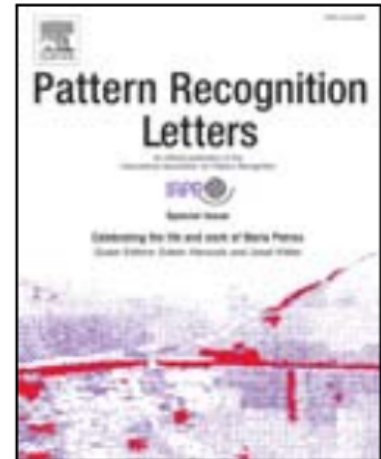


## Accepted Manuscript

Cardiac Arrhythmia Classification Using Statistical and Mixture Modeling Features of ECG Signals

Rashid Ghorbani Afkhami , Ghanbar Azarnia ,  
Mohammad Ali Tinati

PII: S0167-8655(15)00404-3  
DOI: [10.1016/j.patrec.2015.11.018](https://doi.org/10.1016/j.patrec.2015.11.018)  
Reference: PATREC 6395



To appear in: *Pattern Recognition Letters*

Received date: 18 March 2015  
Accepted date: 11 November 2015

Please cite this article as: Rashid Ghorbani Afkhami , Ghanbar Azarnia , Mohammad Ali Tinati , Cardiac Arrhythmia Classification Using Statistical and Mixture Modeling Features of ECG Signals, *Pattern Recognition Letters* (2015), doi: [10.1016/j.patrec.2015.11.018](https://doi.org/10.1016/j.patrec.2015.11.018)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Highlights

- Gaussian mixture modeling is used to fit the probability density function of heartbeats.
- Expectation maximization algorithm estimates the parameters of statistical model.
- Skewness, kurtosis and 5<sup>th</sup> moment of ECG signals express the shape parameters.
- RR interval information represent the time-domain characteristics of ECG signals.
- An ensemble of decision trees are used to perform the final classification.

ACCEPTED MANUSCRIPT

# Cardiac Arrhythmia Classification Using Statistical and Mixture Modeling Features of ECG Signals

Rashid Ghorbani Afkhami<sup>a</sup> \*, Ghanbar Azarnia<sup>a</sup> and Mohammad Ali Tinati<sup>a</sup>

<sup>a</sup>Faculty of Electrical Engineering, University of Tabriz, Tabriz, Iran

---

## ABSTRACT

In this paper we propose a novel method for accurate classification of cardiac arrhythmias. Morphological and statistical features of individual heartbeats are used to train a classifier. Two RR interval features as the exemplars of time-domain information are utilized in this study. Gaussian mixture modeling (GMM) with an enhanced expectation maximization (EM) solution is used to fit the probability density function of heartbeats. Parameters of GMM together with shape parameters such as skewness, kurtosis and 5th moment are also included in feature vector. These features are then used to train an ensemble of decision trees. MIT-BIH arrhythmia database containing various types of common arrhythmias is employed to test the algorithm. The overall accuracy of 99.70% in “class-oriented” scheme and 96.15% in “subject-oriented” scheme is achieved. Both cases express a significant improvement of accuracy in compare to other methods.

2012 Elsevier Ltd. All rights reserved.

---

## Keywords

Decision Tree, Expectation maximization, Gaussian mixture model, Heartbeat classification, Higher order statistics.

## 1. Introduction

The analysis of the electrocardiographic (ECG) signal provides a detailed information on the patient's cardiac health status. Cardiac arrhythmias are group of conditions in which the electrical activity of the heart is irregular, faster or slower than normal. While some types of arrhythmias impose an imminent threat on patient's life (e.g. ventricular fibrillation and tachycardia) other types represent long-term threats, which will still need special care. Cardiac arrhythmias are also one of the most common causes of death and as a result, their accurate detection has been of great interest in biomedical studies.

ECG is a simple and specific diagnostic test for assessment of heart rhythms. In most cases arrhythmias don't show any consistent changes in ECG signal, therefore a Holter monitor is needed. Holter monitor is an ECG recorded over a longer period time (i.e. longer than thirty minutes), which helps the detection of dysrhythmias that may happen briefly and unpredictably throughout that period. While analyzing such a long signal by physicians is extremely time demanding, signal processing techniques can offer rapid, real-time categorization with acceptable accuracies. Our goal in this paper is to introduce a legitimate method for automatic detection of cardiac arrhythmias based on novel techniques in signal processing.

Electrocardiographic based classification of cardiac arrhythmias has been investigated in numerous papers. These methods can usually differ in three main aspects, features, classifiers and evaluation schemes.

Different features used in papers are Hermit coefficients (Jiang & Kong, 2007), (Lagerholm, et al., 2002), (Osowski, et al., 2004), higher order statistical features (de Lannoy, et al., 2012), (Osowski, et al., 2004), morphological features (de Chazal, et al., 2004), (de Oliveira, et al., 2011), (Zeng, et al., 2011), independent component analysis and wavelet features (Ince, et al., 2009), (Jiang, et al., 2006), (Yang & Shen, 2013), (Ye, et al., 2012).

Classifiers such as self-organizing map (SOM) (Lagerholm, et al., 2002), support vector machine (SVM) (Jiang, et al., 2006), (Osowski, et al., 2004), (Ye, et al., 2012), artificial neural network (ANN) (Ince, et al., 2009), (Jiang & Kong, 2007), conditional random field (CRF) (de Lannoy, et al., 2012), linear discrimination analysis (LDA) (de Chazal, et al., 2004), (Llamedo & Martinez, 2011), and ensemble methods (Zeng, et al., 2011) are considered in different papers.

Two different evaluation schemes, namely "class-oriented" and "subject-oriented", are used in literature. Heartbeat segmentation of extensive ECG signals such as MIT-BIH arrhythmia records, results in more than a thousand heartbeats for an individual record. As a result, it is highly probable that in random selection of training set for a supervised classifier there will be samples of the same patients used in both training and testing sets. This method is referred to as "class-oriented" and it is criticized for poor generalization performance. While class-oriented method is used in many papers (e.g., (de Oliveira, et al., 2011), (Jiang, et al., 2006), (Lagerholm, et al., 2002), (Osowski, et al., 2004), (Prasad & Sahambi, 2003), (Rodriguez, et al., 2005)), having particularly correlated samples in training and testing sets may cause the overfitting problem and lead to promising results that might not be reachable in practice. In contrast, Subject-oriented method, which was proposed by (de Chazal, et al., 2004), creates

a patient-based division of dataset into training and testing sets i.e. prior to heartbeat segmentation procedures. This method is believed to be more realistic and is adopted by (de Lannoy, et al., 2012), (Llamedo & Martinez, 2011), (Ye, et al., 2012).

Although promising results have already been seen in cardiac arrhythmia classification, there is a long way to go before applicable methods and totally automatic classifications can be employed in practice. This means more accurate and stable algorithms should be developed.

The paper is organized as follows: Section 2 describes the dataset, section 3 provides background materials and methods and section 4 represents the proposed method; in section 5 we share the simulation results and compare it with previous works and section 6 has the conclusion.

## 2. Background

### 2.1. Higher Order Statistics

Probability distribution moments are the generalization of concept of the expected value and can be used to define the characteristics of probability density function (Ebrahimzadeh & Khazaei, 2011).

Skewness (3<sup>rd</sup> order statistics) and kurtosis (4<sup>th</sup> order statistics) of random variable,  $x$ , are defined respectively as follows,

$$\gamma_3 = \frac{E\{[x - E(x)]^3\}}{(E\{[x - E(x)]^2\})^{3/2}}, \quad (1)$$

$$\gamma_4 = \frac{E\{[x - E(x)]^4\}}{(E\{[x - E(x)]^2\})^2} - 3. \quad (2)$$

In which  $E$  denotes the expected value. While skewness gives a measure of the lopsidedness of the distribution, kurtosis gives a comparative measurement of the signal's distribution with normal distribution of the same variance. Estimates of the two are given respectively by,

$$\hat{\gamma}_3 = \frac{\sum_{i=1}^N (x_i - \hat{m})^3}{(N-1)\hat{\sigma}^3}, \quad (3)$$

$$\hat{\gamma}_4 = \frac{\sum_{i=1}^N (x_i - \hat{m})^4}{(N-1)\hat{\sigma}^4} - 3, \quad (4)$$

where  $x_i$  is a realization of the random variable  $x$ .  $\hat{m}$  and  $\hat{\sigma}$  are estimates of mean and variance for the observed signal with  $N$  samples.

### 2.2. Gaussian Mixture Model

A Gaussian mixture model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs with many degrees of freedom allow arbitrary density modeling. In this model, data  $\{x_1, x_2, \dots, x_n\}$  in  $R^d$  are assumed to arise from a random vector with density,

$$f(\mathbf{x}) = \sum_{i=1}^K p_i \phi(\mathbf{x}|\mu_i, \Sigma_i). \quad (5)$$

In which,  $K$  is the number of mixture components,  $p_i$  is the mixture weight, i.e. prior probability of a certain component and  $\phi(\mathbf{x}|\mu, \Sigma)$  is the normal distribution with mean vector and covariance matrix of  $\mu$  and  $\Sigma$ , respectively. Generally, the mixing parameters  $\theta = (p_1, \dots, p_k, \mu_1, \dots, \mu_k, \Sigma_1, \dots, \Sigma_k)$  are estimated by maximizing the log-likelihood,

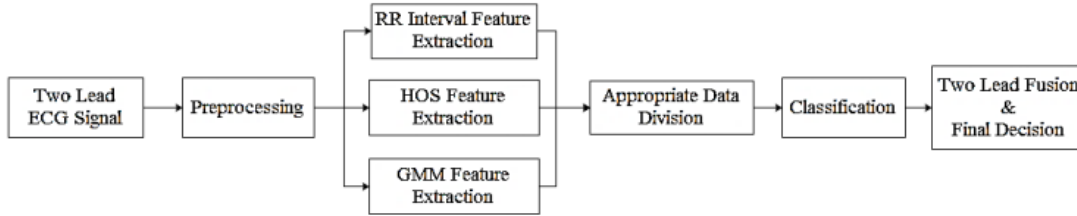


Fig. 1. Block diagram of proposed method for cardiac arrhythmia classification

$$L(\theta|x_1, \dots, x_n) = \sum_{j=1}^n \ln \left[ \sum_{i=1}^K p_i \phi(x_j|\mu_i, \Sigma_i) \right]. \quad (6)$$

A widely used approach for maximum likelihood solution is the Expectation Maximization (EM) algorithm. EM is an iterative algorithm that starts from an initial parameter  $\theta^0$ . The E step computes conditional probabilities of  $\hat{p}_i(x_j)$  where  $(1 \leq i \leq K, 1 \leq j \leq n)$  and  $x_j$  emerges from the  $i$ th mixture component and the M step updates  $\hat{p}_i, \hat{\mu}_i, \hat{\Sigma}_i$  with maximum likelihood estimation using  $\hat{p}_i(x_j)$  as conditional mixing weights (Biernackia, et al., 2003).

### 2.3. Decision Trees

A Decision tree (DT) is a classifier expressed as a recursive partition of the instance space. Discernibility, ability to handle missing attributes, having characteristics of non-parametric classifiers and the implementation simplicity are the main advantages of DTs. A decision tree is made of nodes and edges. The tree starts from a node called *root node*, which has no incoming edge. Nodes with outgoing edges are called *internal nodes* or *test nodes* and all the other nodes are called *leaves* or *terminal nodes*. Internal nodes are a split of their input space into two or more subspaces according to a certain discrete function of the input attributes values. Moreover, each leaf node may hold a probability vector indicating the probability of the target attribute having a certain value. According to the decisions made by internal nodes, instances are navigated from the root of the tree down to a leaf and then classified. This process faces the task of recursive partitioning the input space. The input space is generally represented by a training set like as

$$\mathcal{L} = \{(x_n, y_n) | n = 1, \dots, N\}. \quad (7)$$

This set includes  $N$  instances which are represented by a feature vector  $x$  and its associated class,  $y$ . After training the tree with the set  $\mathcal{L}$ , the new instances propagate through the tree and assign to the class which the leaf belongs (Maimon & Rokach, 2010) (Zaunseder, et al., 2014).

### 2.4. Ensemble Learners

Ensemble is a supervised learning algorithm that combines several *weak learners* (also called *base learners*) which are slightly better than random guessing to construct one *strong learner*. Integrating these weak learners can be done by various methods, like as majority voting or weighted aggregation of individual results. It is evident that combining several identical classifiers provides no gain, so this method can be useful only if there is dissimilarity among learners. Thus, the diversity of weak learners is crucial in the efficiency of ensemble learners. Diversity can be achieved by using different presentations of the input data, as in bagging. Variations in learner design and adding a penalty to the output are other methods to encourage diversity (Maimon & Rokach, 2010), (Opitz & Maclin, 1999), (Zaunseder, et al., 2011).

Bagging was proposed by (Breiman, et al., 1984) and it is the most well-known *bootstrap* ensemble method that processes

samples concurrently. Bagging generates individuals for its ensemble by training each inducer on a random reconstruction of the training set. The training set for a particular classifier is generated by randomly drawing  $N$  samples with replacement, where  $N$  is the size of original training set. As a result, each individual learner in the ensemble is configured with a different random sampling of the training set. This means that the initial training samples might appear zero or multiple times in a certain training set. Bagging combines these learners by majority voting, i.e. the most voted class is selected. The obtained strong learner will have the misclassification error lower than a particular inducer because the dissimilarity among these learners can compensate for the increase in error rate of each single learner. Breiman notes that bagging is more effective on *unstable* learners because it can remove their uncertainty. In this context, a learner is considered unstable if perturbing the training set can result in large changes in predictions. He claimed that decision trees are example of unstable learning algorithms.

## 3. Method

This section introduces the methodology used in the paper. Firstly, we talk about the dataset and then we follow the overall processing steps illustrated in Fig. 1. After preprocessing, which is baseline wandering removal and beat segmentation, we extract three sets of features; higher order statistical features, RR interval features and mixture modeling features. These features will then be fed to the classifier.

### 3.1. Dataset

We have used MIT-BIH arrhythmia database in our study, which includes many common and life-threatening arrhythmias. The database contains 48 half-hour ambulatory ECG recordings, obtained from 47 subjects. Each record consists of two leads. For 45 of the recordings the first lead is modified limb lead II (MLII) and for the others it is modified lead V5. The second lead on the other hand, is a pericardial lead (V1 for 40 recordings and V2, V4 or V5 for the others). Twenty-three recordings of this database were chosen at random from a set of four thousand ambulatory ECG recordings collected from a mixed population of inpatients; the remaining 25 recordings were selected from the same set to include less common but clinically significant arrhythmias that wouldn't be well-presented in a small random set (Mark & Moody, 1997).

The original labeling of the database includes 16 classes of rhythms shown in Table 1. Despite Standards recommended for class labeling and results representation of cardiac arrhythmia classification algorithms by the Association for the Advancement of Medical Instrumentation (AAMI) (Anon., 1998), (Anon., 1987), only (de Chazal, et al., 2004), (de Lannoy, et al., 2012), (Ince, et al., 2009), (Jiang & Kong, 2007), (Llamedo & Martinez, 2011), (Rodriguez, et al., 2005), (Yang & Shen, 2013), (Ye, et al., 2012), have utilized these standards. The AAMI defines the five clinically relevant classes as "N" (beats originating in the sinus node), "S" (supraventricular ectopic beats), "V" (ventricular ectopic beats), "F" (fusion beats) and "Q"



Table 1. MIT-BIH arrhythmia database information

Heartbeat Type	Ann*	Total #	
Normal Rhythm	NOR	N	74607
Left Bundle Branch Block	LBBB	L	8069
Right Bundle Branch Block	RBBB	R	7250
Atrial Premature Contraction	APC	A	2514
Premature Ventricular Contraction	PVC	V	7127
Paced Beat	PB	/	7020
Aberrated Atrial Premature Beat	AP	a	150
Ventricular Flutter Wave	VF	!	472
Fusion of Ventricular and Normal Beat	VFN	F	802
Non-Conducted P-wave (Blocked APC)	BAP	x	193
Nodal (Junctional) Escape Beat	NE	j	229
Fusion of Paced and Normal Beat	FPN	f	982
Ventricular Escape Beat	VE	E	106
Nodal (Junctional) Escape Beat	NP	J	83
Atrial Escape Beat	AE	e	16
Unclassified Beat	UN	Q	35
Total			109655

\* Annotation that is used for each arrhythmia in the database

(unclassified beats). Table 2 shows the mapping from the MIT-BIH arrhythmia database classes to the AAMI heartbeat classes. Table 2 also contains the train and test set division schemes (DS1 and DS2, respectively) introduced by (de Chazal, et al., 2004), for subject-oriented classification. This scheme carefully divides the dataset in a way that training and testing samples are fairly distributed among the mentioned five classes and also discards the four paced records (i.e. 102, 104, 107 and 217) as recommended by AAMI.

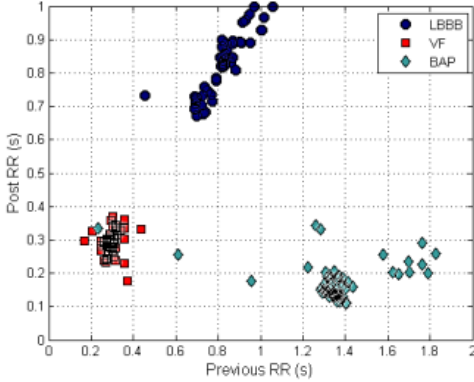


Fig. 2. Samples of post-RR and previous-RR features for three chosen classes of arrhythmia

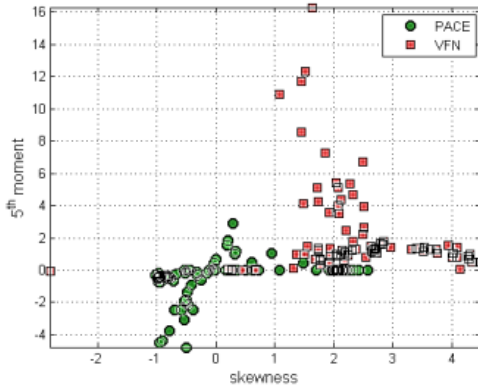


Fig. 3. Samples of higher order statistical features for four chosen classes of arrhythmia

We use all the 48 recordings and 16 classes of MIT-BIH annotations in class-oriented scheme and 44 non-paced recordings with 5 classes of Table 2 as recommended by AAMI in subject-oriented scheme. Thus our proposed method is comparable with most of algorithms developed for ECG arrhythmia classification.

### 3.2. Data Preprocessing

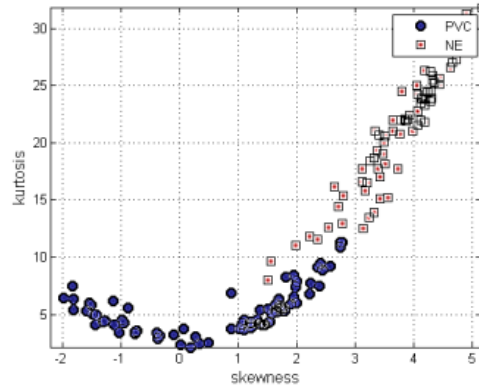
The data of MIT-BIH arrhythmia database are band-pass filtered at 0.1-100 Hz and digitized at 360 samples per second. We remove the baseline wandering of these signals using two stages of median filtering as proposed by (Awodeyi, et al., 2013).

An annotation file comes with each recording, indicating the occurrence sample of major local extremum for each individual heartbeat, which generally appears at R-wave peak. In this study only the information of R peaks are of interest and since the detection of these points are not investigated here, we use the annotation files to automatically obtain these instances and then manually correct them on the beat-by-beat basis as suggested by (de Chazal, et al., 2004).

The QT intervals (measured from the beginning of the QRS complex to the end of the T wave) can vary up to 420 ms in low heart rates and PR intervals (measured from the beginning of the P wave to the beginning of the QRS complex) take values from 120 ms up to 200 ms (Jenkins & Gerred, n.d.). For each heartbeat knowing the R location, we take samples from the interval of 250 ms before R peak and 400 ms after R peak, i.e. total 0.65 s of each heartbeat. This choice of fixed size beat segmentation is long enough to catch the samples representing the repolarization of ventricles and short enough not to get the samples of T wave from the previous beat. While usually longer intervals are selected in studies (e.g. 0.83 s intervals are used by (Ye, et al., 2012)), our prime concern is to exclude any samples of neighbor heartbeats. Narrow QRS complexes and higher heart rates are

Table 2. AAMI recommended labeling with training set (DS1) and testing set (DS2) used in subject-oriented scheme

AAMI Class	MIT-BIH Class	Total #
N	NOR, LBBB, RBBB, AE, NE	89665
S	APC, AP, BAP, NP	2940
V	PVC, VE, VF	7478
F	VFN	802
Q	FPN, UN	17
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230	
DS2	100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234	



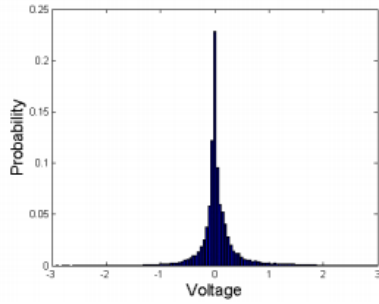


Fig. 4. Probability density of a thousand randomly selected heartbeats

very common among class ‘S’ arrhythmias. These characteristics can cause problems in long beat-segmentation schemes and consequently poor classification results for this class. In the meantime, fixed size segmentation method helps us avoid computational complexity of detecting other time-domain features.

### 3.3. Feature Extraction

Disorders in electrical activity of the heart can affect the time needed for repolarization and/or depolarization (e.g. bundle branch blocks cause prolonged QRS complexes). In order to exploit these characteristics of different heart rhythms, we add two RR interval features. We also believe that statistical features can delineate the ECG signal good enough for an accurate classification, so we focus on higher order statistics and mixture modeling to extract eight statistical features.

#### 1) RR Interval Features

We have chosen RR interval information as the only time-domain features in our study. Two RR intervals are calculated directly from the R locations named as previous-RR and post-RR intervals. Previous-RR is defined as the time distance between current and previous R location while post-RR is the time distance between current R location and the following one.

RR interval features can hold indispensable information about heart rhythms so they were adopted by various studies such as (de Chazal, et al., 2004), (de Lannoy, et al., 2012), (de Oliveira, et al., 2011), (Ince, et al., 2009), (Lagerholm, et al., 2002), (Llamedo & Martinez, 2011), (Osowski, et al., 2004), (Prasad & Sahambi, 2003), (Yang & Shen, 2013), (Ye, et al., 2012). In Fig. 2 we have illustrated these features for 50 randomly selected heartbeats of three chosen arrhythmias.

#### 2) Higher Order Statistical Features

It was proved that higher order statistics (HOS) are less sensitive to morphological changes of ECG and the nonlinear and dynamic nature of these features help to extract the subtle changes in ECG data (Martis, et al., 2011). Three statistical features as skewness, kurtosis and 5<sup>th</sup> moment are calculated for each heartbeat. Fig. 3 shows these three features for 80 randomly selected heartbeats of four chosen arrhythmias.

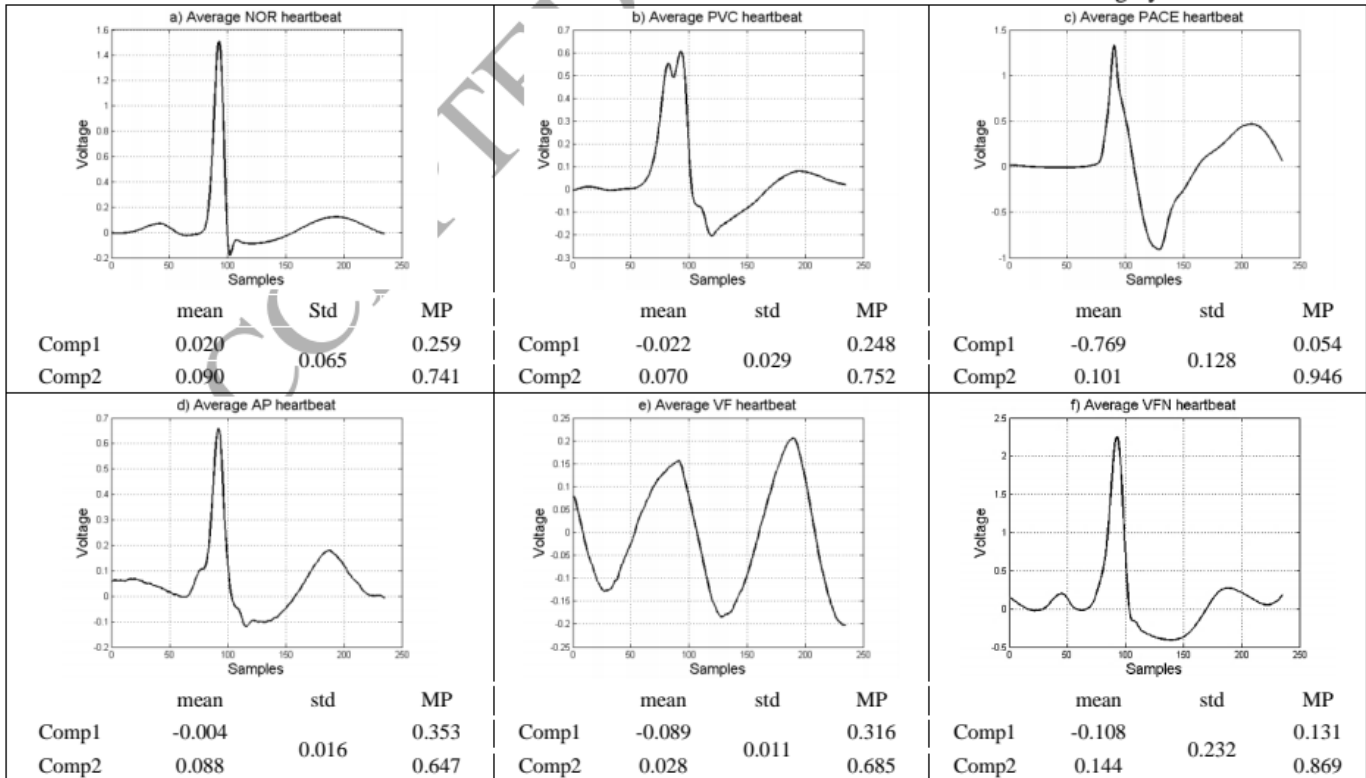
#### 3) Mixture Modeling Features

We perform an intrabeat Gaussian mixture modeling for each individual heartbeat. It’s known that each arrhythmia can cause a particular change in the recorded voltage of ECG signal. Thus, a prudent mixture modeling of single heartbeat samples can provide applicable information to be used in the main classification. We use a 2-component EM based Gaussian mixture modeling for each signal. Since EM is to be calculated for over 100,000 heartbeats, we set a relatively relaxed threshold to stop the iteration, which is,

$$\frac{L^q - L^{q-1}}{L^q} < 10^{-5}, \quad (8)$$

where  $L^q$  denotes the observed log-likelihood at  $q$ th iteration mentioned in (6).

The EM algorithm usually starts from some initial estimate of model parameters,  $\theta^\circ$ , (e.g. random) and then proceeds to iteratively updating  $\theta$  until a certain convergence criteria is detected. However, an important drawback of EM is that in a multivariate context its solution highly relies on the starting



Comp 1&2 are the mixture components; std is the standard deviation and MP is the mixing proportion for each component

Fig. 5. Average heartbeat for six different arrhythmias and parameters of GMM estimation for each one

parameters and consequently the algorithm generates poor maximum likelihood estimates (Biernackia, et al., 2003). Considering our enormous database, the goal is to find a simple method that would give the highest likelihood in the least possible number of iterations. We examine the statistics of one thousand heartbeats extracted randomly from the training set. Fig. 4, shows the probability density function (PDF) of this signal which has the mean and standard deviation of 0.065 and 0.45, respectively. It is apparent that the PDF has a dominant peak around zero, as the most samples of ECG signal have the voltage near zero with the tendency toward positive voltages, resulting from QRS and T peaks. Consequently, we considered initiating the EM algorithm with one assertive Gaussian centered at the origin and an auxiliary Gaussian with a small positive mean. Empirically we found an optimum set of initiation parameters, as  $\theta^0 = (0.6, 0.4, 0, 0.1, 0.01, 0.01)$  (i.e. 60% and 40% mixing proportions with 0 and 0.1 mean for each component, respectively and the two GMMs share the same variance,  $\Sigma_i$ , with initial value of 0.01). Proposed initiation parameters and log-likelihood threshold of EM algorithm leads to a fast convergence within less than 6 iterations for each heartbeat. Final parameters of the two Gaussians are considered as the GMM features, including two mean values, two mixing proportions and one standard deviation. Fig. 5, shows the average signal of training set for six chosen arrhythmias with their calculated GMM parameters.

#### 3.4. Classification

Bootstrap aggregating ensemble method is used to combine 100 decision tree learners and make a more stable and accurate classifier. While a single decision tree has a high variance and is considered as unstable, bagging reduces the variance and helps to avoid overfitting problem. A research by (Wang, et al., 2009) showed that bagging outperforms basic decision trees and decision trees improved by boosting.

As mentioned in section II, each signal in MIT-BIH database includes two leads. All the procedures such as feature extraction and classification that are illustrated in Fig. 1, are independently applied to both leads. Preferably the two outcomes for each heartbeat should point to the same class, so we consider any inconsistency of the results as a misclassification error.

#### 4. Evaluation and Results

Results of classification are evaluated in two schemes: class-oriented and subject-oriented, which differ only in dataset divisions and class labeling.

##### 4.1. Performance Metrics

Various approaches are adopted in literature to evaluate the results. In this study, sensitivity and positive predictivity are used. Sensitivity (Se) can be defined as a measure of successfully classified samples,

$$Se = \frac{TP}{TP + FN} \times 100, \quad (9)$$

where  $FN$  is the total number of falsely missed negative samples and  $TP$  is the total number of correctly classified positive samples. Positive predictivity (Pp) measures exclusive classification of different samples and can be defined as,

$$Pp = \frac{TP}{TP + FP} \times 100, \quad (10)$$

where  $FP$  is the total number of falsely classified positive samples.

##### 4.2. Class-Oriented Evaluation

The class-oriented evaluation is performed on the original 16-class labeling of MIT-BIH database. As shown in Table 4, total number of 109655 heartbeats from different classes are divided into training and testing sets. 12% of normal rhythms, 40% of common arrhythmias (LBBB, RBBB, APC and PVC) and 50% of less common arrhythmias are randomly selected to make the training set, which includes 23984 samples (21.87% of all data). Table 4 also contains the number of *rejected* instances (i.e. number heartbeats that got different labels from classifying each lead).

##### 4.3. Subject-Oriented Evaluation

The subject-oriented evaluation has been introduced in literature to get more realistic results with help of patient-based division of database (i.e. DS1 and DS2 in Table 2). This scheme is based on the AAMI recommendations to use the five-class labeling standard and discard the four paced records. Subsequently, the size of database reduces to 100902 samples as in Table 3, while the training set holds 51507 samples (51.05% of all data). Table 3 also has the classification results for this scheme.

Table 3. Results for subject-oriented classification scheme

Class	Total #	Train	Test	Se (%)	Pp (%)	Rn
N	89665	45826	43839	97.37	98.40	1153
S	2940	999	1941	86.50	90.90	262
V	7478	4260	3218	95.99	77.63	129
F	802	414	388	11.86	24.21	342
Q	17	8	9	0	-	9
total	100902	51507	49395	96.16	96.15	1895

Table 4. Results for class-oriented classification scheme

Rhythm	Total #	Training	Test	Se (%)	Pp (%)	Rn*
NOR	74607	8953	65654	99.77	99.92	148
LBBB	8069	3228	4841	99.98	100	1
RBBB	7250	2900	4350	98.76	96.68	54
APC	2514	1006	1508	100	98.69	0
PVC	7127	2851	4276	100	100	0
PB	7020	3510	3510	99.97	99.97	1
AP	150	75	75	81.33	85.92	14
VF	472	236	236	100	100	0
VFN	802	401	401	97.51	99.74	10
BAP	193	97	96	100	100	0
NE	229	115	114	94.74	88.52	6
FPN	982	491	491	100	98.00	0
VE	106	53	53	86.79	97.87	7
NP	83	42	41	100	100	0
AE	16	8	8	0	-	8
UN	35	18	17	41.18	100	10
Total	109655	23984	85671	99.70	99.69	259

\* Number of rejected heartbeats for each class

Table 5. Comparative results of class-oriented classification

Reference	Features	Accuracy
(Jiang, et al., 2006)	Wavelet + ICA	98.86
(Ye, et al., 2012)	Wavelet + ICA + RR	99.71
(Lagerholm, et al., 2002)	Hermite	98.49
(Osowski, et al., 2004)	HOS + Hermite	98.18
(Prasad & Sahambi, 2003)	Wavelet + RR	96.77
(Rodriguez, et al., 2005)	Waveform	96.13
(de Oliveira, et al., 2011)	Waveform + RR	98
<b>Proposed</b>	<b>HOS + Mixture Model + RR</b>	<b>99.70</b>



#### 4.4. Results

Table 4 shows the results of class-oriented scheme, all 8 cases of atrial escape beats (AE) in the training set has got different labels from each lead and flagged as *rejected* ( $Rn_{AE} = 8$ ), leading to zero accuracy for AE class. Unclassified beats (UN) doesn't show good results either, 10 beats out of 17 has been flagged as rejected but for other 14 classes the results are very accurate. Total 259 beats are rejected which shows a significant improvement in compare to 2054 rejected beats of (Ye, et al., 2012). Table 5, shows methods and overall accuracy of different studies with class-oriented schemes. Total accuracy of our proposed method for this scheme is 99.70% and it's worth mentioning that unlike (Ye, et al., 2012), we have included rejected beats in calculation of accuracy, sensitivity and positive predictivity.

Subject-oriented scheme, as shown in Table 3, has 1895 rejected instances. The truth table for subject-oriented scheme is illustrated in Table 6. As Table 3 shows, classification of F and Q classes show poor results, which is the case in other papers as well. For three main classes of N, S and V our method certainly outperforms previous works as compared in Table 7.

#### 5. Conclusion

In this paper, we have presented a novel feature selection method for classification of cardiac arrhythmias. The comparative results of our proposed method, as shown in Table 5 and Table 7, indicates the high performance of Gaussian mixture models to fit the data and preserve the essential information for an accurate classification. Ensemble learning methods which are seldom used in biomedical signal processing are adopted in this work and the statistical advantages of these classifiers helped in the classification of unevenly distributed arrhythmia groups.

Using both class-oriented and subject-oriented schemes enables us to compare the results with other methods in this area of research. It is apparent that statistical features and mixture model parameters need to get more attention on processing biomedical signals.

In the classification of ECG arrhythmias, different types of features are usually treated the same, training the classifiers. For the future studies we suggest categorizing features based on their characteristics and then combining ensembles of classifiers

trained with each group of features for a final classification.

#### References

- Anon., 1987. Recommended Practice for Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms, s.l.: Association for the Advancement of Medical Instrumentation.
- Anon., 1998. Testing and Reporting Performance Results of Cardiac Rhythm and ST Segment Measurement Algorithms, s.l.: Association for the Advancement of Medical Instrumentation.
- Awodeyi, A., Alty, S. & Ghavami, M., 2013. Median Filter Approach for Removal of Baseline Wander in Photoplethysmography Signals. *Manchester, s.n.*, pp. 261-264.
- Biernackia, C., Celeuxb, G. & Govaertc, G., 2003. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Computational Statistics & Data Analysis*, 41(3), pp. 561-575.
- Breiman, L., Friedman, J., Stone, C. J. & Olshen, R., 1984. *Classification and Regression Trees*. s.l.:Chapman and Hall/CRC.
- de Chazal, P., O'Dwyer, M. & Reilly, R., 2004. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering*, 51(7), pp. 1196-1206.
- de Lannoy, G., Francois, D., Delbeke, J. & Verleysen, M., 2012. Weighted Conditional Random Fields for Supervised Interpatient Heartbeat Classification. *IEEE Transactions on Biomedical Engineering*, 59(1), pp. 241-247.
- de Oliveira, L., Andreao, R. & Sarcinelli-Filho, M., 2011. Premature Ventricular beat classification using a dynamic Bayesian network. *Boston, MA, s.n.*
- Ebrahimzadeh, A. & Khazaei, A., 2011. Higher order statistics for automated classification of ECG beats. *Yichang, s.n.*
- Ince, T., Kiranyaz, S. & Gabbouj, M., 2009. A Generic and Robust System for Automated Patient-Specific Classification of ECG Signals. *IEEE Transactions on Biomedical Engineering*, 56(5), pp. 1415-1426.
- Jenkins, D. & Gerred, S., n.d. Normal ECG. [Online] Available at: <http://www.ecglibrary.com/norm.php> [Accessed February 2015].
- Jiang, W. & Kong, S., 2007. Block-Based Neural Networks for Personalized ECG Signal Classification. *IEEE Transactions on Neural Networks*, 18(6), pp. 1750-1761.
- Jiang, X., Zhang, L., Zhao, Q. & Albayrak, S., 2006. ECG Arrhythmias Recognition System Based on Independent Component Analysis Feature Extraction. *Hong Kong, s.n.*
- Lagerholm, M., Peterson, C., Braccini, G. & Edenbrandt, L., 2002. Clustering ECG complexes using Hermite functions and self-organizing maps. *IEEE Transactions on Biomedical Engineering*, 47(7), pp. 838-848.
- Llamedo, M. & Martinez, J., 2011. Heartbeat Classification Using Feature Selection Driven by Database Generalization Criteria. *IEEE Transactions on Biomedical Engineering*, 58(3), pp. 616-625.
- Maimon, O. & Rokach, L., 2010. *Data Mining and Knowledge Discovery Handbook*. s.l.:Springer.
- Mark, R. & Moody, G., 1997. MIT-BIH Database and Software Catalog. [Online] Available at: <http://ecg.mit.edu/dbinfo.html> [Accessed February 2015].
- Martis, R., Acharya, R. & Ray, A., 2011. Application of Higher Order Cumulants to ECG Signals for the Cardiac Health Diagnosis. *Boston, s.n.*
- Opitz, D. & Maclin, R., 1999. Popular Ensemble Methods: An Empirical Study. *Journal of Artificial Intelligence Research*, Volume 11, pp. 169-198.
- Osowski, S., Hoai, L. T. & Markiewicz, T., 2004. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering*, 51(4), pp. 582-589.
- Prasad, G. & Sahambi, J., 2003. Classification of ECG arrhythmias using multi-resolution analysis and neural networks. *Conference on Convergent Technologies for the Asia-Pacific Region, TENCON, Volume 1*, pp. 227-231.
- Rodriguez, J., Gofii, A. & Illarramendi, A., 2005. Real-time classification of ECGs on a PDA. *IEEE Transactions on Information Technology in Biomedicine*, 9(1), pp. 23-34.
- Wang, H., Jiang, Y. & Wang, H., 2009. Stock return prediction based on Bagging-decision tree. *Nanjing, s.n.*
- Yang, S. & Shen, H., 2013. Heartbeat Classification using discrete wavelet transform and kernel principal component analysis. *Sydney, NSW, s.n.*, pp. 34-38.

Table 6. Confusion matrix for subject-oriented scheme

		Predicted Results				
		N	S	V	F	Q
Reference	N	42686	152	859	138	0
	S	235	1679	22	5	0
	V	121	4	3089	1	0
	F	332	4	6	46	0
	Q	5	1	3	0	0

Table 7. Comparative results of subject-oriented classification

Method	N		S		V	
	Se	Pp	Se	Pp	Se	Pp
(Llamedo & Martinez, 2011)	95	98	77	39	81	87
(Ye, et al., 2012)	88.6	97.5	60.8	52.3	81.5	63.1
(de Chazal, et al., 2004)	87.1	99.2	75.9	38.5	77.7	81.6
<b>Proposed</b>	<b>97.4</b>	<b>98.4</b>	<b>86.5</b>	<b>90.9</b>	<b>96.0</b>	<b>77.6</b>

25. Ye, C., Kumar, B. & Coimbra, M., 2012. Heartbeat Classification Using Morphological and Dynamic Features of ECG Signals. *IEEE Transactions on Biomedical Engineering*, 59(10), pp. 2930-2941.
26. Zaunseder, S., Huhle, R. & Malberg, H., 2011. CinC Challenge - Assessing the Usability of ECG by Ensemble Decision Trees. *Hangzhou, s.n.*, pp. 277-280 .
27. Zeng, X. D., Chao, S. & Wong, F., 2011. Ensemble Learning on Heartbeat Type Classification. *Macao, s.n.*

ACCEPTED MANUSCRIPT