn 2007;40(11):2967—73.

[11] Khadra L, Al-Fahoum AS, Al-Nashash H. Detection of life-threatening cardiac arrhythmias using wavelet transformation. Med Biol Eng Comput 1997;35(6):626—32.

[12] Al-Fahoum AS, Howitt I. Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias. Med Biol Eng Comput 1999;37(1):566—73.

[13] Song MH, Lee J, Cho SP, Lee KJ, Yoo SK. Support vector machine based arrhythmia classification using reduced features. Int J Control Automat Syst 2005;3(4):571—9.

[14] Muirhead RJ, Puff RD. A Bayesian classification of heart rate variability data. Physica A 2004;336:503—13.

[15] Tsipouras MG, Goletsis Y, Fotiadis DI. A method for arrhythmic episode classification in ECGs using fuzzy logic and Markov models. In: Murray A, editor. Proceedings of the computers in cardiology. 2004. p. 361—4.

[16] Acharya RU, Bhat BS, Iyengar SS, Rao A, Dua S. Classification of heart rate data using artificial neural network and fuzzy equivalence relation. Pattern Recogn 2003;36(1):61—8.

[17] Tsipouras MG, Fotiadis DI, Sideris D. An arrhythmia classification system based on the RR-interval signal. Artif Intell Med 2005;33(3):237—50.

[18] Baudat G, Anouar F. Generalized discriminant analysis using a kernel approach. Neural Comput 2000;12(10):2385—404.

[19] Liu QS, Lu HQ, Ma SD. Improving kernel Fisher discriminant analysis for face recognition. IEEE Trans Circuits Syst Video Technol 2004;14(1):42—9.

[20] Yang MH. Kernel eigenfaces vs. kernel Fisherfaces: face recognition using kernel methods. In: Proceedings of the 5th international conference on automation of face gesture recognition; 2002. p. 215—20.

[21] Lu J, Plataniotis KN, Venetsanopoulos AN. Face recognition using kernel direct discriminant analysis algorithms. IEEE Trans Neural Networks 2003;14(1):117—26.

[22] Liu YH, Huang HP, Weng CH. Recognition of electromyographic signals using cascaded kernel learning machine. IEEE/ASME Trans Mechatron 2007;12(3):253—64.

[23] Polat K, Güneş S, Arslan A. A cascade learning system for classification of diabetes disease: generalized discriminant analysis and least square support vector machine. Expert Syst Appl 2008;34(1):482—7.

[24] Polat K, Güneş S. A novel approach to estimation of E. coli promoter gene sequences: combining feature selection and least square support vector machine (FS_LSSVM). Appl Math Comput 2007;190(2):1574—82.

[25] Vapnik V. Statistical learning theory. Berlin: Springer; 1998.

[26] Osuna E, Freund R, Girosi F. Training support vector machines: an application to face detection. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition; 1997. p. 130—6.

[27] Liu YH, Chen YT. Face recognition using total margin-based adaptive fuzzy support vector machines. IEEE Trans Neural Networks 2007;18(1):178—92.

[28] Theodoridis S, Koutroumbas K. Pattern recognition, 2nd ed., San Diego: Elsevier Academic Press; 2003.

[29] Burges JC. A tutorial on support vector machines for pattern recognition. Data Min Knowl Discov 1998;2:121—67.

[30] Mark RG, Moody GB. MIT-BIH Arrhythmia Database 1997 [Online]. Available from: http://ecg.mit.edu/dbinfo.html [accessed: 01.04.2008].

[31] Moody GB, Mark RG. The impact of the MIT/BIH arrhythmia database. IEEE Eng Med Biol Mag 2001;20(3):45—50.

[32] Badilini F, Moss AJ, Titlebaum EL. Cubic spline baseline estimation in ambulatory ECG recordings for the measurement of ST segment displacements. In: Proceedings of the annual international conference of the IEEE Engineering in Medicine and Biology Society 13(2); 1991. p. 584—5.

[33] Pan J, Tompkins WJ. A real time QRS detection algorithm. IEEE Trans Biomed Eng 1985;32(3):230—6.

[34] Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biomed Eng 1986;33(12):1157—65.

[35] Acharya RU, Kannathal N, Krishnan SM. Comprehensive analysis of cardiac health using heart rate signals. Physiol Meas 2004;25(5):1139—51.

[36] Carvalho JLA, Rocha AF, Nascimento FA, Neto JS, Junqueira LF. Development of a Matlab software for analysis of heart rate variability. In: Proceedings of the 6th international conference on signal processing; 2002. p. 1488—91.

[37] Tulppo MP, Makikallio TH, Takala TES, Seppanen T, Huikuri HV. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. Am J Physiol 1996;271(1):244—52.

[38] Ho KLL, Moody GB, Peng CK, Mietus JE, Larson MG, Levy D, et al. Predicting survival in heart failure case and control subjects by use of fully automated methods for deriving nonlinear and conventional indices of heart rate dynamics. Circulation 1997;96(3):842—8.

[39] Pincus SM. Approximate entropy as a measure of system complexity. Proc Natl Acad Sci USA 1991;88(6):2297—301.

[40] Pincus SM, Goldberger AL. Physiological time series analysis: what does regularity quantify? Am J Physiol 1994;266: 1643—56.

[41] Rezek IA, Roberts SJ. Stochastic complexity measures for physiological signal analysis. IEEE Trans Biomed Eng 1998; 45(9):1186—91.

[42] Eckman JP, Kamphorst SO, Ruelle D, Ciliberto S. Lyapunov exponents from time series. Phys Rev A 1986;34(6):4971—9.

[43] Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge, UK: Cambridge University Press; 1997.

[44] Wolf A, Swift JB, Swinney HL, Vastano JA. Determining Lyapunov exponents from a time series. Physica D 1985; 16:285—317.

[45] Uzun IS, Asyali MH, Selebi G, Pehlivan M. Nonlinear analysis of heart rate variability. In: Proceedings of the 23rd annual international conference of the IEEE Engineering in Medicine and Biology Society, vol. 2; 2001. p. 1581—4.

[46] Huikuri HV, Makikallio TH, Peng CK, Goldberger AL, Hintze U, Moller M. Fractal correlation properties of R—R interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infraction. Circulation 2000;101:47—53.

[47] Kim HC, Kim DJ, Bang SY. Face recognition using LDA mixture model. In: Proceedings of the 16th international conference on pattern recognition, vol. 2; 2002. p. 486—9.

[48] Martinez AM, Kak AC. PCA versus LDA. IEEE Trans Pattern Anal Mach Intell 2001;23(2):228—33.

[49] Mercier G, Lennon M. Support vector machines for hyperspectral image classification with spectral-based kernels. In: Proceedings of the international geoscience and remote sensing symposium; 2003. p. 288—90.

[50] Wang L, Liu B, Wan C. Classification using support vector machines with graded resolution. In: Proceedings of the international conference on granular computing, vol. 2; 2005. p. 666—70.

[51] Hsu C-W, Lin C-J. A comparison of methods for multiclass support vector machines. IEEE Trans Neural Networks 2002; 13(2):415—25.