Analysis of 12-lead Classification Models for ECG Classification

M Llamedo[1,2,3], A Khawaja[4], JP Martínez[1,3]

[1]Aragon Inst of Eng Research, Univ of Zaragoza, Aragon, Spain
[2]Universidad Tecnológica Nacional, Buenos Aires, Argentina
[3]CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Spain
[4]Biosigna GmbH, Munich, Germany

Abstract

In this work we studied the improvement achieved by including information from the 12 ECG leads, in a previously developed classification model. This model includes features from the RR interval series and morphology descriptors calculated from the wavelet transform. The experiments were carried out in the INCART database, available in Physionet, and the generalization was corroborated in a private database. In both databases the AAMI recommendations for class labeling and results presentation were followed. Different approaches to integrate the additional information available in the 12-leads were studied. The best performing approach obtained for normal beats, sensitivity (S) 98%, positive predictive value (P^+) 94%; for supraventricular beats, S 88%, P^+ 91%; and for ventricular beats S 91%, P^+ 92%. The generalization capability was confirmed in a private database with comparable results. The performance of the reference two-lead classifier was improved by taking into account additional information from the 12-leads.

1. Introduction

Cardiovascular diseases are currently the biggest single cause of death in developed countries according to their public health agencies. The analysis of the electrocardiographic signal (ECG) provides a noninvasive and inexpensive technique to analyze the heart function for different cardiac conditions. One important analysis performed in the ECG is the classification of heartbeats, which is important for the study of arrhythmias.

Many algorithms for ECG heartbeats classification were developed in the last decades (see references in [1,2]) using the available two-lead databases. Some methodological key-points in the development of these classifiers allowed results comparison [1–3]. Probably the most relevant aspects were the fulfillment of AAMI recommendations [4], the patient-oriented data division [1] and the generalization capability of the classifier [2]. The room for improvement in the field of heartbeats classification, together with the availability of 3 and 12 leads holter devices makes necessary the development of algorithms capable of exploiting the increase of information. Besides, in the last years, the St. Petersburg Institute of Cardiological Technics 12-lead Arrhythmia Database (INCART) is freely available on Physionet [5], making possible the development multilead heartbeat classifiers.

The objective of this work is to develop and evaluate a 12-lead classifier, following the premises and results adopted in [2]: automatic classification, follow AAMI recommendations, use a simple classifier and robust features with physiological meaning.

2. Methods

2.1. ECG databases

In this work we used the INCART database. It consists of 75 annotated recordings extracted from 32 Holter records. Each record is 30 minutes long and contains 12 standard leads, each sampled at 257 Hz. The annotations were produced by an automatic algorithm and then corrected manually, containing over 175000 beat annotations in all. The original records were collected from patients undergoing tests for coronary artery disease (17 men and 15 women, aged 18-80; mean age: 58). None of the patients had pacemakers; most had ventricular ectopic beats. The database includes preferentially subjects whose ECG was consistent with ischemia, coronary artery disease, conduction abnormalities, and arrhythmias.

Besides, a second database was used to evaluate the generalization of the classifier. This is a private database developed at Biosigna GmbH, which consists of 56 recordings containing a broad set of pathologies. Each recording is one hour length, sampled at 500 Hz with an amplitude resolution of 410 increments per mV using a 12-bit ADC, allowing a range of 10mV approximately.

The AAMI Q class (unclassified and paced heart-

Table 1. Databases used in this work. Heart beats classes are N: normal, S: supraventricular, V: ventricular, F: fusion, and Q: unknown.

Database	N	S	V	F	Q	#Rec
INCART	153192	1957	19844	219	6	75
Biosigna	287554	1335	2572	0	0	56
Totals	440746	3292	22416	219	6	131

beats) was discarded since it is poorly represented in both databases. A similar limitation occurs with the fusion (F) AAMI class, but instead of discarding the heartbeats of this class, a class-labeling modification to the AAMI recommendation was adopted, as in [2]. It consists in merging fusion (of normal and ventricular beats) and ventricular classes, into a modified ventricular class denoted as V'. We will refer to this modification as AAMI2 labeling. The class distribution for both databases is summarized in Table 1.

2.2. Heartbeats classification: classifier and features

Under the assumption of independent and normally distributed data, the maximum *a posteriori* criterion (MAP) leads to the quadratic classifier defined by the discriminant functions

$$g_i(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^{\mathrm{T}}\Sigma_{\mathrm{i}}^{-1}\mathbf{x} + \boldsymbol{\mu}_{\mathbf{i}}^{\mathbf{T}}\Sigma_{\mathrm{i}}^{-1}\mathbf{x} - \frac{\mathbf{1}}{\mathbf{2}}\boldsymbol{\mu}_{\mathbf{i}}^{\mathbf{T}}\Sigma_{\mathrm{i}}^{-1}\boldsymbol{\mu}_{\mathbf{i}} - \frac{1}{2}\log(|\Sigma_i|) + \log(P(\boldsymbol{\omega}_i))$$

for the i-th class, where $\mathbf{x}$ represents the feature vector describing each heartbeat, and $\boldsymbol{\mu}_{\mathbf{i}}$, Σ_i and $P(\boldsymbol{\omega}_i)$ are the mean vector, covariance matrix and prior probability of the i-th class. The values of $\boldsymbol{\mu}_i$ and Σ_i were computed from the training data with the sample mean and covariance matrix expressions while the values for the prior probabilities $P(\boldsymbol{\omega}_i)$ were considered the same for all classes. The classification rule assigns $\mathbf{x}$ to the class i which results in the maximum posterior probability $g_i(\mathbf{x})$. In the case that the covariance matrix Σ is considered the same for all classes ($\Sigma_i = \Sigma_j, \forall i \neq j$), the quadratic discriminant classifier (QDC) becomes linear in $\mathbf{x}$ leading to the linear discriminant classifier (LDC) where Σ can be estimated as the weighted sample covariance

$$\boldsymbol{\Sigma} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{C}} \mathrm{w_i} \sum_{\mathrm{m}=1}^{\mathrm{M_i}} (\mathbf{x_m} - \boldsymbol{\mu_i}).(\mathbf{x_m} - \boldsymbol{\mu_i})^{\mathrm{T}}}{\sum_{\mathrm{i}=1}^{\mathrm{C}} \mathrm{w_i}.\mathrm{M_i}}.$$

The class-weighting possibility is of much interest due to the heavy class-size unbalance inherent to this application, where the normal class is in general one order of magnitude more represented that other classes. We refer as LDC to the linear classifier where $w_i = w_j, \forall i \neq j$, any other weight scheme will be referred as compensated linear classifier (LDC-C). In this work, all classification tasks were performed using and modifying the PRtools toolbox [6] for Matlab (The Mathworks Inc., Massachusetts).

Table 2. Features used in the model obtained in [2] only for two-lead recordings.

Feature	Description
$\ln(RR[i])$	Current RR interval
$\ln(RR[i+1])$	Next RR interval
$\ln(RR_1)$	Average RR interval in the last minute
$\ln(RR_{20})$	Average RR interval in the last 20 minutes
k_Z^x	Zero-cross position of the WT autocorrelation sequence in lead 1
k_Z^y	Zero-cross position of the WT autocorrelation sequence in lead 2
k_M^x	Maximum position of the WT autocorrelation sequence in lead 1
k_M^y	Maximum position of the WT autocorrelation sequence in lead 2

Following the results obtained in [2], where we developed a classification model with good generalization capabilities including rhythm and morphological features. As the rhythm features used in the model do not depend on the number of available leads, the first four features in Table 2 remain the same. Therefore we will focus the analysis to those features describing heartbeat morphology, which are the ones that can be improved by the new leads available.

The features k_Z^L and k_M^L described in [2], for each lead L were calculated in four sets of leads to study the best way of including the additional information. The first set includes the 12 standard leads (12L), while the three remaining sets are calculated from it. The second is the vectocardiogram (VCG) set, derived with the inverse Dower matrix from 12L. The following two sets are the result of projecting 12L and its fourth scale wavelet transform (WT), into the two most important basis of a principal component analysis (PCA). The PCA is performed for each beat in a segment including the QRS complex, defined 80 ms before and after the fiducial point (FP), for all available leads. The two most important components obtained are used as projection basis for all available leads. As a result a pair of ECG (or wavelet transformed) signals are obtained, which includes the most important information (in the variance sense) of a multilead set of leads. The resulting sets are the 12L-PCA and the WT-PCA respectively.

For the first three sets described (12L, VCG and 12L-PCA) we calculated the fourth scale of the WT for each lead in the set. For the fourth set, WT-PCA, the WT was calculated previous to the projection. Then for a segment which starts 130 ms before and ends 200 ms after the FP, the autocorrelation sequence ($r_L(k)$) in the fourth WT is calculated. The last step consists in detecting the first zero-crossing (k_Z^L), and the position of the first minimum (k_M^L) as shown in Figure 1, resulting in two features per lead.

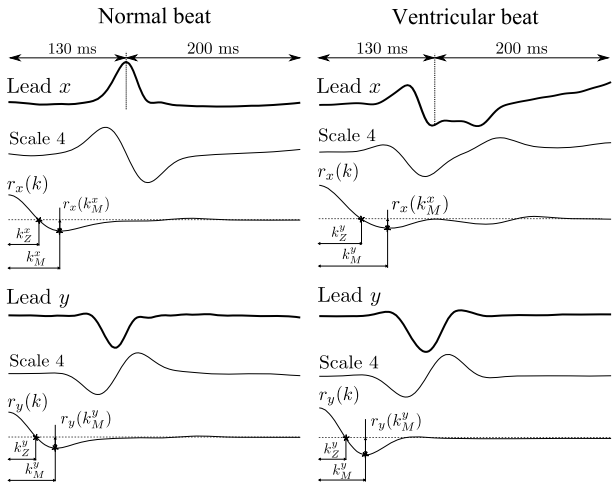

Figure 1. Illustration of the features calculated from the wavelet correlation signals for normal and ventricular beats. The autocorrelation signal of the QRS complex at scale 4 is shown for both leads (r_x and r_y). The zero-crossings and peaks of interest are indicated with an asterisk.

2.3. Experimental setup

The experiment consists in finding a classification model which increases the performance obtained by the reference two-lead model in [2], preserving the generalization capability to other databases. We studied the effect of adding one, two and all leads present for the four sets defined above. The class and global performances were calculated for each experiment by a k-fold crossvalidation, with $k = 10$ folds. It is important that each crossvalidation step implies training in $^9/_{10}$ of the database patients, and testing in the remaining $^1/_{10}$ of the patients. From the resulting (aggregated through all folds) confusion matrices, the performance estimates were calculated following AAMI recommendations [4]. The class imbalance, present in all public arrhythmia databases, is handled by scaling each row in the confusion matrix to sum the same. This results in the balanced performance calculation used in Table 4. Finally the performance for the best model found was also calculated in Biosigna database, to assess the generalization of the model.

3. Results

The results of the experiments described in the previous section are presented in Tables 3 and 4. The best model found, resulted from features calculated from both leads of the WT-PCA set.

4. Discussion and conclusions

In this work we have improved a two-lead heartbeat classifier by including the additional morphology information present in 12-lead recordings. We followed the concept of the morphology features obtained in [2], but calculating these features in different set of leads as reported in Table 3. It was shown that the WT-PCA set obtains the best improvements respect to the baseline classifier obtained in [2]. The selected model confirmed its generalization capability in the Biosigna database, where it also obtained better performance than the reference classifier, as shown in Table 4.

A patient-oriented 10-fold crossvalidation scheme was adopted to evaluate the performance, in order not to force future works to adopt a fixed training/test set to compare their results. This scheme provides also an acceptable bias/variance trade-off in the performance estimation. We also present the confusion matrices to ease the performance comparison, specially for the publicly available INCART database.

One advantage of the proposed approach is that it can be used for an arbitrary number of leads, because after the PCA we only retain the two most important components for the feature calculation. These components are calculated specifically for the QRS complex, and in the fourth scale of the WT, typically where the ECG presents maximum SNR. However in case of a large-scale artifact (such as lead disconnection) during the QRS complex, the PCA calculation would be corrupted, being this the larger limitation found for this approach. As an alternative model, only the first component of the PCA could be selected, resulting in a promising model that not only performed better than the original, but reduced the classification model in two features.

The performance improvement is however mild, maybe because the automatic classification approach is close to the performance limit achievable with the current classification model. Other techniques to improve the classification performance for supraventricular beats, like the classification of heartbeat sequences or the detection of the P wave are being studied by our group. Regarding to the ventricular class, techniques of morphology adaptation as described in [7] are also under development.

This results represents an improvement in performance respect of two-lead approach, concluding that the correct addition of new information to a classifier model improves its performance.

Acknowledgments

This work was supported by projects TEC-2007-68076-C02-02 from CICYT and GTC T-30 from DGA (Spain). The CIBER of Bioengineering, Biomaterials and

Table 3. Performance (in percentages) comparison between the several set of leads separating AAMI2 classes (N, S, V'). In bold the best set of leads found.

				Normal		Suprav.		Ventr.		Total		
Set of leads	Comments	Leads	# Features	S	P^+	S	P^+	S	P^+	A	S	P^+
12L	best lead	III	6	98	92	86	87	83	88	89	89	89
	Ref. model [2]	II-V1	8	97	93	87	87	84	89	90	90	90
	all leads	all	20	97	94	86	87	86	88	90	90	90
VCG	best lead	Y	6	98	93	83	83	81	85	87	87	87
		X Y Z	10	98	93	82	84	83	85	87	87	87
12L-PCA		1	6	98	93	87	87	86	90	90	90	90
		1-2	8	98	93	82	87	86	86	89	89	88
WT-PCA		1	6	99	93	86	90	89	90	91	91	91
	Selected model	**1-2**	**8**	**98**	**94**	**88**	**91**	**91**	**92**	**92**	**92**	**92**

Table 4. Performance comparison between the model suggested in this work and the reference classifier [2] separating AAMI2 classes (N, S, V'). First the confusion matrices for both databases are shown, and below the class and total performances are summarized. The performances are expressed in percentages.

INCART database

		Algorithm			
		n	s	v'	Total
Truth	N	150253	2403	536	153192
	S	81	1724	152	1957
	V'	465	1413	18275	20153
	Total	150799	5540	18963	175302

Biosigna database

		Algorithm			
		n	s	v'	Total
Truth	N	281341	5392	821	287554
	S	85	1211	39	1335
	V'	290	346	1936	2572
	Total	281716	6949	2796	291461

		Normal		Suprav.		Ventr.		Total		
Database	Classifier	S	P^+	S	P^+	S	P^+	A	S	P^+
INCART	This work	98	94	88	91	91	92	92	92	92
	[2]	97	93	87	87	84	89	90	90	90
Biosigna	This work	98	85	91	86	75	96	88	88	89
	[2]	97	84	89	82	70	93	85	85	86

Nanomedicine is an initiative of ISCIII.

References

[1] de Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2004;51:1196–1206.

[2] Llamedo M, Martínez J. Heartbeat classification using feature selection driven by database generalization criteria. IEEE Transactions on Biomedical Engineering In press;.

[3] Park K, Cho B, Lee D, Song S, Lee J. Hierarchical support vector machine. In Computers in Cardiology 2008, volume 35. IEEE Computer Society Press, 2008; 229–232.

[4] Testing and reporting performance results of cardiac rhythm and st-segment measurement algorithms. American National Standard, ANSI/AAMI/ISO EC57, 1998–(R)2008.

[5] Goldberger AL, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220.

[6] Duin R, Juszczak P, Paclik P, Pekalska E, deRidder D, Tax D, Verzakov S. Pr-tools, a matlab toolbox for pattern recognition, 2008. URL `http://www.prtools.org`.

[7] de Chazal P, Reilly RB. A patient-adapting heartbeat classifier using ecg morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2006; 53:2535–2543.

Address for correspondence:

Mariano Llamedo Soria, llamedom@unizar.es
Dept Ing. Electr. y Com. María de Luna 1, Edificio Ada Byron Lab. 2.05 – CP: 50018, Zaragoza, España.