

Weighted SVMs and Feature Relevance Assessment in Supervised Heart Beat Classification

Gael de Lannoy^{1,2}, Damien François¹, Jean Delbeke², and Michel Verleysen^{1,3}

¹ Université catholique de Louvain, Machine Learning Group
Place du levant 3, 1348 Louvain-La-Neuve, Belgium

² Université catholique de Louvain, Departement of physiology and pharmacology
av. Hippocrate 54, 1200 Bruxelles, Belgium

³ Université Paris I Panthéon-Sarbonne, SAMM
Rue de Tolbiac 90, F-75634 Paris Cedex 13, France
{gael.delannoy,damien.francois,jean.delbeke,
michel.verleysen}@uclouvain.be

Abstract. The diagnosis of cardiac dysfunctions requires the analysis of long-term ECG signal recordings, often containing hundreds to thousands of heart beats. In this work, automatic inter-patient classification of heart beats following AAMI guidelines is investigated. The prior of the normal class is by far larger than the other classes, and the classifier obtained by a standard SVM training is likely to act as the trivial acceptor. To avoid this inconvenience, a SVM classifier optimizing a convex approximation of the balanced classification rate rather than the standard accuracy is used. First, the assessment of feature sets previously proposed in the literature is investigated. Second, the performances of this SVM model is compared to those of previously reported inter-patient classification models. The results show that the choice of the features is of major importance, and that some previously reported feature sets do not serve the classification performances. Also, the weighted SVM model with the best feature set selection achieves results better than previously reported inter-patient models with features extracted only from the R spike annotations.

1 Introduction

The analysis of electrocardiogram (ECG) signals provides with critical information on the cardiac function of patients. Long-term recordings of ECG's are required, for example, for the clinical diagnosis of some disease conditions, and for the evaluation of new drugs during phase-one studies. Such long-term recordings are usually obtained using the popular Holter recorders.

Holter ambulatory systems record from 24 to 48 hours of heart activity, providing with data containing thousands of heart beats. The analysis is usually performed off-line by cardiologists, whose diagnosis may rely on just a few transient patterns. Because of the high number of beats to scrutinize, this task is very expensive and reliable visual inspection is difficult. Computer-aided classification of pathological beats is therefore of great importance to help physicians perform correct diagnoses.

Automating this task, however, is difficult in real situations for several reasons. First, several sources of noise pollute the ECG signal. Among these, power line interferences,

muscular artifacts, poor electrode contacts and baseline wandering due to respiration can sometimes be identified. Second, the vast majority of the heart beats are normal healthy beats while just only a small number of beats are pathological, and of course those are of uttermost importance. Third, artificial intelligence methods used to automate classification require the extraction of discriminative features from the ECG signals. Unfortunately, very few information is available to decide how to extract and build those features.

Several features characterizing the heart beats and several classification models have been investigated in the literature [6]. However, very few reported experiments follow the standards defined by the American Association for Medical Instrumentation (AAMI), which makes it very difficult to assess the relative merits of the methods and of the proposed extracted features [1]. Furthermore, critical issues, such as the class imbalance are often not taken specifically into account.

Another important aspect of heart beat classification is the way the training and test sets are composed. Most published methods for classifying the heart beats of a patient require access to previous data from that particular patient. In other words, data for each patient must be present both in training and test sets. We refer to this as “intra-patient” classification. Of course, one cannot always ensure that previous data is known and available for each patient when needed, especially new ones. Therefore, “inter-patient” classification – classifying the beats of a new tested patient according to a reference database built from data from other patients – is much more useful, but at the same time, much harder and demanding in terms of generalization capabilities of the models.

In this paper, inter-patient classification of heart beats following the AAMI guidelines is investigated. The class imbalance is taken into account by using a *weighted* support vector machine classifier. A large number of features extracted from the ECG signals described in the literature are combined and evaluated, and their relevance is assessed.

The remaining of this paper is organized as follows. Section 2 briefly reviews the state of the art in heart beat classification. Section 3 provides a short overview of the theoretical background for the mathematical methods used in this work. Section 4 describes the ECG database and Section 5 describes the features extracted from the heart beats. Finally, Section 6 shows the experiments and the results.

2 State of the Art

This section provides a short review of the state of the art in supervised heart beat classification. Two kinds of heart beat classification paradigms can be distinguished, corresponding to either *intra-patient* classification or *inter-patient* classification.

Inter-patient classification consists in classifying the beats of a new tested patient according to a reference database and a model built from data from other patients. This process thus implies generalization from one patient to another. As far as intra-patient classification is concerned, the reference database must contain previously labeled beats from the tested patient. The results that can be achieved are naturally better than when inter-patient classification is performed, but the patient labeled beats are usually not available in real situations. Furthermore, because pathological beats can be very rare,

there is no guarantee that the few training beats that would be labeled for this patient would contain representatives for each class; and the classifier could possibly fail in predicting something it has not learned.

Despite these major drawbacks, the majority of previously reported work is about intra-patient classification. Different models have been proposed for this task, including neural networks [13], k -nearest neighbors [5], hidden Markov models [4] and support vector machines (SVMs) [12]. A comprehensive review of intra-patient classification methods and their results can be found in [6].

As far as inter-patient classification is concerned, the first study to our knowledge to establish a reliable inter-patient classification methodology following AAMI standards is the work of [3]. A linear discriminant classifier model is trained and the results are evaluated on an independent test set. The unbalanced classes issue is addressed by introducing weights in the linear discriminant functions. Another linear discriminant model and a Mahalanobis distance classifier are used in [17] on the same dataset with different features and worst results. In [15], hierarchical SVMs are used to reduce the effect of the class unbalance on the same dataset.

The classification performances heavily rely on the extraction of relevant features from the heart beat time series. A variety of features have been proposed to characterize the heart beats. The representation of the heart beat signal by the coefficients of Hermite basis functions expansions is introduced for a clustering application in [11], and later used for classification by [14] and [15]. Another type of features that has been proposed is the representation of the heart beats by higher order statistics, and in particular the cumulants of order 2, 3 and 4 [13].

Other widely used groups of features are segmentation interval features and morphological features [5,3]. Morphological features correspond to the raw amplitude values of the ECG signal. Segmentation features require the annotation of the ECG characteristic points and then summarize the shape of the P, QRS and T waves by their duration, area, height, Q-T intervals, S-T intervals, etc. In most of previously reported works, the successive time differences between the R spikes of heart beats (later referred to as R-R interval features) are always combined to the other features. However, the intrinsic relevance of each type of features remains unknown. In this paper, this relevance is investigated using feature selection techniques [7,9].

3 Theoretical Background

This section provides a brief theoretical background on mathematical methods that are used in this work.

3.1 Weighted SVM Classifier

A support vector machine (SVM) is a supervised learning method introduced by Vapnik [18]. The two-class case is described here, because its extension to multiple classes is straightforward by applying the one-against-all or one-against-one methods. Let us first define the p dimensional feature vector $\mathbf{x}_i = \{x_i^1, x_i^2, \dots, x_i^p\}$ and the associated class value $y_i \in \{-1, 1\}$ for a given heart beat i with i ranging from 1 to N , N being the total number of heart beats in the dataset.

SVMs are linear machines that rely on a preprocessing to represent the features in a higher dimension, typically much higher than the original feature space. With an appropriate non-linear mapping $\varphi(\mathbf{x})$ to a sufficiently high-dimensional space, finite data from two categories can always be separated by a hyperplane. In SVMs, this hyperplane is chosen as the one with the largest margin.

Assume each observation \mathbf{x}_i has been transformed to $\mathbf{z}_i = \varphi(\mathbf{x}_i)$. The soft-margin formulation of the SVM allows examples to be misclassified or to lie inside the margin by the introduction of slack variables ξ_i in the problem constraints:

$$\text{Minimize } W(\mathbf{w}, b, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \cdot \Phi(\xi) \quad (1)$$

$$\text{s.t. } \begin{cases} y_i(\langle \mathbf{w}, \mathbf{z}_i \rangle + b) \geq 1 - \xi_i & \forall i = 1..N \\ \xi_i \geq 0 & \forall i = 1..N \end{cases} \quad (2)$$

where \mathbf{w} and b are the parameters of the hyperplane. The $\Phi()$ term introduced in the objective function is used to penalize solutions with many training errors. For any feasible solution (\mathbf{w}, b, ξ) , misclassified examples have an associated slack value ξ greater than 1. Hence, it seems natural to chose a function counting the number of slacks greater or equal to one as penalization function $\Phi()$ (the accuracy). Unfortunately, the optimization of such a function combined with the margin criterion is known to be NP hard. In practice, an approximation of the accuracy is thus commonly used:

$$\Phi(\xi) = \sum_{i=1}^n \xi_i \quad (3)$$

In this classical primal SVM formulation, the model is therefore actually optimizing a convex approximation of the accuracy [2].

However, in a heart beat classification task, around 90% of beats are normal beats and a dummy classifier which would always predict the normal class would get 90% accuracy. For this reason, the classification accuracy itself is a biased performance measure and the balanced classification rate (BCR) is to be preferred. In a two-class problem, the BCR is defined as the average between the sensitivity (SE) and the specificity (SP):

$$BCR = \frac{1}{2}(TP/(TP + FN) + TN/(TN + FP)) \quad (4)$$

$$= (SE + SP)/2 \quad (5)$$

where TP is the count of true positives, FN of false negatives, TN of true negatives and FP of false positives. In a multi-class problem, the BCR is equal to the average of class accuracies. Using this measure, a biased two-class classifier predicting always the normal class would only get 50% of BCR, correctly measuring that only one class has been correctly classified.

In the SVM formulation, different penalties for each class can be included in the objective function [10] so that a convex approximation of the BCR is optimized, rather than the accuray. Let N^+ and N^- denote respectively the number of positive and

negative examples. To eliminate the effect of the unequal class sizes on the classification error rate, the objective function is rewritten as:

$$\Phi(\xi) = \frac{N}{N_+} \sum_{\{i|y_i=1\}} \xi_i + \frac{N}{N_-} \sum_{\{i|y_i=-1\}} \xi_i \quad (6)$$

By constructing the Lagrangian, this *primal* formulation can be rewritten in a so-called *dual* form. The optimization is then achieved by solving the system using quadratic programming. For this type of optimization, there exist many effective learning algorithms. A common method is Platt's Sequential Minimal Optimization (SMO) algorithm, which breaks the problem down into 2-dimensional sub-problems that may be solved analytically, eliminating the need for a numerical optimization algorithm such as conjugate gradient methods [16].

In the dual form, the explicit form of the mapping function φ must not be known as long as the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)\varphi(\mathbf{x}_j)$ is defined. The kernel can for example be the linear kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^t \mathbf{x}_j$ or the radial basis function kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ where γ is a kernel parameter to be tuned.

3.2 Hermite Basis Functions

The representation of the heart beat signal via Hermite basis functions (HBF) was first introduced by [11] for a clustering application and later by [14] for classification. This approach exploits similarities between the shapes of HBF and typical ECG waveforms. Let us denote the heart beat signal by $x(t)$. Its expansion into a Hermite series of order N is written as

$$x(t) = \sum_{n=0}^{N-1} c_n \phi_n(t, \sigma) \quad (7)$$

where c_n are the expansion coefficients and σ is the width parameter. $\phi_n(t, \sigma)$ are the Hermite basis functions of the n th order defined as follows:

$$\phi_n(t, \sigma) = \frac{1}{\sqrt{\sigma 2^n n! \sqrt{\pi}}} e^{-t^2/2\sigma^2} H_n(t/\sigma) \quad (8)$$

where $H_n(t/\sigma)$ is the Hermite polynomial of the n th order. The Hermite polynomials satisfy the following recurrence relation:

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x) \quad (9)$$

with $H_0(x) = 1$ and $H_1(x) = 2x$.

The higher the order of the Hermite polynomial, the higher its frequency of changes in the time domain, and the better the capability of the expansion in Eq. 7 to reconstruct the signal [6]. The width parameter σ can be tuned to provide a good representation of beats with large differences in durations. The coefficients c_n of the HBF expansion can be estimated by minimizing the sum of squared errors using singular value decomposition and the pseudo-inverse technique. These coefficients summarize the shape of the heart beat signal and can be treated as the features used in the classification process.

3.3 Higher Order Statistics

The statistical properties of the heart beat signal can be represented by its higher order statistics (HOS). The cumulants of order two, three and four are usually used [13]. Assuming the heart beat signal $x(t)$ has a zero mean, its cumulant C_x^i of order i can be computed as follows:

$$\begin{aligned} C_x^2(\tau_1) &= E\{x(t)x(t + \tau_1)\} \\ C_x^3(\tau_1, \tau_2) &= E\{x(t)x(t + \tau_1)x(t + \tau_2)\} \\ C_x^4(\tau_1, \tau_2, \tau_3) &= E\{x(t)x(t + \tau_1)x(t + \tau_2)x(t + \tau_3)\} \\ &\quad - C_x^2(\tau_1)C_x^2(\tau_3 - \tau_2) \\ &\quad - C_x^2(\tau_2)C_x^2(\tau_3 - \tau_1) \\ &\quad - C_x^2(\tau_3)C_x^2(\tau_2 - \tau_1) \end{aligned}$$

where E is the expectation operator and τ_1, τ_2, τ_3 are the time lags.

4 ECG Database and Preprocessing

Data from the MIT-BIH arrhythmia database [8] are used in this work. The database contains 48 half-hour long ambulatory recordings obtained from 48 patients, for a total of approximately 110'000 heart beats manually labeled into 15 distinct types. Following the AAMI recommendations, the four recordings with paced beats are rejected.

The sampled ECG signals are first filtered to remove unwanted artifacts. The filtering procedure defined in [3] is used in this work. The ECG signal is filtered by two median filters. The first median filter is of 200 msec width and removes the QRS complexes and the P waves. The resulting signal is then processed with a second median filter of 600 msec width to remove the T waves. The signal resulting from the second filter operation contains the baseline wanderings and can be subtracted from the original signal. Powerline and other high frequency artifacts are then removed from the baseline corrected signal with a FIR filter.

The dataset configuration is the same as in [3,15]. The 44 available recordings are divided in two independent datasets of 22 recordings each with approximately the same ratio of heart beats classes. The first dataset is the training set, and is used to build the model. The second dataset is the test set, and is used to obtain an independent measure of the performances of the classifier.

The ECG characteristic points, corresponding to the onset and offset of P, QRS and T waves, are then annotated in each recording and the heart beats are separated using the standard *ecgpuwave*¹ segmentation software provided with the MIT-BIH arrhythmia database. The original R spike annotations are provided as input to the software to help the detection of the characteristic points. The beats whose Q and S points were not detected are considered as outliers and automatically rejected from our datasets.

The MIT-BIH heart beat labeled types are grouped according to the AAMI recommendations into four more clinically relevant heart beat classes [1] (see Table 2 for grouping details):

¹ See <http://www.physionet.org/physiotools/software-index.shtml>

N-class includes beats originating in the sinus node (normal and bundle branch block beat types);

S-class includes supraventricular ectopic beats;

V-class includes ventricular ectopic beats (VEBs);

F-class includes beats that result from fusing normal and VEBs.

Table 1 shows the number of beats in each class and their frequencies in the two datasets.

Table 1. Distribution of heart beat classes in the two independent datasets

	N	S	V	F	Total
Training	45801	938	3708	414	50861
	90.05%	1.84%	7.29%	0.81%	100%
Test	44202	1835	3204	388	49629
	89.06%	3.7%	6.46%	0.78%	100%

5 Feature Extraction

A large variety of features extracted from a heart beat signal have been proposed in the literature. They can be summarized in five feature groups: R-R intervals, segmentation intervals, morphological features, Hermite basis function expansion coefficients and higher order statistics. The features computed in this work for these different groups are now detailed.

1. R-R intervals: This group consists of three features built from the R-R interval series. The first three features are the R-R interval to the previous beat, the R-R interval to the next beat and the average of R-R intervals in a window of 10 surroundings beats.
2. Segmentation intervals: A large variety of 24 features are computed from the annotated characteristic points. These features include a boolean flag indicating the presence or absence of QRS, P and T waves. If the waves are present, their duration, maximum and minimum values, area, standard deviation, skewness and kurtosis are computed as features. The Q-T, S-T, Q-R, R-S intervals are also included. When the characteristic points needed to compute a feature failed to be detected in the heart beat annotation step, the feature value is set to the patient's mean feature value.
3. Morphological features: Ten features are derived by uniformly sampling the ECG amplitude in a window defined by the onset and offset of the QRS complex, and nine other features in a window defined by the QRS offset and the T-wave offset [3]. As the ECG signals were already sampled, linear interpolation was used to estimate the intermediate values of the ECG amplitude. Here again, when the onset or offset points needed to compute a feature were not detected, the feature value is set to the patient's mean feature value.
4. HBF coefficients: The parameters for computing the HBF expansion coefficients as defined in [15] are used. The order of the Hermite polynomial is set to 20, and the width parameter σ is estimated so as to minimize the reconstruction error for each beat.

5. Higher order statistics: The 2nd, 3rd and 4th order cumulant functions are computed. The parameters as defined in [14] are used: the lag parameters range from -250 msec to 250 msec centered on the R spike and 10 equally spaced sample points of each cumulant are used as features, for a total of 30 features.

Two additional normalized feature groups are also computed:

6. Normalized R-R intervals: These features correspond to the ratio between the previous three R-R values and their mean value for this patient. These last features are thus independent from the mean normal behavior of the heart of patients, which can naturally be very different between individuals, possibly misleading the classifier.
7. Normalized segmentation intervals: This group of features contains the same features as the previous segmentation group, but the values are normalized by their mean value for each patient. The normalization is obviously not applied to boolean segmentation features. Here again, the objective is to make each feature independent from the mean behavior of the heart of a patient, because it can naturally be very different between individuals.

6 Experiments and Results

In the following experiments, weighted SVMs with the *one-against-one* multi-class strategy are used. The training of the model is performed on the training dataset introduced in section 4, and this model is then applied to get a prediction of the class label of new heart beats from other patients in the test set.

Several types of kernels are evaluated in this study, and the linear kernel always outperforms the other kernels. As this is in accordance with previous works on heart beat classification using SVMs [15], only the results with the linear kernel are reported here. The regularization parameter C of the SVM model is optimized between 10^1 and 10^{-5} .

6.1 Feature Relevance Assessment

The first experiment aims at assessing the relevance of each set of features (cfr. Section 5) in an exhaustive, “wrapper”, approach: every possible combination of the seven feature sets are built, and each of them is fed to a weighted SVM model, whose performances are taken as relevance measure of the feature combination. Table 3 holds the most interesting results out of the $2^7 - 1 = 127$ configurations. Figure 1 shows how often (in percentage) each feature set is present in the top 20 best configurations.

Best overall performances are obtained with the combination of R-R features, normalized R-R features and higher order statistics with 83.55% of BCR. Using this selection, the lower class accuracy is of 78.53% for the V class. The addition of any other features to this selection always leads to a lower BCR. Also, these three feature sets are the three most selected sets in the top 20 configurations. In particular, the normalized R-R features are selected in every configuration of the top 15. It is interesting to note that when only R-R and normalized R-R feature sets are included in the model, a BCR of 80.07% can be reached. Also, when HOS are included alone, 76.94% of BCR is

Table 2. Grouping of the MIT-BIH labeled heart beat types according to the AAMI standards

Normal beats (N)	Supraventricular ectopic beats (S)	Ventricular beats (V)	ectopic	Fusion beats (F)
Normal beats	Atrial premature beat	Premature contraction	ventricular	Fusion of ventricular and normal beats
Left bundle branch block beats	Aberrated atrial premature beat	Ventricular beats	escape	
Right bundle branch block beats	Nodal (junctional) premature beats			
Atrial escape beats	Supraventricular premature beats			
Nodal (junctional) escape beats				

Table 3. Selection of the most interesting results out of the 127 configurations with the weighted SVM model (sorted in decreasing balanced classification rate)

	Feature sets						Results					
	RR	Norm-RR	Seg	Norm-Seg	HBF	HOS	Morph	N (%)	S (%)	V (%)	F (%)	BCR (%)
Top 6	•	•				•		80.00	88.07	78.53	87.63	83.55
						•		86.72	90.68	75.72	79.64	83.19
	•	•	•					77.68	89.70	86.99	77.84	83.05
					•	•		83.35	90.79	74.72	79.38	82.06
	•	•			•	•		80.58	79.24	78.21	89.43	81.87
Bottom 3	•	•						79.18	63.49	86.49	94.33	80.87
			•	•			•	80.00	17.28	72.78	3.35	43.35
					•		•	69.44	2.51	29.96	66.75	42.17
							•	42.76	72.97	37.73	12.63	41.52
Individuals						•		51.57	93.13	75.16	87.89	76.94
			•					68.19	78.69	72.00	64.95	70.96
	•							81.70	51.66	45.79	93.04	68.05
		•						83.25	64.96	84.46	1.29	58.49
				•				86.82	3.65	73.63	36.34	50.11
All					•			7.69	71.99	68.98	39.43	47.02
	•	•	•	•	•	•	•	91.57	19.56	84.58	57.73	63.36

obtained. These results indicate that R-R intervals, normalized R-R intervals and HOS features are the most important feature sets to include in the model.

The importance of feature selection is illustrated by the results obtained with the inclusion of all feature sets with a BCR of only 63.36% and an accuracy of 19.56% for the V class. Several feature sets do not seem to serve the classification performances, and some configurations can actually lead to a BCR below 50%, which is clearly not

appropriate. In particular, the normalization of the segmentation features with respect to each patient provides a lower accuracy than their non normalised version. Also, morphological features are not selected in any the top 20 configurations.

6.2 Performance Evaluation

The second experiment evaluates the performance of the weighted SVM with the best configuration in Table 3. It is compared to the results of previously reported work that also followed AAMI standards and inter-patient classification. In [3], the best results are obtained by the combination of two weighted linear discriminant (LD) classifier each built on 26 features extracted from one of the two leads of the holter recordings. In [15], a hierarchical model is built from three individual SVM classifiers. Each individual classifier is built on different features, including R-R intervals, HBF and HOS feature sets. The results of the raw SVM with the best configuration when no weights are defined, therefore optimizing an approximation of the accuracy rather than the BCR, are also displayed. Table 4 shows the results.

Table 4. Comparison of the weighted SVM using the best feature sets with previously reported works. The results when no weights are set in the SVM model (*raw SVM*) are also displayed.

	Results				
	N	S	V	F	BCR
Weighted SVM	80.0%	88.1%	78.5%	87.6%	83.6%
[3]	87.1%	76.0%	80.3%	89.4%	83.2%
[15]	86.2%	82.6%	80.8%	54.9%	76.1%
Raw SVM	96.5%	0.7%	77.84%	11.8%	46.7%

The weighted SVM model yields better overall results than the hierarchical SVM model. In particular, the results for the S and F pathological classes are significantly improved. The performances are slightly higher but very close the weighted LDA model. Nevertheless, the features included in our model only rely on R spike annotations and do not require the computation of segmentation interval features to reach the same level of performances. The detection of ECG characteristic points needed to extract these features can indeed be a difficult task in noisy signals.

The weights included in the SVM to take the class unbalance into account are also of major importance. If no weights are defined, the BCR obtained by the SVM model decreases to 46.7%, with an accuracy of only less than 1% for class S and of 11.8% for class F which is clearly unacceptable.

7 Conclusions

The classification of heart beats is of great importance for clinical applications involving the long-term monitoring of the cardiac function. The two main difficulties are the extraction of discriminative features from the heart beat time series and the class unbalance. For this purpose, this work has motivated and followed the use of:

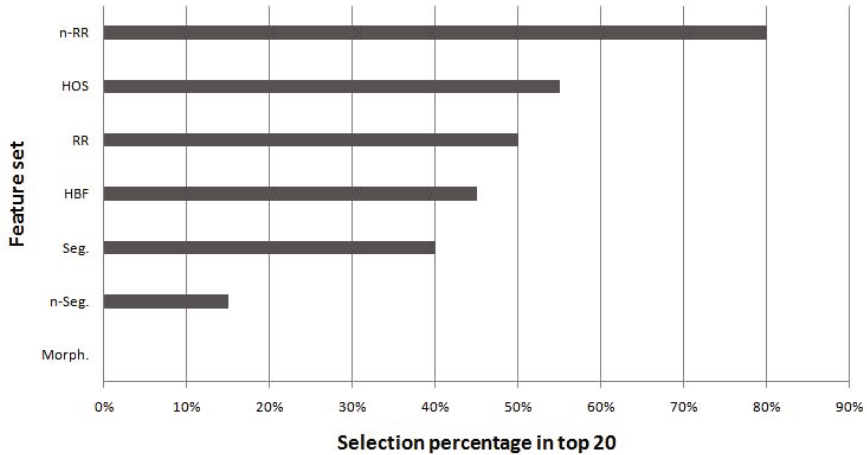


Fig. 1. Percentage of feature set selection in the top 20 best configurations

- AAMI guidelines for the establishment of reliable classifiers and for the evaluation of their relative merits;
- inter-patient rather than intra-patient classification;
- weighted multi-class SVM models to address the class unbalance problem;
- assessment of the relevance of usual feature sets;
- and the balanced classification rate as performance measure.

For the feature set relevance assessment, seven feature sets are considered: R-R intervals, segmentation features, HBF coefficients, higher-order statistics, morphological features, patient-normalized RR intervals and patient-normalized segmentation features. Best results are obtained with the combination of R-R intervals, normalized R-R intervals and high-order statistics with an average class accuracy of 83.55%. Any addition of features to these two sets leads to a lower performance. These results confirm that the choice of the features is a task of major importance, as a bad selection or too many features can lead to unacceptable results.

The results obtained with the weighted SVM model and the best selection of feature sets yield better results than previously reported inter-patient classification models. Furthermore, these performances are obtained with features extracted only from the R spike annotations. Therefore, the model does not require the annotation of the ECG characteristic points which can be very difficult to detect in noisy or pathological beats.

Another important issue for classification of heart beats resides in the class unbalance, which is met with weights being included in the SVM model. Indeed, the average accuracy obtained by the model with our best feature selection decreases from 83.6% to 46.7%, with an accuracy below 12% for class S and F when these weights are removed, leading to rather useless models that are unable to grasp the importance of the pathological cases.

Acknowledgements. G. de Lannoy is funded by a Belgian F.R.I.A. grant. This work was partly supported by the Belgian “Région Wallonne” ADVENS and DEEP projects.

References

1. Association for the Advancement of Medical Instrumentation, Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms. ANSI/AAMI EC38:1998 (1998)
2. Callut, J., Dupont, P.: F_β Support Vector Machines. In: Proceeding of International Joint Conference on Neural Networks, Montreal, Canada, July 31-August 4, pp. 1443–1448 (2005)
3. Chazal, P.D., O'Dwyer, M., Reilly, R.B.: Automatic classification of heartbeats using ecg morphology and heartbeat interval features. *IEEE Transactions on Biomedical Engineering* 51, 1196–1206 (2004)
4. Cheng, W., Chan, K.: Classification of electrocardiogram using hidden markov models. In: Proceedings of the 20th Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, vol. 1, pp. 143–146 (1998)
5. Christov, I., Gómez-Herrero, G., Krasteva, V., Jekova, I., Gotchev, A., Egiazarian, K.: Comparative study of morphological and time-frequency ecg descriptors for heartbeat classification. *Med. Eng. Phys.* 28(9), 876–887 (2006)
6. Clifford, G.D., Azuaje, F., McSharry, P.: *Advanced Methods And Tools for ECG Data Analysis*. Artech House, Inc., Norwood (2006)
7. François, D.: Feature selection. In: Wang, J. (ed.) *Encyclopedia of Data Mining and Warehousing*, 2nd edn. Information Science Reference (2008)
8. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P.C., Mark, R., Mietus, J., Moody, G., Peng, C.-K., Stanley, H.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101(23), e215–e220 (2000)
9. Guyon, I., Gunn, S., Nikravesh, M., Zadeh, L.A.: *Feature Extraction: Foundations and Applications (Studies in Fuzziness and Soft Computing)*. Springer, New York (2006)
10. Huang, Y., Du, S.: Proceeding of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, August 18-21, pp. 4365–4369 (2005)
11. Lagerholm, M., Peterson, C., Braccini, G., Edenbrandt, L., Sornmo, L.: Clustering ECG complexes using hermite functions and self-organizing maps. *IEEE Transactions on Biomedical Engineering* 47(7), 838–848 (2000)
12. Melgani, F., Bazi, Y.: Classification of electrocardiogram signals with support vector machines and particle swarm optimization. *IEEE Transactions on Information Technology in Biomedicine* 12(5), 667–677 (2008)
13. Osowski, S., Hoai, L.: Ecg beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering* 48(11), 1265–1271 (2001)
14. Osowski, S., Hoai, L., Markiewicz, T.: Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering* 51(4), 582–589 (2004)
15. Park, K., Cho, B., Lee, D., Song, S., Lee, J., Chee, Y., Kim, I., Kim, S.: Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function. In: *Computers in Cardiology*, pp. 229–232 (2008)
16. Platt, J.C.: Fast training of support vector machines using sequential minimal optimization. In: Scholkopf, B., Burges, C., Smola, A. (eds.) *Advances in Kernel Methods*. MIT Press, Cambridge (1999)
17. Soria, M., Martinez, J.P.: An ECG classification model based on multilead wavelet transform features. In: *Computers in Cardiology*, pp. 105–108 (2007)
18. Vapnik, V.N.: *The Nature of Statistical Learning Theory (Information Science and Statistics)*. Springer, Heidelberg (1999)