Detection of ECG Arrhythmia Conditions using CSVM and MSVM Classifiers

N. Jannah and S. Hadjiloucas
School of Systems Engineering, The University of Reading
n.m.h.jannah@pgr.reading.ac.uk, s.hadjiloucas@reading.ac.uk

Electrocardiogram (ECG) is widely used for the diagnosis of cardiac arrhythmia conditions. An automatic classification of four beat types Normal (N), premature ventricular contraction (PVC), Supraventricular premature or ectopic beat (SVPB) and Fusion of ventricular and normal beat (FUSION) is implemented using a Multi-class Support Vector Machine (MSVM) and Complex Support Vector Machine (CSVM) algorithms [1]. The ECG signals used in these studies were obtained from the European ST-T Database. A number of beats from different leads and patients were selected for training and evaluating classifier performance. Successful ECG arrhythmia classification usually requires optimizing the following procedures: Pre-processing and beat detection, feature extraction and selection, and classifier optimization. Pre-processing and R peak detection is performed with the WFDB Software Package. This reads the annotation and finds the R (peak) location. R (peak) location used as a reference to detect peaks in other wave such P and T and extract ECG beat. ECG beats are extracted after windowing the signal using 106 samples before the R and 106 samples after the R-peak. Discrete Cosine and Sine transforms or the Discrete Fourier Transform (DFT) were used for feature extraction and dimensionality reduction of the input vector at the input of the classifier. Studies after selecting either 100 or 50 Fourier coefficients for reconstructing individual ECG beats in the feature selection phase were performed. MATLAB software routines were used to train and validate both the CSVM and the Multi-class Support Vector Machine (MSVM) classifier. A Complex kernel function, (Gaussian RBK) with 5-fold cross validation was used for adjusting the kernel values. Sequential minimal optimization (SMO) [2]was used to train the CSVM and compute the corresponding complex hyper-plane parameters. The aim of the study was to improve multi-class SVM by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. Results illustrate that the proposed beat classifier is very reliable, and that it may be adopted for automatic detection of arrhythmia conditions and classification. Accuracies between 86% and 94% are obtained for MSVM and CSVM classification respectively. Using CSVM, a 4 classes problem can be classified rapidly by decomposing it into two distinct SVM tasks. Moreover, the present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier. Future work on wavelet pre-processing to further compress the input space of the classifier by generating wavelets on the basis of higher order moment criteria [3] as well as alternative approaches for extending the CSVM input and output spaces to arbitrary dimension using Clifford algebra SVM [4] will be discussed at the conference.

Feature extraction

Discrete Fourier Transform (DFT)

$$X(k)=\sum_{n=0}^{N-1}x(n)W_N^{nk}\quad ,k=0,\ldots.N-1$$

Inverse DFT

$$x(n)=\frac{1}{N}\sum_{k=0}^{N-1}X(k)W_N^{-nk}\quad ,k=0,\ldots.N-1$$

Discrete Cosine and Sine transforms (DCT and DST)

$$C(k)=\sum_{n=0}^{N-1}x(n)\cos\left(\frac{\pi k(2n+1)}{2N}\right),k=0,1,\ldots.N-1$$

$$S(k)=\sum_{n=0}^{N-1}x(n)\sin\left(\frac{\pi(n+1)(k+1)}{N+1}\right),k=0,1,\ldots.N-1$$

Classification Algorithm

SVM

$$f(x)=\sum_{i=1}^{N}\alpha_i\ y_i\ \mathrm{K}(x_i,x_j)+b$$

Lagrangian function (Dual problem)

$$L=\max\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i,x_j)$$

CSVM

$$g(z)=\underset{i}{\mathrm{sign}}\left(2\sum_{n=1}^{N}\left(a_n d_n^r+ib_n d_n^i\right)\kappa_{\mathbb{C}}^r(z_n,z)+c^r+ic^i\right)$$

Two hyperplanes Algorithm real and imaginary respectively:

$$\underset{a}{\text{maximize}}\ \sum_{n=1}^{N}a_n-\sum_{n,m=1}^{N}a_n a_m\, d_n^r d_m^r\kappa_{\mathbb{C}}^r(z_m,z_n)$$

Wavelet filter bank parametrization algorithm

It proves convenient to write the wavelet transform in matrix form as:

$$\mathbf{t}_{1\times J}=\mathbf{x}_{1\times J}\mathbf{V}_{J\times J}$$

where x = [x_0 x_1 ... x_{J-1}] is the row vector of original variables, t is the row vector of new (transformed) variables and V is the matrix of weights. Let {$h_0, h_1, \ldots, h_{2N-1}$} and {$g_0, g_1, \ldots, g_{2N-1}$} be the impulse responses of the low-pass and high-pass filters respectively.

Choosing V to be unitary

$$\mathbf{V}=\left[\begin{array}{ccccc|ccccc}
0 & 0 & \cdots & h_{2N-4} & h_{2N-2} & 0 & 0 & \cdots & g_{2N-4} & g_{2N-2}\\
h_{2N-1} & 0 & \cdots & h_{2N-5} & h_{2N-3} & g_{2N-1} & 0 & \cdots & g_{2N-5} & g_{2N-3}\\
h_{2N-2} & 0 & \cdots & h_{2N-6} & h_{2N-4} & g_{2N-2} & 0 & \cdots & g_{2N-6} & g_{2N-4}\\
h_{2N-3} & h_{2N-1} & \cdots & h_{2N-7} & h_{2N-5} & g_{2N-3} & g_{2N-1} & \cdots & g_{2N-7} & g_{2N-5}\\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
h_0 & h_2 & \cdots & 0 & 0 & g_0 & g_2 & \cdots & 0 & 0\\
0 & h_1 & \cdots & 0 & 0 & 0 & g_1 & \cdots & 0 & 0\\
0 & h_0 & \cdots & 0 & 0 & 0 & g_0 & \cdots & 0 & 0\\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & h_{2N-2} & 0 & 0 & 0 & \cdots & g_{2N-2} & 0\\
0 & 0 & \cdots & h_{2N-3} & h_{2N-1} & 0 & 0 & \cdots & g_{2N-3} & g_{2N-1}
\end{array}\right]$$

The approximation c and detail d coefficients are stacked in vector t = [c (s)| d(s)], with coefficients s_i in larger scale associated with broad features in the data vector, and coefficients in smaller scales associated sharp peaks.

$$t_j=\sum_{n=0}^{J-1}x_n v_j(n),\quad j=0,1,\ldots,J-1$$

The following conditions ensure orthogonality of the transform:

$$\sum_{n=0}^{2N-1-2l}h_n h_{n+2l}=\begin{cases}1, & l=0\\ 0, & 0<l<N\end{cases}$$

$$g_n=(-1)^{n+1}h_{2N-1-n},\ n=0,1,\ldots,2N-1$$

Parametrization and optimization approach algorithm

$$h_0^{(1)}=\cos(a_1),$$
$$h_0^{(N+1)}=\cos(a_{N+1})h_0^{(N)}$$
$$h_{2i}^{(N+1)}=\cos(a_{N+1})h_{2i}^{(N)}-\cos(a_{N+1})h_{2i-1}^{(N)},\quad i=1,2\ldots N-1$$
$$h_{2N}^{(N+1)}=-\sin(a_{N+1})h_{2N-1}^{(N)}$$
$$h_1^{(1)}=\sin(a_1)$$
$$h_1^{(N+1)}=\sin(a_{N+1})h_0^{(N)}$$
$$h_{2i+1}^{(N+1)}=\sin(a_{N+1})h_{2i}^{(N)}+\cos(a_{N+1})h_{2i-1}^{(N)},\quad i=1,2\ldots N-1$$
$$h_{2N+1}^{(N+1)}=\cos(a_{N+1})h_{2N-1}^{(N)}$$

In order to ensure two vanishing moments for the resulting transform

$$a_N=\frac{\pi}{4}-\sum_{i=1}^{N-1}a_i$$

$$a_{N-1}=\frac{1}{2}\arcsin\left\{-\frac{1}{2}-\sum_{k=1}^{N-2}\left[\sin\sum_{i=1}^{k}2a_i\right]\right\}-\sum_{i=1}^{N-2}a_i$$

The last expression has a real value solution if the set of angles a_i where $1\le i\le N-2$ satisfy a set of constrains that define a non-convex region

$$\chi_i=\sin\sum_{k=1}^{i}2a_k,\quad 1\le i\le N-2$$

$$-\frac{3}{2}\le\sum_{i=1}^{N-2}\chi_i\le\frac{1}{2},\quad -1\le\chi_1,\chi_2\cdots\chi_{N-2}\le 1$$

$$a_1=\frac{1}{2}\arcsin(\chi_1)$$

$$a_i=\frac{1}{2}\arcsin(\chi_i)-\sum_{k=1}^{i-1}2a_k,\quad 2\le i\le N-2$$

ECG beat number	Classification methodology	ECG coefficient number used	Classification accuracy
622	MSVM	100	86
622	CSVM	100	94

References

[1] P. Bouboulis, S. Theodoridis, C. Mavroforakis, and L. Evaggelatou-Dalla, "Complex Support Vector Machines for Regression and Quaternary Classification," *IEEE Trans. Neural Networks Learn IEEE Transactions on Neural Networks and Learning Systems. Syst.*, vol.26, no.6, pp. 1–15, 2014.

[2] J. C. Platt, "Sequential minimal optimization: A fast algorithm for training support vector machines," *Adv. Kernel MethodsSupport Vector Learn.*, vol. 208, pp. 1–21, 1998.

[3] S. Hadjiloucas, N. Jannah, F. Hwang, and R. K. H. Galvão, "On the application of optimal wavelet filter banks for ECG signal classification," *J. Phys. Conf. Ser.*, vol. 490, p. 012142, 2014. ibliography

[4] E. J. Bayro-Corrochano and N. Arana-Daniel, "Clifford support vector machines for classification, regression, and recurrence," *IEEE Trans. Neural Networks*, vol. 21, no. 11, pp. 1731–1746, 2010.

Ministry of Higher Education

Detection of ECG Arrhythmia Conditions using CSVM and MSVM Classifiers

Najlaa Jannah and Dr. Sillas Hadjiloucas

School of Systems Engineering, The University of Reading

School of Systems Engineering

Abstract

- An automatic classification of four beat types Normal (N), premature ventricular contraction (PVC), Supraventricular premature or ectopic beat (SVPB) and Fusion of ventricular and normal beat (FUSION) is implemented using a Multi-class Support Vector Machine (MSVM) and Complex Support Vector Machine (CSVM) algorithm.
- Pre-processing and R peak detection is performed with the WFDB Software Package which reads the annotation and finds the R (peak) location. R (peak) location is used as a reference to detect peaks in other wave such P and T and extract ECG beat.
- Discrete Cosine and Sine transforms or Discrete Fourier Transform (DFT) were used for feature extraction and dimensionality reduction of the input vector at the input of the classifier.
- The aim of the study was to improve multi-class SVM by extending traditional SVM algorithms to complex spaces so as to simultaneously classify four types of heartbeats. The approach can also account for phase changes in the signal.
- Accuracy between 86% and 94% are obtained for MSVM and CSVM classification respectively.
- Results illustrate that the proposed beat classifier is very reliable, and that it may be a useful tool for detection arrhythmia conditions and perform classification automatically.

Introduction

- ECGs provide a graphic representation of the electrical activity of the heart muscle.
- This study evaluates ECG classification using pre-processing routines, feature extraction and selection, and ECG beat classification using MSVM and CSVM.
- DCT and DST coefficients were used to extract features presented at the input vector of the MSVM classifier whereas DFT coefficients are used for creating an input vector to the CSVM classifier.
- The MSVM and CSVM classifier is used to distinguish the four ECG arrhythmias types using the DCT, DST and DFT coefficients as feature vectors at the input vector of the classifier.
- Sequential minimal optimization (SMO) approach is used to train the CSVM and compute the corresponding complex hyper-plane parameters.

Support Vector Machine (SVM) algorithm

- Lagrangian function evaluation based on the dual problem:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i K(x_i, x_j) + b$$

$$L = \max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

Complex Support Vector Machine (CSVM) algorithm

$$g(z) = \operatorname{sign}_i \left(2 \sum_{n=1}^{N} \left(a_n d_n^r + i b_n d_n^i \right) \kappa_{\mathbb{C}}^r (z_n, z) + c^r + i c^i \right)$$

- Two hyperplanes are now defined instead of one defined using standard SVM, real and imaginary respectively:

$$\underset{a}{\text{maximize}} \sum_{n=1}^{N} a_n - \sum_{n,m=1}^{N} a_n a_m d_n^r d_m^r \kappa_{\mathbb{C}}^r (z_m, z_n)$$

$$\underset{b}{\text{maximize}} \sum_{n=1}^{N} b_n - \sum_{n,m=1}^{N} b_n b_m d_n^i d_m^i \kappa_{\mathbb{C}}^r (z_m, z_n)$$

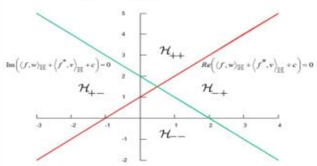

Figure (1) complex hyper- plane in CSVM

Method

Diagrams below summarises the method adopted for ECG classification.

Raw ECG signal

Pre-processing

Baseline removal | ECG Beat extracted

Feature Extraction in Fourier space

DFT | DCT | DST

Classification result for each beat type

N | PVC | SVPB | FUSION

Figure (2) Structure of methodology used for ECG classification.

Feature Extraction Algorithm

- Reconstruction after deleting coefficient below threshold.
  - Discrete Fourier Transform (DFT) and Inverse DFT

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad, k = 0, \ldots ..N-1$$

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-nk} \quad, k = 0, \ldots ..N-1$$

  - Discrete Cosine and Sine transforms

$$C(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{\pi k(2n+1)}{2N}\right), k = 0, 1, \ldots ..N-1$$

$$S(k) = \sum_{n=0}^{N-1} x(n) \sin\left(\frac{\pi(n+1)(k+1)}{N+1}\right), k = 0, 1, \ldots ..N-1$$

- Using DCT coefficients for creating an input vector to MSVM provides better classification than DST.

Experimental Design and Results

- Signal processing was carried out using MATLAB and the Wave Form Database (WFDB) software packages.
- The ECG beats were obtain from the European ST-T Database with 212 samples around the R-peak.
- In feature selection 100 Fourier coefficients were selected for reconstructing individual ECG beats, these form the input vector to the classifier shown in figure (3).
- For the classifier implementation, after feature selection is performed, the datasets are divided into two groups with 311 beats for training and testing purposes.
- The optimal hyper-plane parameters are identified on the basis of the training data set, class label. The output of the SMO algorithm was used for training.
- Finally, on the basis of the training results, test data are imported to the CSVM and MSVM classifiers to perform unknown beat classification.
- The proposed algorithm confirmed classification accuracies of 94%, whereas multi-class SVM 86% accuracies.
- Table (3) illustrates the performance of the classification process using three common measures- sensitivity (ST), specificity(SP), and predictively (PP) .

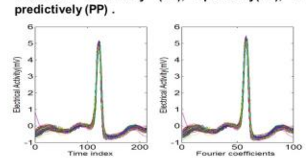

Figure (3) Comparison between original beat feature and reconstructed beats from 100 Fourier coefficients

Annotation	Output result			
	NORMAL	PVC	SVPB	FUSION
NORMAL	150	0	0	3
PVC	0	53	0	3
SVPB	0	3	60	5
FUSION	0	4	0	30

Table (1) confusion matrix using CSVM

Annotation	Output result			
	NORMAL	PVC	SVPB	FUSION
NORMAL	149	0	0	0
PVC	1	59	6	33
SVPB	0	0	54	1
FUSION	0	1	0	7

Table (2) confusion matrix using MSVM

Annotation	Classification performance result					
	CSVM			MSVM		
	PP (%)	ST (%)	SP (%)	PP (%)	ST (%)	SP (%)
NORMAL	98	100	98	100	99	100
PVC	95	88	99	60	98	84
SVPB	88	100	97	98	90	100
FUSION	88	73	99	88	17	100

Table (3) Collective result performance analysis and classification result using MSVM and CSVM

Future Work

- Multidimensional SVM using Clifford Algebras will be used to account for multi-lead signal analysis.
- The proposed approach is to be extended using adaptive wavelet where a different wavelet function is derived at each decomposing level to increase parsimony of the input vector.
- For feature selection wavelet pre-processing will be used to further compress the input to the classifier by generating wavelets on the basis of higher order moment criteria.
- It proves convenient to write the transform in matrix form as: $\mathbf{t}_{1\times J} = \mathbf{x}_{1\times J} \mathbf{V}_{J\times J}$

  where $\mathbf{X} = [x_0 \; x_1 \; \ldots \; x_{J-1}]$ is the row vector of original variables, t is the row vector of new (transformed) variables and $\mathbf{V}$ is the matrix of weights.
- Let $\{h_0, h_1, \ldots, h_{2N-1}\}$ and $\{g_0, g_1, \ldots, g_{2N-1}\}$ be the impulse responses of the low-pass and high-pass filters respectively.
- Choosing $\mathbf{V}$ to be unitary:

$$\mathbf{V} = \left[\begin{array}{ccccc|ccccc}
0 & 0 & \cdots & h_{2N-4} & h_{2N-2} & 0 & 0 & \cdots & g_{2N-4} & g_{2N-2} \\
h_{2N-1} & 0 & \cdots & h_{2N-5} & h_{2N-3} & g_{2N-1} & 0 & \cdots & g_{2N-5} & g_{2N-3} \\
h_{2N-2} & 0 & \cdots & h_{2N-6} & h_{2N-4} & g_{2N-2} & 0 & \cdots & g_{2N-6} & g_{2N-4} \\
h_{2N-3} & h_{2N-1} & \cdots & h_{2N-7} & h_{2N-5} & g_{2N-3} & g_{2N-1} & \cdots & g_{2N-7} & g_{2N-5} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
h_0 & h_2 & \cdots & 0 & 0 & g_0 & g_2 & \cdots & 0 & 0 \\
0 & h_1 & \cdots & 0 & 0 & 0 & g_1 & \cdots & 0 & 0 \\
0 & h_0 & \cdots & 0 & 0 & 0 & g_0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & h_{2N-2} & 0 & 0 & 0 & \cdots & g_{2N-2} & 0 \\
0 & 0 & \cdots & h_{2N-3} & h_{2N-1} & 0 & 0 & \cdots & g_{2N-3} & g_{2N-1}
\end{array}\right]$$

- The approximation $\mathbf{c}$ and detail $\mathbf{d}$ coefficients are stacked in vector $\mathbf{t} = [\mathbf{c}(s) | \mathbf{d}(s)]$, with coefficients s_i in larger scale associated with broad features in the data vector, and coefficients in smaller scales associated sharp peaks.

$$t_j = \sum_{n=0}^{J-1} x_n v_j(n), \quad j = 0, 1, \ldots, J-1$$

$$\sum_{n=0}^{2N-1-2l} h_n h_{n+2l} = \begin{cases} 1, l = 0 \\ 0, 0 < l < N \end{cases} \quad g_n = (-1)^{n+1} h_{2N-1-n}, n = 0, 1, \ldots, 2N-1$$

- Details of the optimization of the wavelet filter band parametrization are discussed in [1].

Summary

- Using CSVM, a 4 classes problem can be classified rapidly by decomposing it into two distinct SVM tasks.
- The present research confirmed that the use of selected number of Fourier coefficients to approximate the ECG beat signal and further compress the input features to the classifier can lead to high classification accuracies and improve the generalization ability of the CSVM classifier.
- Using the CSVM algorithm indicated a significant improvement in multi-class classification over MSVM.

Acknowledgements

- N Jannah acknowledges scholarship support from the Ministry of Higher Education, Kingdom of Saudi Arabia.

References

[1] S. Hadjiloucas, N. Jannah, F. Hwang, and R. K. H. Galvão, "On the application of optimal wavelet filter banks for ECG signal classification," *J. Phys. Conf. Ser.*, 490, 012142, 2014.