A comparison of ECG waveform features for the classification of normal and abnormal heartbeats

Emilien Le Flahat[1 2], Jean-Christophe Billard[1], Eric Plourde[2]

[1]Bitmakers, Paris, France
[2]Université de Sherbrooke, Sherbrooke, Canada

Abstract

This work investigates technics that allow for the automatic classification of normal vs abnormal heartbeats with the goal of assisting general practitioners. In fact, many different ECG waveform features have been proposed over the years as inputs to normal/abnormal heartbeat classifiers. However, there is a need for the formal comparison of the classification performances obtained when using these features, and more importantly their joint combinations, on a single common dataset. This study thus investigates the classification of heartbeats as normal or abnormal using combinations of 5 different types of features and 2 classifiers. Two different supervised classifiers were used: a Multilayer Perceptron (MLP) and a Support Vector Machine (SVM). The best feature set in terms of the accuracy of classification was found to be the combination of the Hermite basis function expansion, the complete higher order statistics of the ECG waveform and the RR intervals. In fact, a classification accuracy of 94.6% was obtained with the MLP for this feature set while a near perfect accuracy of 99.1% was obtained with the SVM.

1. Introduction

Electrocardiogram (ECG) is the most common medical examination in cardiology and is frequently practiced in general medicine. It consists in the recording of the electrical activity of the heart to diagnose cardiovascular disease (CVD). With 17,7 million death each year, CVDs are the first cause of death in the world (31% of all deaths worldwide) [1]. Among CVD, myocardial infarction, which is affecting 120 000 persons/year in France [2], remains a leading cause of death (with almost 16,000 deaths attributable to this condition in 2012) and the ECG is the first-line examination for the diagnosis of this condition. Nevertheless, general practitioners are facing a lack of experience and specificity when they interpret ECG waveforms. This research work investigates technics that allow the automatic classification of ECG waveforms as normal or abnormal.

The first step in an ECG classification approach is to extract relevant features of the ECG waveform, which are then used as inputs to normal/abnormal classifiers. The objective is to find features that will further improve the performances of these classifiers. Many approaches have been used over the years to extract ECG waveform features [3-11]. One of these is simply to use morphological characteristics, which are directly related to cardiac physiology, and exploit the heartbeat signal similarly to the general approach of a specialist. Some classification approaches have used only morphological features [3] whereas others have combined them with different features such as higher order statistics [5] or Hermite polynomials [4]. Cumulants, which are a type of higher order statistics, are an extension of second order technics such as autocorrelation. They can be directly applied on waveforms [6] or on alternative representations such as after applying a wavelet transform [7, 8]. Hermite basis function expansion of the QRS complex is another feature that has been proposed in the literature [9, 10]. Its purpose is to replace the QRS complex with Hermite polynomials and therefore use a reduced number of coefficients. Moreover, the previously mentioned features are often combined with features based on RR intervals.

Once the relevant features have been extracted, they are classified with supervised machine learning technics such as a support vector machine (SVM) [11] or an artificial neural network (ANN) [7]. These machine-learning techniques are typically able to classify many different pathologies simultaneously.

While several features have been proposed over the years as inputs to normal/abnormal heartbeat classifiers, there is still a need for the formal comparison of the classification performances obtained when using these features, and more importantly their joint combinations, on a single common dataset. This study thus investigates the classification of heartbeats as normal or abnormal using combinations of 5 different types of features and 2 classifiers. The best feature set in terms of the accuracy of classification was found to be the combination of the Hermite basis function expansion, the complete higher order statistics of the ECG waveform and the RR intervals. In fact, a classification accuracy of 94.6% was obtained with the ANN classifier (i.e. a multilayer perceptron

(MLP)), for this feature set while a near perfect accuracy of 99.1% was obtained with the SVM (normal: precision=98.7%, recall=99.5%, F1=99.1%; abnormal: precision=99.5%, recall=98.7%, F1=99.1%). This feature set classified with the SVM was thus found to outperform all other feature combinations studied.

2. Methodology

2.1. Database

This work makes use of the MIT-BIH (Massachusetts Institute of Technology – Beth Israel Hospital) arrhythmia database. In this database, each heartbeat is annotated on two leads and belongs to one of 16 classes, the most frequent ones are: Normal, Left Branch Block, Right Branch Block, Premature Ventricular Contraction and Paced. Waveforms were sampled at 360 Hz and only the MLII lead was used. Files 102 and 104 from MIT-BIH do not contain lead II, they were thus excluded. Heartbeats were downsampled in order not to favor a particular class. In other words, the number of beats from the over represented class (i.e. the normal class) was reduced such that the number of normal beats was equal to the number of abnormal ones. The total sum of beats from both classes was 71,860 beats.

Artefacts were removed from the ECG waveforms using Matlab's *wdencmp* denoising function with the symlet 4 wavelet. Moreover, low frequencies were removed using a Butterworth high-pass digital filter with a 0.7 Hz cut-off frequency.

2.2. Feature extraction

Five different types of features were used in this study mainly: morphological characteristics, higher order statistics on both the ECG directly as well as it wavelet representation, Hermite basis functions and RR intervals.

In the following, QRS detection was performed using *ecgpuwave* [12] from Physionet. This algorithm also performed the delineation on beats and returned the start and end of the QRS complexes as well as P and T waves. Moreover, a standardization followed by a normalization was performed on each feature. This is necessary since some features are negative such as P wave polarity. Standardization thus allows to rescale the data to have a mean of 0 and a standard deviation of 1 whereas normalization rescales the standardized data into a range of [0, 1]. These steps were applied on each signal independently.

A first set of features was extracted based on morphological characteristics of heartbeats [3]. Thirteen features related to the QRS complex were obtained for each beat. The first 10 of these features simply consisted in 10 samples between the QRS onset and the QRS offset, which informs about the morphology of the QRS. Then, the length and QRS area between the onset and the offset were also calculated. Finally, the QRS amplitude was calculated from the baseline according to CSE recommendations [13]. Five features related to the P wave were determined, namely: the length between its onset and its offset, the distance between the P offset and the QRS onset, the P area, its amplitude from the baseline and the sign of the P wave. Finally, 4 features were extracted from the T wave. The first of these features is a variable that indicates its type as a function of Physionet's morphological classification (i.e. normal, inverted, positive monophasic, negative monophasic, biphasic negative-positive, or biphasic positive-negative). The length between its onset and its offset were also calculated, as well as its area and finally its amplitude from the baseline.

A second set of features was extracted from each ECG beat using higher order statistics, namely cumulants [6]. More specifically, second, third and fourth orders were used. Five samples were extracted for each order for a total of 15 features.

Cumulants were also obtained from the wavelet (symlet 6) of each beat yielding a third set of features [7]. Only the three mid-band signals D3, D4 and D5 were considered from the wavelet decomposition. Cumulants were calculated with second, third and fourth orders on each subband. Consequently, nine cumulants were calculated for each beat from which were extracted the features. In total, 9 features were extracted from Normalized Summation which is defined as the summation of a cumulant divided by the area between the cumulant and zero, 9 features from the cumulant's variance, 9 features from the number of zero-crossing which allows to characterize the variations in the signal. Finally, 6 features were calculated with symmetry defined by the following equation where C_{ij} is the j^{th} cumulant of the D^i subband, and 2L the length of beats:

$$SYM_{ij} = \frac{\sum_{l=1}^{L} |C_{ij}(l) - C_{ij}(-l)|}{\sum_{l=-L}^{L} |C_{ij}(l)|}$$

These represent a total of 33 features.

A fourth set of features was extracted from the Hermite basis function expansion of the QRS complexes [9]. A window of 91 samples was used to perform the expansion with 45 samples before the R Peak and 45 samples after it. The mean between the first and the last sample was subtracted from the signal. Then, 45 zeros were added before and after the windows. The expansion was then calculated with a Singular Value Decomposition.

Finally, RR intervals are known to facilitate the detection of pathologies [3]. In this work, 4 features were extracted that are related to the RR intervals: RR interval, RR intervals of the previous beat, RR intervals of the next beat, the mean of RR intervals of 10 closest beats and the mean of RR intervals of the complete signal.

2.3. Classification

Two different classification approaches were used in this study: a MLP and an SVM.

The ANN used in this study is a standard feedforward MLP. The training is realized by backpropagation of the error along with the Adam optimization algorithm. The learning rate used is 0.003. The number of iterations, i.e. the number of times the artificial neural networks travels the database, is chosen as 200 and the batch size is 100. The activation function is ReLU (rectified linear unit) and the cost function used is the softmax cross entropy. In order to prevent for any overlearning, the number of neurons was reduced as a function of the number of inputs. Consequently, the number of neurones was chosen as being equal to the number of entries divided by 3 plus the number of outputs as per [14].

The SVM separates entry vectors into two classes with a maximum distance between them. The equilibrium between the maximization of this distance and the reduction of classification errors is controlled with a penalty parameter (C) and a Gamma parameter, which can be seen as the inverse of the radius of influence of samples selected by the model as support vectors. Values of Gamma = 1 and C = 1 were chosen following a cross-validation study, these values prevented overfitting. Moreover, the Gaussian kernel function was used here.

Training and test datasets were randomized and separated in 3. The first dataset was for training and contained 50,302 beats. The second one was for testing and contained 14,372 beats. A third dataset was used for further testing and is not described in this work.

3. Results

The mean accuracy of 10 classifications is presented in Table 1 for 10 different feature sets and for both normal and abnormal heartbeats, the MLP was used for classification. RR interval features were included in every feature set. The accuracy is the ratio of correctly predicted observations over the total observations. As mentioned previously, the number of neurons in the hidden layer was adjusted as a function of the number of features. The number of input layer neurons is equal to the number of inputs and the number of output layer neurons is 2.

A high number of neurons does not necessarily mean a high performance. In fact, the cumulants feature set with 19 features and 9 neurons has a better performance than morphological characteristics combined with cumulants applied on wavelets. Then, among feature sets with only one feature, i.e. that are not combined with one other feature or more, the cumulants provide the best score with an accuracy rate of 93.21%. Once features are combined together, some improvements are further observed. The best results are obtained with the Hermite basis function expansion of the QRS complex combined with the cumulants. Furthermore, Hermite basis function expansion of the QRS complex with morphological features also show good performances. However, combining both the cumulants obtained from temporal waveforms and the cumulants obtained from wavelets does not seem to show promising results.

Table 1. Mean accuracy of 10 classifications for 10 different feature sets for both normal and abnormal heartbeats with MLP. The RR interval features are included in every feature set.

Feature set (all include RR interval features)	Number of features	Neurons in hidden layer	Accuracy
Hermite	19	9	89.68
Cumulants	19	9	93.21
Cumulants on wavelets	37	15	85.70
Morphological	26	11	90.39
Cumulants & cumulants on wavelets	31	13	92.20
Hermite & cumulants	34	14	94.57
Hermite & morphological	41	16	93.56
Morphological & cumulants	41	16	94.10
Hermite & cumulants on wavelets	52	20	91.43
Morphological & cumulants on wavelets	59	22	92.18
All	87	31	94.01

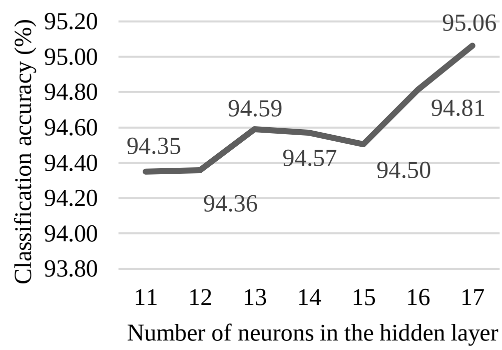

Figure 1. Classification accuracy as a function of the number of neurons in the hidden layer for Hermite and cumulants with a MLP.

Given the good performance of the Hermite features combined with cumulants, we further studied the classification performance by changing the number of neurons in the hidden layer (Figure 1). This feature set also has a lower number of input data compared to other sets which would mean a lower computational cost. Even though a larger number of neurons seems to give slightly better results, a smaller number would be preferable in order to limit overfitting. From Fig. 1, the best outcomes for a lower number of neurons was found to be 13 with a result of 94.59% and 11 with a result of 94.35%.

Table 2. Comparison of classification accuracy, precision and recall between MLP with 13 neurons and SVM with Gaussian kernel for both normal and abnormal heartbeats.

	SVM	MLP
Accuracy	99.09	94.57
Precision (normal)	98.54	94.97
Recall (normal)	99.49	94.50
Precision (abnormal)	99.47	94.14
Recall (abnormal)	98.68	94.78

Table 2 compares the performance of the SVM and the MLP on the Hermite expansion features combined with cumulants. The SVM outperforms the MLP for all measures. For example, a classification accuracy of 94.6% was obtained with the MLP for this feature set while a near perfect accuracy of 99.1% was obtained with the SVM (normal: precision=98.6%, recall=99.5%, F1=99.1%; abnormal: precision=99.5%, recall=98.7%, F1=99.1%).

4. Discussion and conclusion

Several aspects could affect the results of the study. For example, the detection and delineation of the ECG heartbeat was performed with Physionet. Like every algorithm, it is not 100% accurate and some errors could have occurred and affected the results. Moreover, standardization and normalization were performed with respect to each signal independently. However, other scaling could be possible, e.g. with respect to the total database. Finally, all of the feature sets do not have the same size. A hypothesis was made that the number of neurons necessary for a MLP is about the same as the number of entries data divided by 3 plus the number of the output layer. This hypothesis is not necessarily true but it allows to compare feature sets and avoids overlearning.

The best feature set in terms of the accuracy of classification was found to be the combination of the Hermite basis function expansion, the cumulants of the ECG waveform and the RR intervals. In fact, a classification accuracy of 94.6% was obtained with the MLP for this feature set while a near perfect accuracy of 99.1% was obtained with the SVM. This feature set classified with the SVM was thus found to outperform all other feature combinations studied.

References

[1] Global health estimates: deaths by cause, age, sex and country, 2000-2012. *Geneva : World Health Organization*; 2014

[2] Pinaire J, Azé J, Bringay S, Landais P. Infarctus du myocarde : quelles sont les trajectoires de soins pronostiques du décès à l'hôpital ? *Ingénierie des Connaissances*, 2017. p. 14-25.

[3] de Chazal P., O'Dwyer M., B. Reilly R. Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features. *IEEE Trans. on Biomed. Eng.*; Vol. 51, No. 7, 2004: 1196-1205.

[4] de Lannoy G., François D., Delbeke J. and Verleysen M. Weighted Conditional Random Fields for Supervised Interpatient Heartbeat Classification. *IEEE Trans. on Biomed. Eng.*; 2012: 241-247.

[5] Kutlu Y., Kuntalp D. A multi-stage automatic arrythmia recognition and classification system. *Computers in Biology and Medicine*; 2011: 37-45.

[6] Osowski S., Linh T. H. ECG Best Recognition Using Fuzzy Hybrid Neural Network. *IEEE Trans. on Biomed. Eng.*; 2001: 1265-1271.

[7] Yu S-N., Chen Y-H. H. Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. *Pattern Recognition Letters*; 2007: 1142-1150.

[8] Martis R. J., Acharya U. R., Chakraborty C. Application of higher order cumulants to ECG signals for the cardiac health diagnosis. *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society*: 2011.

[9] Osowski S., Linh T.H., Stodolski M. On-line heart beat recognition using Hermite polynomials and neuro-fuzzy network. *IEEE Transactions on Instrumentation and Measurement*; 2003: 1224 - 1231.

[10] Lagerholm M., Peterson C., Braccini G., Edenbrandt L., and Sörnmo L. Clustering ECG Complexes Using Hermite Functions and Self-Organizing Maps. *IEEE Trans. on Biomed. Eng.*; 2000: 838 - 848.

[11] Besrour R., Lachiri Z., Ellouze N. Clustering ECG Beat Classifier Using Support Vector Machine. *3rd International Conference on Information and Communication Technologies: From Theory to Applications*, Damascus, 2008: 1-5.

[12] Laguna P, Jané R, Caminal P. Automatic Detection of Wave Boundaries in Multilead ECG Signals: Validation with the CSE Database. *Computers and Biomedical Research*, vol. 27, no. 1, 1994: 45-60.

[13] The CSE Working Party. Recommendations for measurement standards in quantitative electrocardiography. *European Heart Journal*; 1985: 815-825.

[14] Jeff Heaton. *Introduction to neural networks with Java 2nd edition*; 2008: 158.

Address for correspondence:
Eric Plourde
2500 Université Blvd, Sherbrooke (Quebec), Canada J1K 2R1
eric.plourde@usherbrooke.ca