Noname manuscript No.
(will be inserted by the editor)

Robust Automated Cardiac Arrhythmia Detection in ECG Beat Signals

Victor Hugo C. de Albuquerque · Thiago M. Nunes · Danillo R. Pereira · Eduardo José da S. Luz · David Menotti · João P. Papa · João Manuel R. S. Tavares

the date of receipt and acceptance should be inserted later

Abstract Nowadays, millions of people are affected by heart diseases worldwide, whereas a considerable amount of them could be aided through an electrocardiogram (ECG) trace analysis, which involves the study of arrhythmia impacts on electrocardiogram patterns. In this work, we carried out the task of automatic arrhythmia detection in ECG patterns by means of supervised machine learning techniques, being the main contribution of this paper to introduce the Optimum-Path Forest (OPF) classifier to this context. We compared six distance metrics, six feature extraction algorithms and three classifiers in two variations of the same dataset, being the performance of the techniques compared in terms of effectiveness and efficiency. Although OPF revealed a

Victor Hugo C. de Albuquerque
Programa de Pós-Graduação em Informática Aplicada, Laboratório de Bioinformática, Universidade de Fortaleza, Fortaleza, CE, Brazil
E-mail: victor.albuquerque@unifor.br

Thiago M. Nunes
Centro de Ciências Tecnológicas Universidade de Fortaleza, Fortaleza, CE, Brazil
E-mail: tmnun@hotmail.com

Danillo R. Pereira and João P. Papa
Departamento de Ciência da Computação, Universidade Estadual Paulista, Bauru, São Paulo, Brazil
E-mail: dpereira@ic.unicamp.br and papa@fc.unesp.br

Eduardo José da S. Luz
Universidade Federal de Ouro Preto, Departamento de Computação, Ouro Preto-MG, Brazil
E-mail: eduluz@gmail.com

D. Menotti
Universidade Federal do Paraná, Departamento de Informática, Curitiba-PR, Brazil
E-mail: menotti@inf.ufpr.br

João Manuel R. S. Tavares (Corresponding author)
Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Departamento de Engenharia Mecânica, Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
E-mail: tavares@fe.up.pt

higher skill on generalizing data, the Support Vector Machines (SVM) based classifier presented the highest accuracy. However, OPF shown to be more efficient than SVM in terms of the computational time for both training and test phases.

Keywords ECG heart beats · Electrophysiological signals · Cardiac dysrhythmia classification · Feature extraction · Pattern recognition · Optimum-Path Forest

1 Introduction

The automatic detection and classification of arrhythmias in electrocardiography-based signals (ECG) has been widely studied in the last years in order to aid the diagnose of heart diseases. One way to perform this type of test is to conduct a long-time recording of the cardiac activity of an individual in his/her normal routine in order to obtain a reasonable amount of information about the individual's heartbeats. However, the posterior task of analysing such data may be tiresome and more prone to errors when interpreted by human beings, since there is a huge amount of information to be processed.

In order to cope with such problem, several works have been carried out arrhythmia classification in EEG signals by means of machine learning-oriented techniques [5, 14, 18, 15, 1]. However, regardless of the classification algorithm used, some processing steps are crucial to design a reasonable approach to detect arrhythmia. The quality of classification when dealing with ECG signals is directly dependent on the pre-processing phase, which aims at filtering noise frequencies that might interfere with ECG signal [20]. After preprocessing, it is required to detect and segment each heartbeat of the ECG signal. In order to perform this task, an important step is the detection of the QRS complex (three deflections from ECG signal), specifically the R wave, since most part of techniques for the detection and segmentation of heartbeats are based on the location of such deflection. Because of the steep angular coefficient and amplitude of the R wave, the QRS complex becomes more obvious than any other part of the ECG signal, being easier to be detected for later segmentation.

The final step is the classification of ECG signals, which is usually accomplished in a supervised fashion. Support Vector Machines (SVMs) [29, 27, 32, 7, 1, 8, 12, 6] and Artificial Neural Networks (ANNs) [9, 33, 34, 13, 11, 31, 23, 28, 21, 30] are among the most used machine learning techniques for this purpose. Other approaches such as Linear Discriminant Analysis [5] and a hybridization of Support Vector Machines and Artificial Neural Networks [10] are also applied for heartbeat classification. However, one of the main shortcomings related to the aforementioned pattern recognition techniques concerns with their parameters, which need to be fine-tuned prior to their application over the unseen samples (test set). SVMs are known due to their good skills on generalizing over test samples, but with the cost of having a high computational burden when learning the statistics of the training data, since each different kernel has its own parameters to be set up. ANNs are usually very fast for classifying samples, but its training step may be trapped in local optima, as well as it is not straightforward to choose a proper neural architecture.

Based on such assumptions, Papa et al. [26, 24] proposed the Optimum-Path Forest (OPF) classifier, which is a framework for designing classifiers based on graph partitions, being the samples (feature vectors) encoded by graph nodes and connected to each other by means of a predefined adjacency relation. A set of key nodes (prototypes) competes among themselves in order to conquer the remaining nodes offering to them optimum-path costs. This competition process generates a set of optimum-path trees rooted at each prototype node, meaning that a sample of a given tree is more strongly connected to its root than to any other in the forest.

The OPF classifier has gained considerable attention in the last years, since it has some advantages over traditional classifiers: (i) it is free of hard-to-calibrate control parameters; (ii) it does not assume any shape/separability of the feature space; (iii) it runs the training phase usually much faster; and (iv) it can take decisions based on global criteria. However, to the best of our knowledge, the OPF classifier has never been employed to aid the diagnosis of arrhythmias in heart rate by means of ECG signals so far. Therefore, the main contribution of this paper is to evaluate OPF effectiveness in ECG-based arrhythmia classification, being its results compared against some state-of-the-art pattern recognition techniques in terms of accuracy, computational time, sensitivity and specificity. Finally, another contribution of this work is to assess the performance of six different feature extraction methods in the aforementioned context, mainly: the approaches proposed by Chazal et al. [5], Güler and Übeyli [9], Song et al. [29], Yu and Chen [33], You and Chou [34], and Ye et al. [32].

2 Methodology

In this section, we describe the methodology employed in this work. Initially, the MIT-BIH (Massachusetts Institute of Technology - Beth Israel Hospital Boston) *Arrhythmia Database* [19] is described addressing considerations of ANSI/AAMI standard EC57 [3], which standardizes the evaluation of computational tools for the classification of cardiac arrhythmia datasets. After that, the feature extraction techniques used to generate the feature vectors are then described, followed by the description of the statistical parameters used to evaluate the performance of the classifiers under comparison.

2.1 MIT-BIH *Arrhythmia Database*

The MIT-BIH *Arrhythmia Database* is composed of signals from electrocardiography exams, being widely used to evaluate the performance of algorithms concerning the task of detecting arrhythmias [22]. The data consists of 48 records, 30 minutes-long, taken from 24 hours of ECG acquisition, being the samples obtained from two different channels. The signals were acquired from 47 patients between 1975 and 1979 at the Laboratory of Arrhythmia *Boston's Beth Israel Hospital*, which are aged between 23 and 89 years of which 22 females and 25 males. The analog records were digitized according to a sampling rate of $360Hz$, and the heartbeats marked and manually classified by experts in 15 classes regarding the type of arrhythmia. The types of arrhythmia identified in the database are indicated in Table 1.

Table 1 Types of heartbeats presented in the MIT-BIH database grouped according to AAMI Standard.

AAMI class	MIT-BIH original class	Type of beat
Normal (N)	N	Normal beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
	e	Atrial escape beat
	j	Nodal (junctional) escape beat
Supraventricular ectopic beat (S)	A	Atrial premature beat
	a	Aberrated atrial premature beat
	J	Nodal (junctional) premature beat
	S	Supraventricular premature beat
Ventricular ectopic beat (V)	V	Premature ventricular contraction
	E	Ventricular escape beat
Fusion beat (F)	F	Fusion of ventricular and normal beat
Unknown beat (Q)	/	Paced beat
	f	Fusion of paced and normal beat
	Q	Unclassifiable beat

Since the detection and segmentation of beats in ECG signals is not the main goal of this work, we have employed precomputed annotations of R waves provided by the database in order to accomplish the signal segmentation. In addition, 4 records derived from patients that make use of pacemakers that were discarded, following the recommendation of ANSI/AAMI standard EC57 [3], which also recommends to group the 15 classes reported in the database's annotations into 5 classes (Table 1). Figure 1 depicts some ECG signals for each class, being class Q represented by 10 signals, and the remaining ones represented by 100 signals. The signals were randomly picked up from the database.

2.2 Training and Test Set

The database was partitioned into two sets of records in order to separate the patients in training and testing groups. The composition of both sets was based on the study of Chazal et al. [5], which proposed to separate the patients by balancing each heartbeat class, as presented in Table 2. Besides the division of heartbeats into 5 classes as defined in [3], it was also considered the classification of heartbeats proposed by Llamedo and Martínez [15], which divided the 5 classes proposed in [3] into three main classes: N, S and V. Classes F and Q, which are less significant, were added to class V.

Fig. 1 MIT-BIH heartbeat signals grouped according to [3].

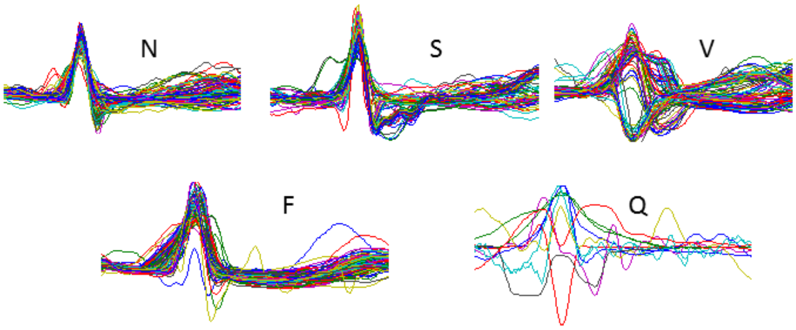

Table 2 Composition of the training and test sets according to Chazal et al. [5].

Set	Records
Training	101, 106, 108, 109, 112, 114, 115, 116, 118, 119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223 e 230
Test	100, 103, 105, 11, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233 e 234

2.3 Feature Extraction

Six feature extraction approaches (associated with Dataset A-F) were chosen based on the work of Luz and Menotti [16], which performed a comparison among some of the most used approaches for such purpose, mainly: Discrete Wavelet Transform (DWT), Independent Component Analysis (ICA), Principal Component Analysis (PCA), as well as information about RR range/interspace, which is the distance between peaks of two successive R waves in an ECG signal. For each dataset, the following methods were considered in this work:

– Dataset A - morphology of the signal and RR range [5];
– Dataset B - DWT [9];
– Dataset C - DWT [29];
– Dataset D - DWT, RR range and signal energy [33];
– Dataset E - DWT, ICA and RR range [34] and
– Dataset F - DWT, ICA, PCA and RR range [32].

The distribution of heartbeats by class and feature extraction approach considering the division of classes proposed by [3] is shown in Table 3, while Table 4 displays the same information considering the distribution into 3 classes proposed by [15]. In this Table T_b and n_f stand for the number of heartbeats of the set and the number of features extracted by each technique, respectively. One can noticed the variation in the number of beats among the methods concerns with the feature extraction techniques, that usually do not allow using the entire database. Samples located at the extremities of the signal, for instance, do not contain enough neighboring samples/segments to perform the proper feature extraction.

Table 3 Description of the experimental datasets according to AAMI classes [3].

	Dataset	Feature extraction	n_f	Hearbeat class					T_b
				N	S	V	F	Q	
Training	A	[5]	155	45,747	940	3,777	415	8	50,887
	B	[9]	19	45,845	943	3,788	415	8	50,999
	C	[29]	21	45,825	943	3,788	414	8	50,978
	D	[33]	13	45,844	943	3,788	415	8	50,998
	E	[34]	31	45,511	929	3,770	412	8	50,630
	F	[32]	100	45,844	943	3788	415	8	50,998
Test	A	[5]	155	44,181	1786	3,218	388	7	49580
	B	[9]	19	44,238	1836	3221	388	7	49690
	C	[29]	21	44,218	1836	3219	388	7	49,668
	D	[33]	13	44,238	1836	3221	388	7	49,690
	E	[34]	31	43,905	1823	3197	388	7	49,320
	F	[32]	100	44238	1836	3221	388	7	49,690

Table 4 Description of the experimental datasets according to [15].

	Dataset	Method	n_f	Heartbeat class			T_b
				N	$SVEB$	VEB	
Train	A	[5]	155	45,747	940	4,200	50,887
	B	[9]	19	45,845	943	4211	50,999
	C	[29]	21	45,825	943	4,210	50,978
	D	[33]	13	45,844	943	42,11	50,998
	E	[34]	31	45,511	929	4,190	50,630
	F	[32]	100	45,844	943	4,211	50,998
Test	A	[5]	155	44,181	1786	3,613	49580
	B	[9]	19	44,238	1,836	3616	49,690
	C	[29]	21	4,4218	1,836	3614	49,668
	D	[33]	13	44,238	1,836	3616	49,690
	E	[34]	31	43,905	1,823	3592	49,320
	F	[32]	100	44,238	1836	3,616	49,690

2.4 Optimum Path Forest classifier

Let's $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ be a λ-labeled dataset, where $\mathcal{D}_1$ and $\mathcal{D}_2$ denote the training and test sets, respectively. Let's $\mathcal{S} \subset \mathcal{D}_1$ be a set of prototypes of all classes (i.e., the key samples that best represent each samples class). The complete graph $(\mathcal{D}_1, A)$ is composed of nodes that represent samples in $\mathcal{D}_1$, and any pair of samples defines an edge in $A = \mathcal{D}_1 \times \mathcal{D}_1$ (Figure 2a)[1]. Additionally, let's $\pi_s =< s_1, s_2, \ldots, s_n, s >$ be a path with terminus at node $s \in \mathcal{D}_1$.

Roughly speaking, the OPF classifier contains two distinct phases, being the first one employed for training purposes, and the latter used to assess the robustness of the classifier designed in the previous phase. The training phase aims at building the optimum-path forest, and the test step classifies each

[1] The edges are weighted by the distance between their corresponding samples/nodes.

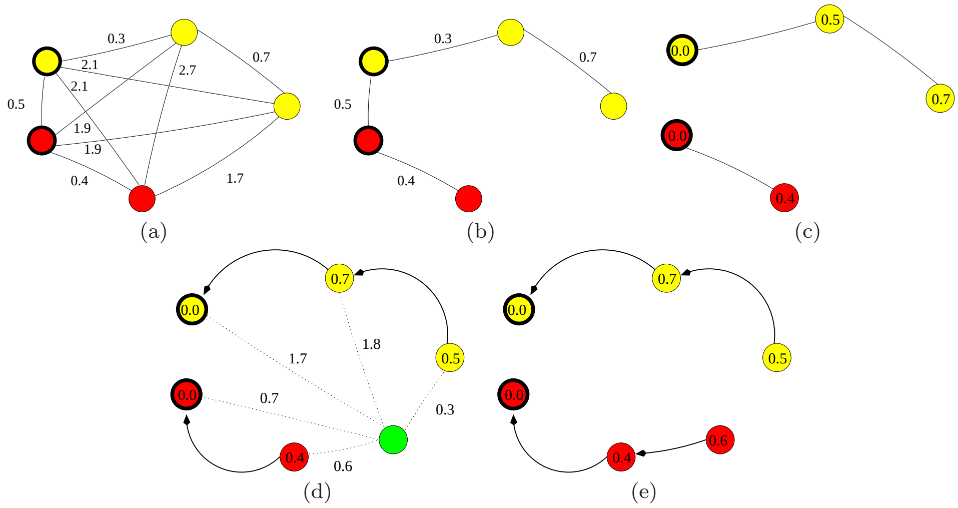

Fig. 2 (a) In the training step the training set is modeled as a complete graph, (b) a minimum spanning tree over the training set is computed (prototypes are highlighted), (c) optimum-path forest over the training set, (d) classification process of a test sample (in green), and (e) test sample classification.

test node individually, i.e. ,they are added to the training set for classification purposes only, and further removed.

2.4.1 Training step

$\mathcal{S}^*$ is an optimum set of prototypes when the OPF algorithm minimizes the classification errors for every $s \in \mathcal{D}_1$. Such set S^* can be found by the theoretical association between the minimum-spanning tree (MST) and the optimum-path tree for f_{max} [2]. Briefly, the training is the process of finding the S^* and an OPF classifier rooted at S^*. The MST in the complete graph $(\mathcal{D}_1, A)$ (Figure 2b) is represented by a connected acyclic graph whose nodes are all samples of $\mathcal{D}_1$, and the edges are undirected and weighted by the distances d between two adjacent samples. Every pair of samples is connected by a single path, which is minimum according to f_{max}. Hence, the minimum-spanning tree contains one optimum-path tree for any selected root node.

The optimum prototypes are the closest nodes of the MST with different labels in $\mathcal{D}_1$ (i.e., samples that fall in the frontier of the classes, as highlighted in Figure 2b). Removing the edges between different classes, their adjacent nodes become prototypes in S^*. The OPF algorithm can define an optimum-path forest with minimum classification errors in $\mathcal{D}_1$ (Figure 2c).

Soon after finding prototypes, the OPF algorithm is used, which essentially aims at minimizing the cost of every training sample. Such cost is computed using the f_{max} path-cost function, given by:

$$f_{max}(\langle s\rangle) = \begin{cases} 0 & \text{if } s \in S \\ +\infty & \text{otherwise,} \end{cases}$$
$$f_{max}(\pi_s \cdot \langle s,t\rangle) = \max\{f_{max}(\pi_s), d(s,t)\}, \quad (1)$$

where $\langle s\rangle$ is a trivial path, $\langle s,t\rangle$ is the arc between the adjacent nodes s and t such that $s,t \in \mathcal{D}_1$, $d(s,t)$ denotes the distance between nodes s and t, and $\pi_s \cdot \langle s,t\rangle$, is the concatenation of path π_s with the arc $\langle s,t\rangle$. One can note that $f_{max}(\pi_s)$ computes the maximum distance between adjacent samples in π_s when π_s is not a trivial path. Roughly speaking, the OPF algorithm aims at minimizing $f_{max}(\pi_t)$, $\forall t \in \mathcal{D}_1$.

2.4.2 Classification step

For any node $t \in \mathcal{D}_2$, we consider all edges connecting t with samples $s \in \mathcal{D}_1$, as though t were part of the training graph (Figure 2d). Considering all possible paths from S^* to t, OPF finds the optimum path $P^*(t)$ from S^* and labels t with the class $\lambda(R(t))$ of its most strongly connected prototype $R(t) \in S^*$ (Fig. 2e). This path can be identified incrementally evaluating the optimum cost $C(t)$:

$$C(t) = \min\{\max\{C(s), d(s,t)\}\}, \; \forall s \in D_1. \quad (2)$$

Let the node $s^* \in \mathcal{D}_1$ be the one that satisfies Equation 2 (i.e., the $P(t)$ in the optimum path $P^*(t)$). Given that $L(s^*) = \lambda(R(t))$, the classification simply assigns $L(s^*)$ as the class of t. An error occurs when $L(s^*) \neq \lambda(t)$.

3 Results and Discussion

In this section, we present the experimental results concerning the effectiveness and efficiency of each pair classifier/feature extraction technique employed in this work. First of all, the OPF classifier is evaluated considering six distance metrics: Euclidean, Chi-Square, Manhattan, Chi-Squared and Squared Bray-Curtis. After that, a comparison among OPF with the best metrics, Support Vector Machines with Radial Basis Function (SVM-RBF) and a Bayesian classifier (BC) is then presented.

3.1 Experimental Analysis of Optimum-Path Forest

In this section, we evaluate the performance and the computational time of the OPF classifier using six distance metrics[2]. The evaluation is performed considering the classification according to five [3] and three classes [15].

[2] For such purpose, we used the LibOPF library [25].

3.1.1 Five-class Problem

Here, we present the results considering the experimental dataset divided into five classes. Table 5 displays the recognition rates obtained by OPF using each distance metric[3] in the datasets defined by each feature extraction approach.

Table 5 OPF accuracy considering 5 classes. (The most accurate result is indicated in bold.)

Dataset	Euclidean [%]	Chi-Square [%]	Manhattan [%]
A	80.68	83.26	77.57
B	79.63	88.80	79.43
C	81.25	87.60	84.46
D	90.70	89.12	**91.21**
E	86.54	89.05	86.47
F	89.12	85.28	90.39

Set	Canberra [%]	Squared Chi-Squared [%]	Bray-Curtis [%]
A	77.93	76.14	79.81
B	80.51	80.61	87.69
C	84.90	82.63	76.55
D	90.88	90.75	88.90
E	86.53	86.62	81.79
F	86.60	85.70	78.41

We can observe that OPF with Manhattan distance obtained the best recognition rate with dataset D (91.21%), and that is approximately 0.35% higher than the second best result obtained with the Canberra distance metric (90.88%), as well as 0.5% higher than the result obtained with the Squared Chi-Squared metric (90.75%). Additionally, the results using dataset D were the best for all employed distances, suggesting that the method proposed by Yu and Chen [33] might be a good feature extractor to be used together with OPF. In addition to the recognition rate, we also computed the sensitivity (Se) and specificity (Sp), as well as the harmonic mean (H) of these two parameters (Table 6).

The best values of H considering class N were obtained using Canberra (0.78) and Squared Chi-Squared (0.78) distances and feature extractor C. The combination of Squared Chi-Squared metric and extractor C resulted in the best value of H for class S (0.60). In regard to classes V and F, Euclidean

[3] The recognition rates were computed using the standard formula, i.e., the ratio of the number of correct classifications by the number of database samples and H the harmonic mean between sensitivity and specificity.

Table 6 Specificity, sensitivity and their harmonic mean considering the OPF classifier and the AAMI five-classes categorization. (The best values for the harmonic mean are indicated in bold.) Notice the H, Se and Sp values are not divided by 100 due to the lack of space.

		Heartbeat classes				
		N	S	V	F	Q
Metrics	Dataset	H — Se — Sp	H — Se — Sp	H — Se — Sp	H — Se — Sp	H — Se — Sp
Euclidean	A	066 — 085 — 054	002 — 001 — 097	084 — 078 — 091	**055** — 038 — 097	000 — 000 — 100
	B	050 — 086 — 035	005 — 002 — 097	056 — 041 — 090	001 — 001 — 098	000 — 000 — 100
	C	074 — 085 — 066	031 — 018 — 095	084 — 078 — 091	014 — 007 — 097	000 — 000 — 100
	D	073 — 096 — 059	030 — 018 — 099	084 — 075 — 097	007 — 004 — 099	000 — 000 — 100
	E	065 — 092 — 050	006 — 003 — 098	076 — 062 — 097	029 — 017 — 097	000 — 000 — 100
	F	074 — 093 — 062	022 — 012 — 099	**091** — 086 — 097	031 — 018 — 097	000 — 000 — 100
Chi-Square	A	017 — 093 — 009	003 — 001 — 099	015 — 008 — 095	005 — 002 — 099	000 — 000 — 100
	B	002 — 100 — 001	000 — 000 — 100	000 — 000 — 100	001 — 001 — 100	000 — 000 — 100
	C	015 — 098 — 008	002 — 001 — 100	017 — 009 — 098	000 — 000 — 100	000 — 000 — 100
	D	004 — 100 — 002	000 — 000 — 100	004 — 002 — 100	000 — 000 — 100	000 — 000 — 100
	E	004 — 100 — 002	000 — 000 — 100	004 — 002 — 100	000 — 000 — 100	000 — 000 — 100
	F	012 — 095 — 006	002 — 001 — 099	011 — 006 — 097	001 — 001 — 100	000 — 000 — 100
Manhattan	A	066 — 081 — 056	003 — 002 — 095	086 — 082 — 090	051 — 035 — 096	000 — 000 — 100
	B	046 — 087 — 031	006 — 003 — 097	051 — 035 — 090	001 — 000 — 099	000 — 000 — 100
	C	076 — 088 — 066	031 — 018 — 097	085 — 078 — 093	013 — 007 — 098	000 — 000 — 100
	D	071 — 096 — 056	023 — 013 — 099	085 — 075 — 098	008 — 004 — 099	000 — 000 — 100
	E	064 — 092 — 048	006 — 003 — 098	075 — 061 — 097	023 — 013 — 097	000 — 000 — 100
	F	072 — 095 — 058	010 — 006 — 099	090 — 083 — 098	034 — 021 — 098	000 — 000 — 100
Canberra	A	071 — 081 — 063	047 — 031 — 098	080 — 073 — 087	025 — 015 — 096	000 — 000 — 100
	B	042 — 088 — 028	005 — 003 — 098	045 — 029 — 091	002 — 001 — 099	000 — 000 — 100
	C	**078** — 088 — 071	056 — 039 — 096	084 — 077 — 093	016 — 009 — 099	000 — 000 — 100
	D	071 — 096 — 057	026 — 015 — 099	085 — 075 — 098	011 — 006 — 099	000 — 000 — 100
	E	064 — 092 — 049	007 — 004 — 098	076 — 062 — 097	024 — 014 — 097	000 — 000 — 100
	F	064 — 092 — 049	007 — 004 — 099	083 — 071 — 098	009 — 005 — 095	000 — 000 — 100
Squared Chi-Sq.	A	065 — 080 — 056	006 — 003 — 097	081 — 076 — 086	037 — 023 — 097	000 — 000 — 100
	B	048 — 088 — 033	004 — 002 — 098	053 — 037 — 091	001 — 001 — 098	000 — 000 — 100
	C	**078** — 085 — 072	**060** — 044 — 095	085 — 078 — 093	022 — 012 — 097	000 — 000 — 100
	D	073 — 096 — 060	032 — 019 — 099	085 — 075 — 097	009 — 005 — 099	000 — 000 — 100
	E	065 — 092 — 050	006 — 003 — 098	076 — 062 — 097	027 — 016 — 097	000 — 000 — 100
	F	069 — 090 — 056	020 — 011 — 099	087 — 079 — 098	008 — 004 — 093	000 — 000 — 100
Bray-curtis	A	051 — 086 — 036	018 — 010 — 098	057 — 041 — 090	016 — 009 — 098	000 — 000 — 100
	B	005 — 098 — 002	000 — 000 — 100	000 — 000 — 098	000 — 000 — 100	000 — 000 — 100
	C	055 — 083 — 041	008 — 004 — 096	057 — 043 — 088	001 — 000 — 099	000 — 000 — 100
	D	002 — 100 — 001	001 — 000 — 100	000 — 000 — 100	000 — 000 — 100	000 — 000 — 100
	E	036 — 090 — 022	005 — 003 — 098	044 — 028 — 096	006 — 003 — 096	000 — 000 — 100
	F	053 — 085 — 039	003 — 002 — 096	054 — 038 — 091	021 — 012 — 098	000 — 000 — 100

distance has provided the best results with feature extractor F. As to class Q, OPF did not classify any sample properly due to the following main factors: the non-concentrated distribution of samples from that class, and the low representation of samples in the training and test sets ($\sim$ 0.00015% of the total number of samples).

However, a high recognition rate not always reflects a satisfactory performance in terms of classes separation, once that only class N (patient without cardiac arrhythmia) represents $\approx$ 90% of all dataset. For instance, let's consider the case of Chi-Square metric, which presented the best accuracy rates for feature extractor B (Table 5). The good results of such metric did not lead us to a satisfactory performance in terms of classes separation, since it presented low values for sensitivity and specificity for all classes, except for class N. This is due to the misclassification of most samples of classes S, V, F and Q, as belonging to class N, leading to a low harmonic mean (2%). In order to clarify this, the confusion matrix related to feature extractor B and Squared Chi-Square metric was built, Table 7. From the data obtained, one can verify that the dataset is dominated by class N, which clearly influenced all other classes. This can be confirmed by analyzing the results obtained for classes S, V, F and Q, that had the majority of the samples misclassified as being from class N (first column of Table 7). Also, it is important to stress that the accuracy calculated in this work do consider unbalanced datasets [26].

Table 7 Confusion matrix obtained for Chi-Square and feature extractor B.

		True class				
		N	S	V	F	Q
Predicted class	N	44,115	1,834	3,209	350	7
	S	27	0	2	4	0
	V	76	0	7	32	0
	F	19	2	3	2	0
	Q	1	0	0	0	0

3.1.2 Three-class Problem

We have also evaluated OPF considering the three-class dataset division proposed by Llamedo and Matínez [15], where classes F and Q are merged into class V. Table 8 presents the accuracy results obtained considering the three-class problem. Once again, the best result was obtained with Manhattan distance and feature extractor D (91.42%), as happened in the five-class problem (Table 5). Although some classes have been merged, we still have an umbalanced dataset. The aggregation of classes F and Q into class V has smoothed such problem, but class C still concentrates approximately 90% of the samples. Table 9 presents the results obtained in terms of sensitivity, specificity and harmonic mean.

Table 8 OPF accuracy considering three classes. (The most accurate result is indicated in bold.)

	Distance metric		
Dataset	Euclidean [%]	Chi-Square [%]	Manhattan [%]
A	81.00	83.41	77.82
B	80.43	88.89	80.18
C	81.41	87.68	84.61
D	90.92	89.13	**91.42**
E	86.81	89.08	86.80
F	89.46	85.35	90.78
	Distance metric		
Dataset	Canberra [%]	Squared Chi-Squared [%]	Bray-Curtis [%]
A	78.29	76.43	80.20
B	81.21	81.46	88.48
C	85.18	82.84	76.88
D	91.07	90.94	88.92
E	86.84	86.90	82.11
F	86.87	85.94	78.73

Considering class N, Canberra and Squared Chi-Squared distances together with the feature extractor C presented the best values for the harmonic mean (H) (0.78). Additionally, Squared Chi-Squared and the same feature extractor achieved the best result over class S. This may indicate that aggregation into 3 classes does not influence the measure H for classes N and S, the same values where obtained in the five-class problem (Table 6). In regard to class V, the best value ($H = 0.88$) was obtained with Euclidean and Manhattan distances over the feature extractor F. Therefore, the aggregation into three classes seemed to improve the results for the classes V, F and Q, which are now clustered into class V.

Table 9 Specificity, sensitivity and their harmonic mean considering the OPF classifier and the three-class categorization. (The best values for the harmonic mean are indicated in bold.)

		Heartbeat class		
		N	S	V
Metric	Dataset	H—Se—Sp	H—Se—Sp	H—Se—Sp
Euclidean	A	066 — 085 — 054	002 — 001 — 097	082 — 077 — 088
	B	050 — 086 — 035	005 — 002 — 097	062 — 047 — 090
	C	074 — 085 — 066	031 — 018 — 095	080 — 072 — 089
	D	073 — 096 — 059	030 — 018 — 099	081 — 070 — 096
	E	065 — 092 — 050	006 — 003 — 098	074 — 060 — 094
	F	074 — 093 — 062	022 — 012 — 099	**088** — 082 — 094
Chi-Square	A	016 — 093 — 009	003 — 001 — 099	016 — 009 — 094
	B	002 — 100 — 001	000 — 000 — 100	002 — 001 — 100
	C	015 — 098 — 008	002 — 001 — 100	016 — 009 — 098
	D	004 — 100 — 002	000 — 000 — 100	004 — 002 — 100
	E	004 — 100 — 002	000 — 000 — 100	004 — 002 — 100
	F	012 — 095 — 006	002 — 001 — 099	011 — 006 — 097
Manhattan	A	066 — 081 — 056	003 — 002 — 095	083 — 080 — 086
	B	046 — 087 — 031	006 — 003 — 097	056 — 041 — 090
	C	076 — 088 — 066	031 — 018 — 097	081 — 072 — 091
	D	071 — 096 — 056	023 — 013 — 099	081 — 070 — 097
	E	064 — 092 — 048	006 — 003 — 098	073 — 060 — 094
	F	072 — 095 — 058	011 — 006 — 099	**088** — 081 — 096
Canberra	A	071 — 081 — 063	047 — 031 — 098	077 — 072 — 083
	B	042 — 088 — 028	005 — 003 — 098	051 — 035 — 091
	C	**078** — 088 — 071	056 — 039 — 096	081 — 073 — 092
	D	071 — 096 — 057	026 — 015 — 099	081 — 070 — 097
	E	064 — 092 — 049	007 — 004 — 098	074 — 061 — 094
	F	064 — 092 — 049	007 — 004 — 099	078 — 068 — 093
Squared Chi-Sq.	A	066 — 080 — 056	006 — 003 — 097	078 — 074 — 083
	B	048 — 088 — 033	004 — 002 — 098	059 — 044 — 090
	C	**078** — 085 — 072	**060** — 044 — 095	081 — 074 — 090
	D	074 — 096 — 060	032 — 019 — 099	081 — 070 — 096
	E	065 — 093 — 050	006 — 003 — 098	074 — 061 — 094
	F	069 — 090 — 056	020 — 011 — 099	082 — 074 — 091
Bray-curtis	A	051 — 086 — 036	018 — 010 — 098	057 — 042 — 088
	B	001 — 099 — 001	000 — 000 — 100	001 — 000 — 099
	C	055 — 083 — 041	008 — 004 — 096	057 — 042 — 087
	D	002 — 100 — 001	001 — 000 — 100	000 — 000 — 100
	E	035 — 090 — 022	005 — 003 — 098	045 — 030 — 092
	F	053 — 085 — 039	003 — 002 — 096	055 — 039 — 089

Table 10 presents the OPF computational time (in seconds) for the training and test phases, being the fastest approaches the ones using Bray-Curtis and Manhattan metrics, since they are simpler to compute. It is important to highlight that these results are accompanied by a satisfactory classification performance, since OPF with Manhattan distance obtained generally very good classification results.

Table 10 OPF computational time (in seconds) considering the three-class problem. (Best values are indicated in bold.)

	Distance metrics		
	Euclidean	Chi-Square	Manhattan
	Training — Test — Total	Training — Test — Total	Training — Test — Total
A	445.18 — 682.89 — 1128.07	3230.23 — 1987.14 — 5217.37	335.07 — 584.79 — 919.86
B	145.53 — 173.48 — 319.01	416.65 — 6.57 — 423.21	54.50 — 94.53 — 149.03
C	142.70 — 177.63 — 320.34	464.50 — 102.25 — 566.74	55.19 — 97.24 — 152.43
D	128.90 — 132.96 — 261.86	299.52 — 2.74 — 302.26	40.09 — 53.07 — 93.15
E	172.76 — 181.13 — 353.89	650.34 — 14.66 — 665.00	80.58 — 108.42 — 189.00
F	317.86 — 444.96 — 762.82	2103.07 — 818.28 — 2921.35	219.07 — 360.48 — 579.55
	Distance metric		
	Canberra	Squared Chi-Squared	Bray-Curtis
	Training — Test — Total	Training — Test — Total	Training — Test — Total
A	1479.33 — 1588.78 — 3068.11	1478.00 — 1588.67 — 3066.67	1143.30 — 1168.50 — 2311.81
B	187.74 — 161.54 — 349.28	189.63 — 183.48 — 373.11	43.35 — **0.06** — 43.41
C	203.89 — 225.74 — 429.63	204.11 — 230.33 — 434.44	104.88 — 118.07 — 222.96
D	130.45 — 142.55 — 273.00	132.00 — 136.51 — 268.52	**35.22** — 6.01 — **41.22**
E	297.03 — 198.04 — 495.07	299.07 — 227.00 — 526.07	150.78 — 125.34 — 276.12
F	969.35 — 997.24 — 1966.59	9682.9 — 976.22 — 1944.50	431.96 — 519.69 — 951.66

3.1.3 Comparative Analysis of the Classifiers considering the five-class problem

In order to compare the performance of OPF over traditional classifiers (SVM-RBF[4] and Bayesian classifier), we considered only the two best distance metrics found in the previous section, i.e., Manhattan and Squared Chi-Squared distances. Therefore, we can summarize the techniques to be compared as follows:

- OPF-L1: OPF with Manhattan distance;
- OPF-SCS: OPF with Squared Chi-Squared distance;
- SVM-RBF: Support Vector Machines using RBF kernel[5];
- BC: Bayesian Classifier.

Table 11 shows the accuracy obtained for each feature extractor and classifier considering five classes of heartbeats. The most accurate technique was SVM-RBF with 94.09% of classification accuracy, followed by OPF-L1, BC and OPF-SCS, which obtained 91.21%, 90.95% and 90.75% of classification

[4] SVM parameters were optimized through cross-validation procedure.
[5] SVM implementation used was based on LIBSVM [4].

accuracies, respectively, considering the feature extractor D. Additionally, Table 12 presents the sensitivity, specificity and harmonic mean results.

Table 11 Accuracy rates obtained considering AAMI five classes. (The best accuracy value is indicated in bold.)

	Classifier			
Dataset	OPF-L1 [%]	OPF-SCS [%]	SVM-RBF [%]	BC [%]
A	77.57	76.14	88.21	80.69
B	79.43	80.61	84.06	79.52
C	84.46	82.63	89.82	81.37
D	91.21	90.75	**94.09**	90.95
E	86.47	86.62	87.06	86.82
F	90.39	85.70	87.12	89.14

Table 12 Harmonic mean, specificity and sensitivity obtained considering five classes and all classifiers. (The best values are indicated in bold.)

		Heartbeat class				
		N	S	V	F	Q
Metric	Dataset	H—Se—Sp	H—Se—Sp	H—Se—Sp	H—Se—Sp	H—Se—Sp
OPF-L1	A	066 — 081 — 056	003 — 002 — 095	086 — 082 — 090	051 — 035 — 096	000 — 000 — 100
	B	046 — 087 — 031	006 — 003 — 097	051 — 035 — 090	001 — 000 — 099	000 — 000 — 100
	C	076 — 088 — 066	031 — 018 — 097	085 — 078 — 093	013 — 007 — 098	000 — 000 — 100
	D	071 — 096 — 056	023 — 013 — 099	085 — 075 — 098	008 — 004 — 099	000 — 000 — 100
	E	064 — 092 — 048	006 — 003 — 098	075 — 061 — 097	023 — 013 — 097	000 — 000 — 100
	F	072 — 095 — 058	010 — 006 — 099	090 — 083 — 098	034 — 021 — 098	000 — 000 — 100
OPF-SCS	A	065 — 080 — 056	006 — 003 — 097	081 — 076 — 086	037 — 023 — 097	000 — 000 — 100
	B	048 — 088 — 033	004 — 002 — 098	053 — 037 — 091	001 — 001 — 098	000 — 000 — 100
	C	078 — 085 — 072	**060** — 044 — 095	085 — 078 — 093	022 — 012 — 097	000 — 000 — 100
	D	073 — 096 — 060	032 — 019 — 099	085 — 075 — 097	009 — 005 — 099	000 — 000 — 100
	E	065 — 092 — 050	006 — 003 — 098	076 — 062 — 097	027 — 016 — 097	000 — 000 — 100
	F	069 — 090 — 056	020 — 011 — 099	087 — 079 — 098	008 — 004 — 093	000 — 000 — 100
SVM-RBF	A	074 — 092 — 063	006 — 003 — 099	093 — 091 — 096	**082** — 072 — 097	000 — 000 — 100
	B	049 — 091 — 033	001 — 000 — 100	061 — 045 — 093	000 — 000 — 098	**008** — 005 — 100
	C	070 — 094 — 056	015 — 008 — 098	089 — 083 — 097	003 — 002 — 100	000 — 000 — 100
	D	**080** — 098 — 067	051 — 035 — 099	090 — 082 — 099	004 — 002 — 100	000 — 000 — 100
	E	070 — 092 — 057	012 — 007 — 099	089 — 082 — 098	001 — 001 — 094	000 — 000 — 100
	F	074 — 091 — 063	026 — 015 — 099	**093** — 089 — 096	030 — 018 — 096	000 — 000 — 100
BC	A	066 — 084 — 054	002 — 001 — 097	084 — 078 — 091	055 — 038 — 097	000 — 000 — 100
	B	050 — 086 — 035	005 — 002 — 097	056 — 041 — 090	001 — 001 — 099	000 — 000 — 100
	C	074 — 085 — 066	031 — 018 — 095	084 — 078 — 091	013 — 007 — 097	000 — 000 — 100
	D	073 — 096 — 059	029 — 017 — 099	085 — 076 — 097	007 — 004 — 099	000 — 000 — 100
	E	064 — 093 — 049	005 — 003 — 098	076 — 062 — 097	027 — 016 — 097	000 — 000 — 100
	F	074 — 093 — 062	021 — 012 — 099	091 — 086 — 097	031 — 018 — 097	000 — 000 — 100

From Table 12, one can realized that the best results in terms of harmonic mean were obtained for class N with SVM-RBF and feature extractor D (80.00%). This result is about 2% higher than the second best result obtained by OPF-SCS with feature extractor C (78.00%). In regard to class S, the best classifier was OPF-SCS using feature extractor C, followed by SVM-RBF with 51% of classification accuracy, which achieved the best recognition rates for classes V and F.

Table 13 displays the mean execution times considering the training, testing and total time (training+testing) required by each classifier[6]. The fastest classifier in the training phase was the BC in all datasets, followed by OPF-L1. The OPF-SCS was faster than SVM-RBF in all datasets as well, except for dataset F, where SVM-RBF had the third best time. The excessive times of SVM-RBF were due to the grid search that is necessary to fine-tune its parameters.

Table 13 Mean computational time (in seconds) required in the AAMI five-class problem. The standard deviation is also displayed. (The lowest times are indicated in bold.)

Dataset	Classifier											
	OPF-L1						OPF-SCS					
	train		test		total		train		test		total	
A	337.0	(1.6)	584.2	(0.7)	921.2	(1.8)	1487.2	(10.8)	1604.7	(12.4)	3091.9	(22.3)
B	54.8	(0.7)	102.9	(13.4)	157.7	(13.7)	191.9	(2.4)	181.2	(4.0)	373.1	(3.6)
C	55.5	(0.4)	95.1	(4.0)	150.6	(3.9)	206.9	(2.0)	242.5	(7.9)	449.4	(9.9)
D	40.3	(0.2)	53.3	(7.3)	**93.6**	(7.2)	132.4	(0.8)	131.3	(4.9)	263.7	(5.4)
E	81.1	(0.8)	115.1	(3.5)	196.2	(3.2)	302.2	(3.7)	223.8	(4.9)	525.9	(6.9)
F	220.4	(2.1)	380.3	(6.9)	600.8	(6.1)	974.9	(6.7)	990.0	(3.4)	1964.8	(9.9)
Dataset	Classifier											
	SVM-RBF						BC					
	train		teste		total		train		test		total	
A	2668.4	(26.2)	32.2	(6.7)	2700.6	(31.3)	62.9	(0.3)	1622.2	(8.6)	1685.2	(8.9)
B	576.0	(474.0)	12.6	(1.7)	588.6	(472.3)	11.1	(0.2)	236.0	(2.7)	247.1	(2.8)
C	195.3	(8.6)	6.9	(2.1)	202.2	(10.5)	11.9	(0.0)	253.8	(3.0)	265.7	(2.9)
D	170.7	(5.7)	**6.7**	(0.0)	177.4	(5.7)	**8.7**	(0.1)	173.0	(1.9)	181.7	(1.9)
E	546.5	(48.2)	9.4	(0.7)	555.9	(47.5)	15.6	(0.1)	354.1	(4.1)	369.6	(4.0)
F	608.7	(9.5)	15.3	(0.1)	624.0	(9.6)	42.0	(0.1)	1058.8	(9.2)	1100.8	(9.2)

In the test phase, the best computational time was obtained by SVM-RBF (6.7 seconds), being almost 8 times faster than OPF-L1 (53.3 seconds), both with feature extractor D. The third fastest technique was OPF-SCS (131.3 seconds) while BC, despite being the fastest in the training phase, took 173 seconds to classify the samples. In resume, SVM-RBF was the fastest in the classification phase, followed by OPF-L1, OPF-SCS and BC. Usually, SVM is fast for classifying samples, since it only considers the support vectors for such purpose, while OPF may need to evaluate a considerable number of training samples for that. However, if we consider the total time, OPF-L1 was the most efficient technique, which may lead us to consider it as a very suitable classifier concerning the trade-off between low computational time and high recognition rate.

Table 14 presents the confusion matrix related to SVM-RBF classifier in the five-class problem for the Dataset A [5]. It can be noted a confusion of class $SVEB$ with class N, where only 37 (2 %) samples were classified correctly for class $SVEB$. However, using the OPF-SCS classifier with Dataset C [29], the amount of samples correctly classified in the same class was around 43 %.

[6] We have executed all techniques 10 times for statistical purposes.

Thus, to detect Cardiac arrhythmia, also known ascardiac dysrhythmia or irregular heartbeat, the accuracy over class $SVEB$ is usually considered most important. As such, the OPF-SCS accuracy obtained for this class, which is much higher than the one of SVM-RBF, is of greater clinical relevance.

Table 14 Confusion matrices obtained for SVM-RBF and OPF-SCS classifiers.

SVM-RBF - Dataset A [5]						
		True Class				
		N	SVEB	VEB	F	Q
Classified as	N	40099	1705	221	83	3
	SVEB	361	37	3	0	0
	VEB	2285	23	2933	15	4
	F	1436	21	61	290	0
	Q	0	0	0	0	0
OPF-SCS - Dataset C [29]						
		True Class				
		N	SVEB	VEB	F	Q
Classified as	N	37677	612	591	322	2
	SVEB	2495	802	33	2	0
	VEB	2798	412	2514	17	5
	F	1245	9	81	47	0
	Q	3	2	0	0	0

3.1.4 Comparative Analysis of the Classifiers considering the three-class problem

In this section, we analyze the performance and computational time of all classifiers considering the three-class division proposed by [15]. Table 15 presents the recognition rates for each pair classifier/feature extractor method, being the sensitivity, specificity and harmonic mean results displayed in Table 16.

Table 15 Classification accuracy considering the three-class problem. (The best accuracy value is indicated in bold.)

	Classifier			
Dataset	OPF-L1	OPF-SCS	SVM-RBF	BC
A	77.82	76.43	80.01	80.98
B	80.18	81.46	84.29	80.31
C	84.61	82.84	90.01	81.53
D	91.42	90.94	**93.72**	91.17
E	86.80	86.90	88.45	87.07
F	90.78	85.94	83.66	89.47

In regard to class N, the SVM-RBF classifier has obtained the best harmonic mean value with feature extractor D, meanwhile OPF-SCS was the most accurate technique for class S using feature extractor C. These results are consistent with those obtained considering five classes. With respect to class V, three classifiers obtained the best harmonic mean values: SVM-RBC with feature extractor A, and OPF-L1 and BC with feature extractor F. However, in all these three cases, the values are followed by low sensitivity values for class S.

Table 16 Harmonic mean, specificity and sensitivity considering three classes and all classifiers. (The best values are indicated in bold.)

		Heartbeat class		
		N	S	V
Metric	Dataset	H—Se—Sp	H—Se—Sp	H—Se—Sp
OPF-L1	A	066 — 081 — 056	003 — 002 — 095	083 — 080 — 086
	B	046 — 087 — 031	006 — 003 — 097	056 — 041 — 090
	C	076 — 088 — 066	031 — 018 — 097	081 — 072 — 091
	D	071 — 096 — 056	023 — 013 — 099	081 — 070 — 097
	E	064 — 092 — 048	006 — 003 — 098	073 — 060 — 094
	F	072 — 095 — 058	011 — 006 — 099	**088** — 081 — 096
OPF-SCS	A	066 — 080 — 056	006 — 003 — 097	078 — 074 — 083
	B	048 — 088 — 033	004 — 002 — 098	059 — 044 — 090
	C	078 — 085 — 072	**060** — 044 — 095	081 — 074 — 090
	D	074 — 096 — 060	032 — 019 — 099	081 — 070 — 096
	E	065 — 093 — 050	006 — 003 — 098	074 — 061 — 094
	F	069 — 090 — 056	020 — 011 — 099	082 — 074 — 091
SVM-RBF	A	072 — 082 — 065	007 — 004 — 098	**088** — 092 — 084
	B	053 — 090 — 038	001 — 000 — 100	069 — 056 — 091
	C	070 — 095 — 056	012 — 007 — 099	083 — 074 — 095
	D	**080** — 098 — 067	052 — 035 — 099	084 — 074 — 098
	E	071 — 093 — 057	012 — 006 — 099	086 — 080 — 093
	F	072 — 087 — 062	023 — 013 — 099	086 — 085 — 088
BC	A	066 — 085 — 054	002 — 001 — 097	082 — 077 — 088
	B	050 — 086 — 035	005 — 002 — 097	062 — 047 — 089
	C	074 — 085 — 066	031 — 018 — 095	080 — 072 — 089
	D	073 — 096 — 059	029 — 017 — 099	082 — 071 — 096
	E	065 — 093 — 049	005 — 003 — 098	074 — 060 — 094
	F	074 — 093 — 062	021 — 012 — 099	**088** — 083 — 094

Table 17 presents the mean computational time in seconds concerning all techniques. Once again, the lowest computational time for training was achieved by BC and followed by OPF-L1 for all datasets. Except for feature extractors C and F, where OPF took longer to train, SVM-RBF classifier was the most costly technique for training the samples. Relatively to the five-class problem, similar computational times could be observed for BC and OPF-based classifiers, evidencing the robustness of these classifiers when dealing with different number of classes. As expected, the SVM computational time decreased, since we have less classes to be analyzed during the pair-wise comparison against them[7]. Last but not least, SVM-RBF was the fastest technique for the classification phase, while OPF-L1 obtained the lowest execution time considering both training and test phases.

[7] LIBSVM implements the one-against-one method for multi-class tasks.

Table 17 Mean computational time (in seconds) considering the three-class problem. The standard deviation is also displayed. (The lowest times are indicated in bold.)

Dataset	Classifier											
	OPF-L1						OPF-SCS					
	train		test		total		train		test		total	
A	336.0	(1.7)	575.2	(11.5)	911.2	(9.9)	1486.1	(7.1)	1597.9	(8.1)	3084.1	(15.1)
B	54.5	(0.2)	93.0	(2.8)	147.5	(3.0)	191.2	(1.7)	192.3	(7.8)	383.4	(9.4)
C	55.5	(0.3)	92.3	(9.2)	147.7	(9.3)	206.2	(2.0)	233.5	(12.7)	439.7	(14.1)
D	40.2	(0.2)	52.2	(5.2)	**92.4**	(5.4)	132.6	(0.5)	134.1	(3.8)	266.7	(3.5)
E	81.0	(0.4)	105.9	(4.2)	186.9	(4.0)	301.6	(2.6)	224.8	(4.2)	526.4	(2.2)
F	221.7	(2.7)	368.8	(11.6)	590.5	(14.1)	973.4	(4.7)	977.3	(2.8)	1950.7	(6.8)

Dataset	Classifier											
	SVM-RBF						BC					
	train		test		total		train		test		total	
A	2069.7	(23.7)	25.9	(5.4)	2095.6	(28.6)	62.5	(0.1)	971.2	(5.2)	1033.7	(5.3)
B	280.5	(2.9)	11.6	(0.7)	292.2	(3.6)	11.0	(0.1)	141.5	(0.7)	152.5	(0.7)
C	194.8	(2.7)	7.7	(0.0)	202.5	(2.7)	11.7	(0.1)	152.9	(0.3)	164.6	(0.4)
D	165.9	(1.0)	**6.0**	(0.0)	171.9	(0.9)	**8.7**	(0.1)	105.4	(1.4)	114.1	(1.5)
E	536.1	(58.9)	9.1	(0.8)	545.2	(58.1)	15.5	(0.1)	213.5	(2.2)	229.0	(2.3)
F	553.3	(26.2)	17.0	(6.8)	570.3	(32.9)	41.6	(0.3)	638.5	(5.1)	680.1	(5.3)

Also, based on a similar analysis to the one carried out with the data in Table 14, it could be confirmed that also in the three-classes problem OPF-SCS is the most appropriate to identify the pathological classes, i.e., the ones with greater clinical interest. Luz et al. [17] considered only the Euclidean metric, and obtaining highest accuracy rates of 90.7% and 90.9% in the 3- and 5-classe problems, respectively, considering in both cases the extraction method proposed by [34]. However, the present work could improve the accuracy of OPF with Manhattan distance, obtaining 91.42 and 91.21% in the 3- and 5-classe problems, respectively, for the same dataset and with computational time inferior to the one achieved by Luz et al. [17]. This considerable increase in accuracy directly leads to a more accurate detection of pathological classes. As such, it is possible to identify more precisely a cardiac arrhythmia with the Manhattan distance than with Euclidean one. Again, it should be stressed that the aforementioned classes are of great importance for clinical analysis, and that the SVM classifier could not detect accurately enough the samples of these classes.

4 Conclusions and Future Works

In this paper, a detailed study about the performance and computational time of supervised classification algorithms regarding the task of arrhythmia detection in ECG signals was presented. The main contributions of this work are: (i) to evaluate the OPF classifier in the task of arrhythmia detection, (ii) to evaluate six distances with OPF, among which the best accuracy rates were obtained by the Manhattan metric, while better generalization (i.e., the accuracy achieved per class) was attained using *Square Chi-Square* distance, (iii) to test six feature extraction techniques and investigate which one leads to better recognition rates and generalization, (iv) to compare OPF against Support Vector Machines and a Bayesian classifier, being found that OPF was the less generalist, while the SVM classifier was the most accurate, and, finally, (v) to find that OPF achieved the best trade-off between computational load and recognition rate.

Being OPF less generalist with respect to classes V and S, which are of great clinical significance regarding to class N, one can concluded that this classifier is more appropriate for the classification of arrhythmias in ECG signals that the SVM and Bayesian classifiers.

Since we observed that OPF and SVM-RBF were the most accurate classifiers, our future works will be guided to explore the synergy between these classifiers in order to build an ensemble of classifiers aiming at increasing the recognition rate of arrhythmia detection in ECG signals, as well as to evaluate other traditional and most recent feature extraction methods.

Acknowledgments

The first author thanks the Brazilian National Council for Research and Development (CNPq) for providing financial support through grants # 470501/2013-8 and # 301928/2014-2.

The sixth author is grateful to CNPq grants #306166/2014-3 and #470571/2013-6, as well as to São Paulo Research Foundation (FAPESP) grant #2014/16250-9.

The last author gratefully acknowledge the funding of Project NORTE-01-0145-FEDER-000022 - SciTech - Science and Technology for Competitive and Sustainable Industries, cofinanced by “Programa Operacional Regional do Norte (NORTE2020)”, through “Fundo Europeu de Desenvolvimento Regional (FEDER)”.

References

1. Abawajy, J.H., Kelarev, A.V., Chowdhury, M.: Multistage approach for clustering and classification of ECG data. Computer Methods and Programs in Biomedicine **112**, 720–730 (2013)
2. Allène, C., Audibert, J.Y., Couprie, M., Keriven, R.: Some links between extremum spanning forests, watersheds and min-cuts. Image and Vision Computing **28**(10), 1460–1471 (2010)
3. ANSI/AAMI: Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms. Association for the Advancement of Medical Instrumentation -AAMI / American National Standards Institute, Inc.-ANSI (2008). ANSI/AAMI/ISO EC57, 1998-(R)2008
4. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2**, 27:1–27:27 (2011). Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`

5. Chazal, P., O'Dwyer, M., Reilly, R.B.: Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering **51**(7), 1196–1206 (2004)
6. Daamouchea, A., Hamamib, L., Alajlanc, N., Melgani, F.: A wavelet optimization approach for ECG signal classification. Biomedical Signal Processing and Control **7**(4), 342–349 (2012)
7. de Lannoy, G., François, D., Delbeke, J., Verleysen, M.: Weighted SVMs and feature relevance assessment in supervised heart beat classification. In: Biomedical Engineering Systems and Technologies (BIOSTEC), pp. 212–223 (2010)
8. Dutta, S., Chatterjee, A., Munshi, S.: Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification. Medical Engineering & Physics **32**(10), 1161–1169 (2010)
9. Güler, I., Übeyli, E.D.: ECG beat classifier designed by combined neural network model. Pattern Recognition **38**(2), 199–208 (2005)
10. Homaeinezhad, M.R., Atyabi, S.A., Tavakkoli, E., Toosi, H.N., Ghaffari, A., Ebrahimpour, R.: ECG arrhythmia recognition via a neuro-SVM-KNN hybrid classifier with virtual QRS image-based geometrical features. Expert Systems with Applications **39**, 2047–2058 (2012)
11. Ince, T., Kiranyaz, S., Gabbouj, M.: A generic and robust system for automated patient-specific classification of ECG signals. IEEE Transactions on Biomedical Engineering **56**(5), 1415–1427 (2009)
12. Khazaee, A., Ebrahimzadeh, A.: Classification of electrocardiogram signals with support vector machines and genetic algorithms using power spectral features. Biomedical Signal Processing and Control **5**(4), 252–263 (2010)
13. Korüek, M., Doĝan, B.: ECG beat classificartion using particle swarm optimization and radial basis function neural network. Expert Systems with Applications **37**, 7563–7569 (2010)
14. Korürek, M., Dogan, B.: ECG beat classification using particle swarm optimization and radial basis function neural network. Expert Systems with Applications **37**(12), 7563–7569 (2010)
15. Llamedo, M., Martínez, J.P.: Heartbeat classification using feature selection driven by database generalization criteria. IEEE Transactions on Biomedical Engineering **58**(3), 616–625 (2011)
16. Luz, E., Menotti, D.: How the choice of samples for building arrhythmia classifiers impact their performances. In: Engineering in Medicine and Biology Society (EMBC), Annual International Conference of the IEEE, pp. 4988–4991. IEEE, Boston, EUA (2011)
17. Luz, E.J.S., Nunes, T.M., Albuquerque, V.H.C., Papa, J.P., Menotti, D.: ECG arrhythmia classification based on optimum-path forest. Expert Systems with Applications **40**(9), 3561–3573 (2013)
18. Mar, T., Zaunseder, S., Martínez, J.P., Llamedo, M., Poll, R.: Optimization of ECG classification by means of feature selection. IEEE Transactions on Biomedical Engineering **58**(8), 2168–2177 (2011)

19. Mark, R.G., Schluter, P.S., Moody, G.B., Devlin, P.H., Chernoff, D.: An annotated ECG database for evaluating arrhythmia detectors. IEEE Transactions on Biomedical Engineering **29**(8), 600 (1982)
20. Martis, R.J., Acharya, R., Adeli, H.: Current methods in electrocardiogram characterization. Computers in Biology and Medicine **48**, 133–149 (2014)
21. Martis, R.J., Acharya, U.R., Mandana, K., Ray, A., Chakraborty, C.: Application of principal component analysis to ECG signals for automated diagnosis of cardiac health. Expert Systems with Applications **39**, 11,792–11,800 (2012)
22. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEEE Engineering in Medicine and Biology Magazine **20**(3), 45–50 (2001)
23. Nejadgholi, I., Mohammad, M.H., Abdolali, F.: Using phase space reconstruction for patient independent heartbeat classification in comparison with some benchmark methods. Computers in Biology and Medicine **41**, 411–419 (2011)
24. Papa, J.P., Falcão, A.X., de Albuquerque, V.H.C., Tavares, J.M.R.S.: Efficient supervised optimum-path forest classification for large datasets. Pattern Recognition **45**(1), 512–520 (2012)
25. Papa, J.P., Falcão, A.X., Suzuki, C.T.N.: LibOPF: A library for the design of optimum-path forest classifiers. Campinas, SP (2009). Version 2.1, available at `http://www.ic.unicamp.br/~afalcao/LibOPF`
26. Papa, J.P., Falcão, A.X., Suzuki, C.T.N.: Supervised pattern classification based on optimum-path forest. International Journal of Imaging Systems and Technology **19**(2), 120–131 (2009)
27. Park, K.S., Cho, B.H., Lee, D.H., Song, S.H., Lee, J.S., Chee, Y.J., Kim, I.Y., Kim, S.I.: Hierarchical support vector machine based heartbeat classification using higher order statistics and hermite basis function. In: Computers in Cardiology, pp. 229–232 (2008)
28. Rai, H.M., Trivedi, A., Shukla, S.: Ecg signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier. Measurement **46**, 3238–3246 (2013)
29. Song, M.H., Lee, J., Cho, S.P., Lee, K.J., Yoo, S.K.: Support vector machine based arrhythmia classification using reduced features. International Journal of Control, Automation, and Systems **3**(4), 509–654 (2005)
30. Wang, J.S., Chiang, W.C., Hsu, Y.L., Yang, Y.T.C.: ECG arrhythmia classification using a probabilistic neural network with a feature reduction method. Neurocomputing **116**, 38–45 (2013)
31. Y. Chen, S.Y.: Selection of effective features for ECG beat recognition based on nonlinear correlations. Artificial Intelligence in Medicine **54**, 43–52 (2012)
32. Ye, C., Coimbra, M.T., Kumar, B.V.K.V.: Arrhythmia detection and classification using morphological and dynamic features of ECG signals. In: IEEE International Conference on Engineering in Medicine and Biology Society, pp. 1918–1921. IEEE, Buenos Aires, Argentina (2010)

33. Yu, S., Chen, Y.: Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network. Pattern Recognition Letters **28**(10), 1142–1150 (2007)
34. Yu, S., Chou, K.: Integration of independent component analysis and neural networks for ECG beat classification. Expert Systems with Applications **34**(4), 2841–2846 (2008)