Biomedical Signal Processing and Control 19 (2015) 56–67

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Classification of AAMI heartbeat classes with an interactive ELM ensemble learning approach

Mohamad M. Al Rahhal [a], Yakoub Bazi [a,*], Naif Alajlan [a], Salim Malek [a], Haikel Al-Hichri [a], Farid Melgani [b], Mansour A. Al Zuair [a]

[a] Dept. of Computer Engineering, College of Computer and Information Sciences, King Saud University, PO Box 51178, Riyadh 11543, Saudi Arabia
[b] Dept. of Information Engineering and Computer Science, Univ. of Trento, Via Sommarive, 14, I-38123 Trento, Italy

ARTICLE INFO

Article history:
Received 24 December 2014
Received in revised form 24 February 2015
Accepted 18 March 2015
Available online 11 April 2015

ABSTRACT

In recent years, the recommendations of the Association for the Advancement of Medical Instrumentation (AAMI) for class labeling and results presentation are closely followed as a possible solution for standardization. Regardless of the class normalization, this standard basically recommends for performance evaluation to adopt inter-patient scenarios, which renders the classification task very challenging due to the strong variability of ECG signals. To deal with this issue, we propose in this paper a novel interactive ensemble learning approach based on the extreme learning machine (ELM) classifier and the induced ordered weighted averaging (IOWA) operators. While ELM is adopted for ensemble generation the IOWA operators are used for aggregating the obtained predictions in a nonlinear way. During the iterative learning process, the approach allows the expert to label the most relevant and uncertain ECG heart beats in the data under analysis and then adds them to the original training set for retraining. The experimental results obtained on the widely used MIT-BIH arrhythmia database show that the proposed approach significantly outperforms state-of-the-art methods after labeling on average 100 ECG beats per record. In addition, the results obtained on four other ECG databases starting with the same initial training set from MIT-BIH confirm its promising generalization capability.

1. Introduction

The electrocardiogram (ECG) signal contains valuable information about patient's heart activity. The monitoring and analysis of ECG signals represents an efficient way for the early detection of different cardiac diseases. To this end, many researchers devoted their efforts over the years to develop computer-based methods for arrhythmia detection and classification. Nevertheless, the comparison across most of these methods could not be performed fairly, due to the lack of standardization in the development and evaluation criteria.

In recent years, the recommendations of the Association for the Advancement of Medical Instrumentation (AAMI) for class labeling and results presentation are closely followed as a possible solution for standardization [1–6]. Typically, the AAMI standard defines five classes of interest: normal (N), ventricular (V), supraventricular (S), fusion of normal and ventricular (F) and unknown beats (Q). Regardless of the class definition, this standard recommends essentially for performance evaluation to adopt an inter-patient scenarios, which is not usually adopted in most of the works published in the literature. This requirement renders the automatic classification task very challenging as the test subjects are unseen during the classifier design. Although, various feature representation (e.g., morphological, temporal, wavelets, higher order statistics, etc.) as well as many classifiers (e.g., linear discriminant analysis, neural networks, support vector machines, etc.) were used, the results obtained by the automatic methods remain unsatisfactory.

To tackle this issue semiautomatic approaches allowing expert interaction were introduced [7–11]. Usually, these approaches train a global classifier on a large dataset and another local-classifier on the first few minutes from the record under analysis labeled by the expert. Then the outputs of both classifiers are fused using voting rules to classify the entire record. However, this mode of labeling does not take into consideration how much are these signals relevant for boosting the classification accuracy. Indeed, the generalization ability of the classifiers depends strongly on the samples

* Corresponding author. Tel.: +966 506469808.
E-mail addresses: malrahhal@ksu.edu.sa (M.M.A. Rahhal), ybazi@ksu.edu.sa (Y. Bazi), najlan@ksu.edu.sa (N. Alajlan), smalek@ksu.edu.sa (S. Malek), haikel@ksu.edu.sa (H. Al-Hichri), melgani@disi.unitn.it (F. Melgani), zuair@ksu.edu.sa (M.A.A. Zuair).

http://dx.doi.org/10.1016/j.bspc.2015.03.010
1746-8094/

that represent well the statistical distribution of the data. So, it would be necessary to design a system that allows us to define better mechanisms for selecting and labelling samples fundamental to the correct discrimination between the set of considered classes.

In this work, we propose a novel ensemble method for the interactive classification of AAMI heart beat classes. The choice of ensemble strategies is mainly motivated by their robustness compared to single-based classifier designs [12–16]. Fig. 1 provides a general view of the proposed method which is composed of the following iterative steps: (1) ensemble construction, (2) fusion; and (3) ECG beat selection and labeling. For the ensemble construction step, we use the extreme learning machine (ELM) classifier that gained popularity in recent years as an efficient class of learning algorithms for single-hidden layer feedforward neural networks (SLFNs) [16,17]. The main concept behind the ELM lies in the random choice of the SLFN hidden layer weights and biases, i.e. these hidden layer parameters need not be tuned, unlike regular neural networks or SVM variant methods. The output weights are determined analytically, thus the network is solved with very few steps and with low computational cost. Another interesting feature of the ELM is that it provides a unified learning platform with widespread type of feature mappings which could be done either in a known space similar to neural networks or in an infinite space similar to kernel methods. In addition, it can be used for regression and multi-class classification applications directly.

For the fusion step, we use the induced ordered weighted averaging (IOWA) operators [18,19]. This nonlinear operator is an extension of the standard OWA operator proposed by Yager [20]. However, the difference is that the reordering step is done using an auxiliary value, called order-inducing variable, rather than using the actual outputs called argument values. To obtain automatically the weights associated with the IOWA fusion operators, we tailor the prioritized aggregation idea to the classification scenario [21]. Finally in the selection and labeling step, unlike state-of-the-art methods we do not allow the user to label the first few minutes but instead we use uncertainty criteria [22] to rank the ECG beats in terms of their ambiguity with respect to the fusion result obtained in the previous step. The most ambiguous samples are given to the expert for labeling and then injected in the training set for retraining. It is expected that this process will increase the generalization ability of the classification system on the difficult samples for the next iterations. The experimental results obtained on the MIT-BIH arrhythmia database show that the proposed system can provide significant improvements in terms of classification accuracy with a reduced number of expert interaction (on average 100 ECG beat per record). In addition, the results obtained on four other ECG databases using the same initial training from MIT-BIH confirm its promising generalization capability.

The rest of the paper is organized as follows. Section 2 provides a detailed description of the proposed interactive classification system. The experimental results obtained on five ECG databases from physionet are reported in Section 3. Finally, conclusions and future developments are drawn in Section 4.

2. Proposed interactive classification system

Let us consider $D=\left\{(X_i,y_i)\right\}_{i=1}^{N}$ a training set composed of N training ECG feature vectors $\mathbf{X}_i$ of dimension d (e.g., morphological and temporal features [1,22]) and $y_i \in \left\{1,\ldots,N_o\right\}$ is the corresponding class label where N_o represents the number of classes. Given this training set D, we aim to classify a new ECG test record using the interactive classification system shown in Fig. 1. The following algorithm provides a general description of the proposed approach, whereas details descriptions are provided in next subsections.

Algorithm 1 Interactive classification
Input:
- Training set $D=\left\{(X_i,Y_i)\right\}_{i=1}^{N}$
- Test record: $Rec=\left\{(X_\ell)\right\}_{\ell=1}^{M}$
- Ensemble size: P
- Number of interactions: $ITER$
- Number signals to label at each iteration: N_s

Output: Classification result
- Step 1: Generate an ensemble of P diverse training sets each of size L from D with *k-means; for* $Iter = 1 : ITER$
- Step 2: Train P-ELM estimators on the p training sets;
- Step 3: Classify Rec by aggregating the P-ELM predictions with IOWA^{PA} (see Algorithm 2);
- Step 4: Rank the signals of the test record Rec based on their uncertainty;
- Step 5: Ask an expert to label the top ranked N_s signals;
- Step 5: Augment the P training sets with these new labeled signals;
- Step 6: Output the final classification result.

2.1. Ensemble generation with k-means and ELM

It is well known that the key success in ensemble methods is to design accurate and diverse models. Diversity can be obtained in many ways such as using different classifiers, different feature representations or by sampling strategies. For more details, we refer the reader to [15] for a comprehensive review. In our context, we address this issue by clustering the global training set $D=\left\{(X_i,y_i)\right\}_{i=1}^{N}$ with the k-means clustering algorithm. This choice is clearly justified as the commonly used the training set (composed of 22 records of the MIT-BIH arrhythmia database) is very large and highly unbalanced. To this end, for each AAMI class we run the *k-means algorithm* to group it into N_{clust} clusters. Then we select the samples closest to the centeroids of each cluster to form a balanced training set of size $L=4\times N_{\text{clust}}$. This process is repeated P times with different initializations to generate P different training sets.

As base learning model, we consider in this work the ELM classifier which is characterized by several attractive proprieties: (i) it has a unified formulation for binary, multiclass and regression problems; (ii) the solution of these problems is given in a unified compact form; (iii) the feature mapping could be done either in known space similar to neural networks or in infinite space similar to kernel methods; (iv) for multiclass classification, ELM uses a configuration of multi-output nodes where the number of nodes is equal to the number of classes. Recently, ELM has shown notable results in several applications compared to other kernel methods [17,23].

Let $\mathbf{h}(\mathbf{x}_i) \in \Re^{1\times N_h}$ be the row output vector of the hidden layer with respect to $\mathbf{x}_i$ and $\boldsymbol{\beta} \in \Re^{N_h\times N_o}$ the output weights that connect the hidden layer with the output layer (which represents the number of classes). Then, the ELM output $\mathbf{f}(\mathbf{x}_i) \in R^{1\times N_o}$ is given by [16]:

$$\mathbf{f}(\mathbf{x}_i)=\mathbf{h}(\mathbf{x}_i)\beta, \quad i=1,\ldots,L. \tag{1}$$

ELM aims to determine the weights $\boldsymbol{\beta}$ by minimizing the following objective function:

$$\min_{\beta}\frac{1}{2}\left\|\beta\right\|_F^2+C\frac{1}{2}\sum_{i=1}^{L}\left\|e_i\right\|^2 \tag{2}$$

Subject to the constraints:

$$\mathbf{h}(\mathbf{x}_i)\beta=\eta_i^T,-\mathbf{e}_i^T, \quad i=1,\ldots,L. \tag{3}$$