Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

Comparison of seven approaches for holter ECG clustering and classification

Václav Chudáček, Milan Petrík, George Georgoulas, Miroslav Čepek, Lenka Lhotská, Chrysostomos Stylios

Abstract - **In this work we present a comparative study, testing selected methods for clustering and classification of holter electrocardiogram (ECG). More specifically we focus on the task of discriminating between normal 'N' beats and premature ventricular 'V' beats**

Some of the tested methods represent the state of the art in pattern analysis, while others are novel algorithms developed by us. All the algorithms were tested on the same datasets, namely the MIT-BIH and the AHA databases.

The results for all the employed methods are compared and evaluated using the measures of sensitivity and specificity.

I. Introduction

Long term holter [1] monitoring is used for patients with heart problems such as arrhythmias. Heart beats with unusual timing or unusual electrocardiogram (ECG) morphology can be very helpful in early diagnosis of hearts with pathological electrophysiology. Many different methods have been proposed to solve the problem of discrimination between normal ('N') and premature ventricular ('V') beats. Some are based on beat-shape description parameters [2], [3] others use frequency-based features [4].

A number of papers dealing with the classification of ECG usually test their proposed approaches on some non-standard dataset using only one method/algorithm. An objective comparison of such classification schemes obviously becomes problematic. In this paper we have tried to provide a thorough investigation of different pattern analysis techniques using two datasets.

The paper is structured as follows: In Section II we present seven approaches (some of them novel) for the task of distinguishing between 'V' and 'N' beats. Our intention is to compare the methods in a standardized way. Each presented method deserves a thorough description. Unfortunately, due to space limitations, we have described each method in few paragraphs. The interested reader can refer to the cited literature for more details.

For the purpose of comparison we use MIT-BIH and AHA (American Heart Association) databases [5], [6], described in Section III, which both have undergone the same preprocessing steps and from which the same features have been computed. In Section IV, the results are compared using the well established measures of sensitivity and specificity. Finally, Section V concludes the paper presenting a brief discussion and some hints for future work.

II. Methods

A. Rule-Based Decision Tree

In case of the *Rule-Based Decision Tree* (RBDT) [7], the domain of each feature (value measured on each ECG beat) is divided into several intervals, which are given by the experience and knowledge of clinicians. ECG beats are then sorted to "clusters" defined by the cartesian product of the intervals. For every cluster a representing ECG beat is found as the median of the cluster (see Fig.1). The medians are then used in the final classification step – along with a template matching approach- described in Section II.C.

The strength of the method lies on the significant decrease of the computational complexity as we do not need to work with the whole set of the beats.

B. Rule-Based Decision Tree Clustering Based on Fuzzy Intervals

Borders of the intervals in the RBDT method are given, as mentioned above, based on experience and knowledge. Moreover the intervals often represent notions, which are vague by their nature (e.g. ampR is expected to be "small", "medium", or "big"). This fact has inspired us to apply fuzzy logic and fuzzy sets [8]-[10] to enhance this method [11]. Therefore the intervals have been replaced by fuzzy intervals and the clusters by fuzzy clusters. Thus, some ECG beats may partially belong to more than one fuzzy cluster. To be able to determine a median for each fuzzy cluster and to compare the results with other methods we needed to

Manuscript received April 16, 2007. This work has been supported partially by the research program "Information Society" under grant No. 1ET201210527 "Knowledge-based support of diagnostics and prediction in cardiology" and by the joint research project "Intelligent System for Medical Data Analysis and Evaluation Based On Computational Intelligence Techniques" under the Greek-Czech bilateral cooperation agreement, funded by Greek GSRT and Czech MEYS.

V. Chudáček is with Gerstner laboratory at dept. of Cybernetics of Czech Technical University in Prague, Czech Republic (phone: 420-224357325; fax: 420-224916848; e-mail: chudacv@fel.cvut.cz).

M. Petrík is PhD. student of Czech Technical University in Prague, Czech Republic, dept. of Cybernetics, Center for Machine Perception (e-mail: petrikm@fel.cvut.cz).

G. Georgoulas is with Intelligent Control Systems Laboratory, Georgia Tech, Atlanta, USA (e-mail: ggeorgoulas3@mail.gatech.edu)

M. Čepek is PhD. student of Czech Technical University in Prague, Czech Republic, dept. of Computer Science and Engineering, Neural Computing Group (e-mail: cepekm1@fel.cvut.cz).

C. Stylios is with the Dept. of Informatics and Communications Technology, TEI of Epirus, Artas, Greece (e-mail: stylios@teiep.gr)

L. Lhotská is an assoc. professor of the Czech Technical University in Prague, Czech Republic, dept. of Cybernetics, Gerstner laboratory. (e-mail: lhotska@fel.cvut.cz)

propose a median for fuzzy sets and also to generalize the measures of sensitivity and specificity [11].

C. Self organizing map ANN

The *Self Organizing Map* (SOM) *artificial neural network* (ANN) is an unsupervised clustering algorithm [12]. SOM consists of neurons organized in a regular low-dimensional grid. Each neuron is represented by a *d*-dimensional weight vector, where *d* is equal to the dimension of the input vectors. The network calculates the distance between input and weight vectors of all neurons. The neuron whose weight is closest to the input is activated as the output of the network [12].

The implementation of SOM has been performed using MATLAB SOMToolbox [13] and consisted of 15x9 neurons in a hexagonal grid arrangement. We tested several different sizes before choosing the 15x9 grid architecture, which gave the best results.

D. Template matching

For clustering methods as RBDT, fuzzy RBDT and SOM we had to carry out the classification step separately using the *template matching* method.

Median of each cluster is compared with ten different 'N' and 'V' templates – five for each classification group. Then correlation coefficients [14] are computed. For the final decision on the cluster median majority, 2 out of the first 3 coefficients is used. All the beats in the cluster represented by the median are classified according to the classification of the corresponding median.

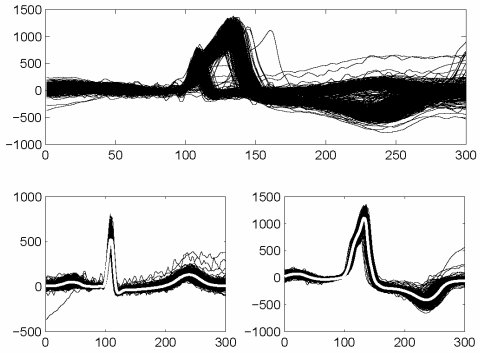

Fig.1 Example of full set (top) and computed clusters based on RBDT method (bottom) - white lines represent medians of the clusters.

E. Support Vector Machines

Support Vector Machines (SVMs) are learning systems that are trained using an algorithm based on optimization theory [15], [16]. The SVM solution finds a hyperplane in the feature space that keeps the empirical error small while maximizing the margin between the hyperplane and the instances closest to it. Every new pattern **x** is classified to either of the two categories (in case of dichotomizing problems $y_i \in \{-1,1\}$) through:

$$f(\mathbf{x}) = sign\left(\sum_{i=1}^{n} y_i a_i K(\mathbf{x}, \mathbf{x}_i) + b\right) \quad (1)$$

where b is a threshold parameter. The coefficients a_i are found by solving a maximization quadratic programming problem, which is "controlled" by a penalty factor C and are assigned to each of the training patterns $\mathbf{x}_i$. The kernel function K implicitly performs the mapping from the input space to the feature space. In our experimental procedure we have only employed the radial basis function kernels, where the width σ, which is common to all kernels, was specified a priori by the user [17].

F. Group of Adaptive Models ANN

The *Group of Adaptive Models* (GAME) ANN is based on an inductive approach. This means that both the parameters and the structure of the ANN are parts of a learning process (values for the parameters are selected and the ANN is constructed from some basic blocks during the learning process).

The GAME is a feed-forward ANN, which extends the concept of the *Group Method of Data Handling* (GMDH) network [18], [19]. GMDH allows only one type of basic block – neurons can have one transfer function only-whereas in GAME ANN there are neurons with many different transfer functions.

GAME ANN is built during the training phase from scratch. Each new layer is constructed in the following way: firstly a large number of new neurons is generated; the neurons differ in the transfer function and in the number of neurons in the previous layer that the new neuron is connected to. The next step is to find the optimal set of internal parameters and the best connections to neurons in the previous layer. To do this, GAME uses an advanced genetic algorithm. The population consists of neurons in every new layer and every neuron is coded to a genome. Genome coding contains information about neurons in the previous layer, which are connected to the neuron, the type of the transfer function and some parameters of the transfer function [20].

At the end of the genetic algorithm all neurons in a new layer are evaluated using a separated testing set and the worst neurons are removed from the layer. The remaining neurons in the layer are "frozen" and the algorithm continues with the creation of a new layer. This is repeated until a neuron with the desired accuracy is found; this neuron is the output of the network.

G. Back Propagation ANN

Back Propagation (BP) learning technique is a classical method for training feed-forward ANNs with one or more hidden layers [21]. In the recall phase of the algorithm, a sample is presented to the network and values are propagated from the input to the output of the network. During the training phase the training samples are presented to the network. The difference between the desired and the actual output is calculated formulating the network's error. This error is propagated backwards from output neurons toward the input and the weights are modified based on that.

The back-propagation network used in this study (selected after thorough testing of different configurations) had 7 neurons in the first hidden layer and 4 neurons in the second. For training standard BP algorithm was implemented in WEKA [22], with momentum and decreasing learning rate.

H. Radial Basis Function ANN

Radial Basis Function (RBF) ANN is trained with a fast supervised learning algorithm which is suitable both for regression and classification. It consists of one input, one output and one hidden layer, which contain RBF neurons. Each RBF neuron is described by a transfer function that represents a d-dimensional Gaussian "bump" (where d is the dimension of the input vectors), with the center at a point μ and width σ. Output neurons calculate the weighted sum of RBF neurons output. More details on RBF learning can be found in [23].

III. Data preprocessing

A. Data

For training and testing of the methods described above, the MIT-BIH arrhythmia database [5] and the AHA database [6] have been employed. From both databases, thirty minute long segments annotated by experts have been used. Since we have focused on the discrimination between ventricular and normal beats, only beats labeled as 'V' or 'N' were selected for the classification purposes. The beats annotated as right ('R') or left ('L') bundle branch blocks ('BBB') were relabeled as 'N' since the annotations 'R' or 'L' represent morphology of the beat instead of the site of beat origin we are interested in.

Our final set contained 74413 'N' and 6954 'V' beats from MIT database; while from the AHA database 160325 'N' and 15253 'V' beats were chosen.

B. Preprocessing of the signal

Power line interference, high-frequency muscle artifacts and low-frequency drift were filtered during the pre-processing phase using methods similar to those described in [24]. Then the preprocessed signal was analyzed and the typical points on the ECG curve were measured out. For this task, well-known methods were applied as described in our previous work [25].

C. Features

We have computed 13 parameters that characterize the shape of each ECG beat as shown in Table I. Amplitude features represent maxima of amplitudes. Ratio features represent ratios of the amplitudes of the main deflections. Two well-known intervals – width of QRS complex and corrected QT interval (using Fridericia equation [26]) have been computed. Generally they discriminate pathological beats from the normal ones.

Finally, for qualitative description of P and T waves, terms like 'positive', 'negative-positive', 'positive-negative', 'negative' have been assigned. For QRS complex description additional R and S peaks had been searched for.

TABLE I
FEATURES USED FOR ECG BEAT DESCRIPTION

Amplitude of wave peaks	Ratios of peak amplitudes	Width of intervals	Description of the wave morphology
ampR	ratRT	intQRS	morphP
ampS	ratRS	intQTc	morphQRS
ampQ	ratQR		morphT
ampTpos			
ampTneg			

D. Training and testing sets

Each database has been divided into two sets – a training set and a testing set. The training set has been constructed using an even representation of 'N' and 'V' beats. From each of the two databases we have randomly selected twenty 'N' and twenty 'V' beats from each recording, if available. The testing set then consisted of the rest of each of the two databases.

IV. Results

Separate global evaluations have been performed on both databases with training sets as described in the previous section. The well-established sensitivity and specificity measures have been used for comparison of the different methods and the results are shown in Table II.

For classification of RBDT, fuzzy RBDT, and SOM methods that are primarily acting as clustering methods, template matching technique has been used. In the case of ANN, the algorithms have been implemented in WEKA [22]

SVM classifiers have been trained using different combinations of the hyperparameters C and σ, which has led to different values of sensitivity and specificity. The results of the SVM method shown in Table II are those with the most balanced performance and zero training error.

The results for each one of the seven methods for both data sets are summarized in Table II.

V. Conclusion and discussion

As can be seen in Table II, the results are mutually comparable and none classifier seems to outperform all the others.

The RBDT method, as well as the most of other presented methods, performs on the MIT database better than on the AHA database. The main advantage of the RBDT approach is its independence of the training set, because in fact it does not use any. The rules are based on a generalized knowledge and therefore they are not affected by the structure of the database. Moreover, the RBDT algorithm is very fast and its decisions are easy to explain; this is an advantage in comparison with other "black-box" methods.

Comparing the fuzzy RBDT to the classical RBDT, the fuzzy RBDT performs better on the MIT database, yet in the case of the AHA database the sensitivity and the specificity is 1% lower than the classical RBDT. It seems that after a more detailed analysis of the fuzzy RBDT, the performance

TABLE II
RESULTS FOR ALL USED METHODS

Method	MIT Sensitivity [%]	MIT Specificity [%]	AHA Sensitivity [%]	AHA Specificity [%]
RBDT	96.63	92.64	85.79	74.35
fuzzyRBDT	95.78	91.31	87.81	79.21
BP NN	93.12	76.13	86.86	77.06
GAME NN	41.59	94.04	49.57	95.41
RBF	89.88	75.94	72.28	87.16
SOM	97.95	35.41	97.00	35.21
SVM	94.10	94.01	93.80	91.47

can still be increased. The intention of fuzzy RBDT is to prepare a background for an automatic diagnosis based on fuzzy rules for a future work. Therefore, the fact that the performance is similar to RBDT can be considered as a good result.

GAME ANN is a new paradigm developed at the Czech Technical University. It often outperforms many data mining techniques [20]. In this particular case its performance is not satisfactory. Some testing records were classified with great accuracy over 90%. But other results are very bad, sometimes even worse than 50%. It seems to be a case of over-fitting of the GAME ANN even though the used configuration usually does not lead to overfitting.

BP, SOM and RBF ANNs are well known methods with many applications in pattern analysis [17]. SOM exhibits a very high sensitivity on both testing sets. However, the specificity is very low. The reason may be an overtraining as in the case of the underperforming GAME NN. Both BP and RBF NN achieve relatively high specificity with RBF performing slightly better in general and both of them seem to perform slightly worse compared to fuzzy RBDT.

The performance of the SVM method, even though it does not achieve the best overall performance in any of the four measurements, is rather good and balanced on both data sets but the time cost for training is quite high.

It is clear that the performance of most algorithms decreases significantly when using the AHA database. Due to the fact that the first database we employed was the MIT database, we have adjusted rules and features according to this database. Therefore, the features extracted were more "tuned" towards the description of the MIT database. It turned out that the AHA database is a bit different containing some cases that deviate from those in the MIT database. What is clearly needed towards better results is a normalization of the data sets. In our future work, we will try to extract features that are more insensitive to specific database in an attempt to produce a classification scheme that will be able to generalize well under any data set condition.

VI. REFERENCES

[1] N. J. Holter, "New Method for Heart Studies." *Science*, vol. 134, pp 1214–1220, Oct. 1961.

[2] P. de Chazal, O. O'Dwyer, and R.B. Reilly, "Automatic Classification of Heartbeats Using ECG Morp. Heartbeat Interval Features", *IEEE Trans. on Biom. Eng.*, vol. 51, No. 7, pp 1196-1206, 2004.

[3] D. Cuesta-Frau, J. C. Perez-Cortes, and G. Andreu-Garcıa, "Clustering of electrocardiograph signals in computer-aided Holter analysis," *Computer methods and programs in Biomedicine,* vol. 72, pp. 179-196, 2003.

[4] M. Lagerholm, C. Peterson, G. Braccini, L. Edenbrandt, and L. Sornmo, "Clustering ECG complexes using hermite functions and Self-organizing maps," *IEEE Trans. Biomed. Eng.,* vol. 47, pp. 838–848, 2000.

[5] A. L. Goldberger, L. Amaral, L. Glass, J. M. Hausdorf, P. C. Ivanov, R. G, Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley,. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation* vol. 101, no. 23, pp.e215-e220.

[6] AHA-DB. Aha database series 1. The American Heart Association, 1997.

[7] V. Chudáček, L. Lhotská, C. Stylios, G. Georgoulas and M. Staviař, "Comparison of Methods for Premature Ventricular Beat Detection In: Int. Special Topics." *ITAB Conference,* Piscataway: IEEE, 2006.

[8] G. J. Klir, and B. Yuan, *Fuzzy Sets and Fuzzy Logic : Theory and Applications*. Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

[9] H. T. Nguyen and E. A Walker, *A First Course in Fuzzy Logic*. CRC Press, 1997.

[10] L. A. Zadeh, "Fuzzy sets," *Inform. Control*, pp. 338–353, 1965.

[11] M. Petrik, and V. Chudacek, "Generalization of Rule-Based Decision Tree to Fuzzy," 2007 *IEEE Int. Workshop on Machine Learning for Signal Processing,* submitted for publication.

[12] T. Kohonen, *Self-Organizing Maps*. Springer, 1995.

[13] J. Vesanto, J. Himberg, E. Alhoniemi, and J. Parhankangas, "Self-organizing map in Matlab: the SOM Toolbox", TOOLMET, 2000.

[14] C. Chiu, T. Lin, and B. Liau, "Using Correlation Coefficient in ECG waveform for arrhythmia detection," *Biomedical Engineering applications, Basis and Communication*, Taiwan.

[15] C. J. C Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121-167, 1998.

[16] K. R. Muller, S, Mika, G. Ratsch, K. Tsuda, and B. Scholkopf, "An Introduction to Kernel-Based Learning Algorithms," *IEEE Trans. Neural Networks*, vol. 2, no. 2, pp. 181-201, 2001.

[17] S. Haykin, *Neural Networks: A Comprehensive Foundation*, NJ: Prentice Hall, 1999.

[18] H. Madala and A. Ivakhnenko. *Inductive Learning Algorithm for Complex System Modelling*. CRC Press, 1994.

[19] J. A. Muller, and F. Lemke. *Self-Organising Data Mining*. Berlin, 2000.

[20] P. Kordík Fully Automated Knowledge Extraction using Group of Adaptive Models Evolution, PhD thesis, CTU Prague, 2006.

[21] D. E Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533-536. 1986.

[22] I. Witten, and F. Eibe, *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, San Francisco, 2005.

[23] M. D Buhmann and M. J Ablowitz,. *Radial Basis Functions: Theory and Implementations*. Cambridge University, 2003.

[24] I. K. Daskalov, I. A Dotsinsky and I. I,Christov, "Developments in ECG acquisition, preprocessing, parameter measurement, and recording," *IEEE Eng. Med. Biol*. vol. 17, 1998.

[25] V. Chudacek, M. Hanuliak, and L. Lhotska, "Clustering of Heartbeats for Automated ECG Holter Analysis," *In proc. EMBEC05* [CD-ROM]. Praha: LS JEP, 2005, vol. 11.

[26] L. S. Fridericia, "Die Systolendauer im Elektrokardiogramm bei normalen Menschen und bei erzkranken," *Acta Medica Scandinavica*, Stockholm, vol. 57, pp. 469-486, 1920.