

Detection of Life-Threatening Arrhythmias Using Feature Selection and Support Vector Machines

Felipe Alonso-Atienza*, Eduardo Morgado, Lorena Fernández-Martínez, Arcadi García-Alberola, and José Luis Rojo-Álvarez, *Senior Member, IEEE*

Abstract—Early detection of ventricular fibrillation (VF) and rapid ventricular tachycardia (VT) is crucial for the success of the defibrillation therapy. A wide variety of detection algorithms have been proposed based on temporal, spectral, or complexity parameters extracted from the ECG. However, these algorithms are mostly constructed by considering each parameter individually. In this study, we present a novel life-threatening arrhythmias detection algorithm that combines a number of previously proposed ECG parameters by using support vector machines classifiers. A total of 13 parameters were computed accounting for temporal (morphological), spectral, and complexity features of the ECG signal. A filter-type feature selection (FS) procedure was proposed to analyze the relevance of the computed parameters and how they affect the detection performance. The proposed methodology was evaluated in two different binary detection scenarios: shockable (FV plus VT) versus nonshockable arrhythmias, and VF versus nonVF rhythms, using the information contained in the medical imaging technology database, the Creighton University ventricular tachycardia database, and the ventricular arrhythmia database. sensitivity (SE) and specificity (SP) analysis on the out of sample test data showed values of SE = 95%, SP = 99%, and SE = 92%, SP = 97% in the case of shockable and VF scenarios, respectively. Our algorithm was benchmarked against individual detection schemes, significantly improving their performance. Our results demonstrate that the combination of ECG parameters using statistical learning algorithms improves the efficiency for the detection of life-threatening arrhythmias.

Index Terms—Feature selection (FS), support vector machines (SVM), ventricular fibrillation (VF) detection.

I. INTRODUCTION

SUDDEN cardiac arrest (SCA) is a major health problem that accounts approximately for six millions deaths in Europe and in the United States [1]. SCA is a sudden, abrupt loss of heart function, most often caused by a rapid ventricular tachycardia (VT) that quickly degenerates into ventricular

fibrillation (VF). Prompt detection of VT and VF episodes is crucial to deliver an electric shock therapy and in this way increase the probability of survival from a SCA incident. This has impelled the development of automated external defibrillators that analyze the surface electrocardiogram (ECG) signal and advise/deliver and electric shock if either rapid VT or VF is detected. However, though extensively tested and studied during the last decades both by the industry and by the scientific community, reliable detection of life-threatening arrhythmias remains an open problem [2]–[5].

A wide variety of detection algorithms have been developed based on temporal/morphological [4]–[8], spectral [9], [10], or complexity parameters [11]–[14] extracted from the ECG signal. For each detector, different separation scenarios have been considered [7], such as VF versus nonVF rhythms, VF plus VT versus nonVTVF, or VF versus VT, making it difficult to assess the real performance of the proposed algorithms. When compared in a standardized way [4], their real performance is reduced from the values presented in the original investigations. The combination of ECG parameters using machine learning techniques, such as neural networks [15]–[17], or support vector machines (SVM) [18]–[20], has been suggested as a useful approach to improve the detection efficiency. This strategy, however, raises additional requirements to be considered. First, the need of feature selection (FS) techniques to select those relevant and informative parameters in order to increase the efficiency of the learning task, to improve the performance of the detection process, and to better understand how data affect the learning process [18], [21]–[23]. And second, the evaluation and comparison of the proposed algorithms should be assessed over the out of sample test set. Broadly, this task has been carried out over the entire [12], [24] or the validation [20] datasets, making it difficult to compare different detection strategies.

The present study aimed to build a high-performance life-threatening arrhythmias detector by combining 13 previously defined ECG parameters using SVM learning algorithms. In this context, the objective is twofold: to assess the performance of the proposed SVM detection algorithm over previously defined methods by carrying out a comparative analysis on the out of sample test data. The second aim was to examine the discriminatory properties of each ECG parameter individually and how, in combination, these affect the learning process. We used a novel FS filter-type method based on combining three different FS filter-type techniques into a single ranking score, allowing us to determine the relevance of each ECG parameter. Using this score, we applied a backward selection procedure with SVM classifiers to yield a robust classifier using a reduced set of ECG parameters. Previous studies have used genetic

Manuscript received November 5, 2013; accepted November 5, 2013. Date of publication November 13, 2013; date of current version February 14, 2014. This work was partially supported by Research Grants IPT-2012-1126-300000, TSI-020100-2009-332, and TEC2010-19263/TCM. *Asterisk indicates corresponding author.*

*F. Alonso-Atienza is with the Department of Signal Theory and Communications, Rey Juan Carlos University, Fuenlabrada, 28943 Madrid, Spain (e-mail: felipe.alonso@urjc.es).

E. Morgado, L. Fernández-Martínez, and J. L. Rojo-Álvarez are with the Department of Signal Theory and Communications, Rey Juan Carlos University, 28943 Madrid, Spain (e-mail: eduardo.morgado@urjc.es; lorena.fernandez@urjc.es; joseluis.rojo@urjc.es).

A. García-Alberola is with the Arrhythmia Unit, University Hospital Virgen de la Arrixaca, 30120 Murcia, Spain (e-mail: arcadi@secardiologia.es).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2013.2290800

algorithms (GA) [20] or discriminant analysis [24] as FS methodology. In [20], a subset of nine features was selected using GA. However, this procedure did not provide a score metric, and therefore, it required to analyze all possible combination of a subset of selected features, which could be a heavily computationally expensive task, especially when the number of selected features increases. On the other hand, the discriminant analysis procedure [24] only accounts for linear relationships among features and labels, thus, missing possible nonlinear information that might be of relevance.

The proposed methodology is here applied in two different binary detection scenarios: shockable (FV plus TV) versus nonshockable arrhythmias, and VF versus nonVF rhythms. We used the public databases medical imaging technology database (MITDB) [25], the Creighton University Ventricular Tachycardia Database (CUDB) [26], and the MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB) [27], to evaluate our algorithms, showing that the results significantly outperform individual detection schemes.

We note that a preliminary version of this paper appeared in [19], showing the usefulness of SVM classification methodology for the detection of life-threatening arrhythmias. Here, we present a much-extended version of this study that includes: 1) two additional ECG parameters [5], [14]; 2) comparative analysis with previously defined detectors on the out of sample test set; and 3) a novel FS procedure that provides with insights about the relevance of each ECG parameter and the learning process using SVM algorithms.

The paper is organized as follows. Section II shows the preprocessing steps to build the dataset of computed parameters from the ECG signal databases. Section III provides a brief background on SVM classifiers and bootstrap resampling techniques. In Section IV, the FS method used in this study, and the algorithm combining FS and SVM classifiers are explained. Then, the detection performance of the proposed methodology is presented in Section V. Finally, discussion and conclusions are drawn in Section VI.

II. FEATURE CONSTRUCTION

This section illustrates the process of building the input space data to feed the SVM classifier from the ECG raw data signals.

A. ECG Collection

We used the complete ECG signal recording files from the MITDB, the CUDB, and the VFDB, which are available at the PhysioNet repository [28]. The MITDB contains 48 Holter recording files of slightly over 30-min length, two channels per file, sampled at 360 Hz. The MITDB includes 15 rhythm labels differentiating between VT, ventricular flutter (VFL), normal sinus rhythm (NSR), among other rhythms. The CUDB contains 35 Holter records of 8-min length from patients who experienced episodes of sustained VT, VFL, and VF. Each record is sampled at 250 Hz and includes only two rhythm annotations, namely, VF and nonVF. The VFDB contains 22 files of 30-min length, two channels per file, sampled at 250 Hz. As the CUDB, the VFDB includes patients who experienced episodes of sustained

VT, VFL, and VF. In this database, annotation labels contain 15 different rhythms, including VT, VF, VFL, NSR, among other rhythms.

B. Preprocessing

All ECG signals were preprocessed using the filtering process proposed in [4], which works in four successive steps: 1) mean subtraction; 2) five-order moving average filtering; 3) high-pass filtering with $f_c = 1$ Hz (drift suppression); and 4) low-pass Butterworth filtering with $f_c = 30$ Hz. Then, noise, asystole, and low-quality (artifacts) episode segments were removed according to the corresponding annotation labels. Finally, only the first channel of the MITDB and the VFDB has been considered, to avoid redundancy of samples during the learning process.

C. ECG Parameters

Each preprocessed ECG signal is divided in nonoverlapping 8-s segments. This window length has demonstrated to give the best performance in a number of investigated detection algorithms [4]. For each $L_e = 8$ s segment, a set of 13 previously defined parameters were computed. These can be broadly classified in three major categories (a detailed explanation of each parameter can be found at the original manuscripts).

1) *Temporal/Morphological Parameters*: are defined in the time domain.

- Threshold crossing interval (TCI) [6] is the time interval between consecutive pulses (threshold crossings) within a 1-s ECG segments. TCI requires a 3-s window to be computed. On a L_e duration episode, TCI is evaluated by averaging $L_e - 2$ consecutive values.
- Threshold crossing sample count (TCSC) [5] refers to the number of samples that cross a given threshold V_0 within a 3-s ECG interval. On a L_e duration episode, TCSC is evaluated by averaging $L_e - 2$ consecutive TCSC values.
- Standard exponential (STE) [4] is calculated as the ratio between the number of crossing points of the ECG signal with a decreasing exponential curve centered at the time instant where maximum amplitude value occurs, and the time duration of the considered ECG segment L_e .
- Modified exponential (MEA) [4] first adjusts a decreasing exponential function positioned at the peak values of an ECG segment. Then, MEA is computed as ratio between the number of liftings, and the time duration of the considered ECG segment L_e .
- Mean absolute value (MAV) [7] is the MAV of 2-s ECG segments. On a L_e duration episode, MAV is obtained by averaging $L_e - 1$ consecutive 2-s values.

2) *Spectral parameters*: are calculated in the frequency domain.

- VF filter (VFleak) [10] is a measure of the residue after applying a narrowband elimination filter centered at the mean signal frequency of the considered ECG signal segment.
- Spectral algorithm (M and $A2$ parameters) [9] analyzes the energy content in different frequency bands by means of Fourier analysis. Let F be the peak frequency (component with largest amplitude) in the range of 0.5–9 Hz. Then, M

measures of the frequency content between 0 and the minimum of (20 F, 100 Hz), while A2 measures the frequency content between 0.7 and 1.4 F.

- Median frequency (FM) [29] is the central frequency of the spectral mass contained in the power spectrum of the considered ECG signal segment. This parameter was defined to estimate the duration of the cardiac arrest, and therefore it has not been usually use for detection purposes. However, since it provides information about the duration of the VF episode, we included it here to analyze its discriminatory properties.

3) *Complexity parameters*: provide with different measures of the complexity of the ECG signal.

- Complexity measurement (CM) [11] is the normalized value of the Lempel-Ziv complexity measure of a binary sequence extracted from the ECG signal segment.
- Phase space reconstruction (PSR) [12] measures the sparsity of the phase plot representation when considering the original ECG signal segment and a time-delayed version of it.
- Hilbert transform (HILB) [13] measures the sparsity of the phase plot representation when considering the original ECG signal segment and its HILB signal.
- Sample entropy (SpEn) [14] is a measure of similarity within an ECG signal segment. A lower value of SpEn indicates more self-similarity. Thus, VF/VT rhythms are characterized by higher values os SpEn.

After computing all the aforementioned parameters, labels were assigned to each 8-s segments. In order to analyze the VF versus nonVF, and the shockable versus nonshockable problems, we considered three types of rhythms labels: VF (including VFL), VT, and other rhythms (O). Labels were assigned according to the mode of the annotation samples within the analyzed segment. For instance, in a transition ECG segment in which 40% of samples are labeled as NSR and the remaining samples are labeled as VF, then we labeled the whole segment as a VF.

The parameterization of the ECG signal segments resulted in a dataset of binary labeled data $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$, where $\mathbf{x}_i \in \mathbb{R}^d$, with $d = 13$ (number of computed parameters), $N = 17\,857$ (number of 8-s segments), and labels $y_i \in \{+1, -1\}$. Two binary detection scenarios were considered: VF episodes versus nonVF, and shockable (VF plus VT) versus nonshockable rhythms. Both problems resulted in unbalanced datasets with the following prior probabilities: VF versus nonVF, ($p_{+1} = 4.8\%$, $p_{-1} = 95.2\%$); and shockable versus nonshockable, ($p_{+1} = 8.5\%$, $p_{-1} = 91.5\%$). Before the classification process, each input feature example $\mathbf{x}^{(j)} \in \mathbb{R}^N$ was scaled so that $0 \leq \mathbf{x}^{(j)} \leq 1$.

III. SVM CLASSIFIERS

We used two different SVM classifiers to discriminate VF versus nonVF rhythms, and shockable versus nonshockable (from now on Shock versus nonShock) episodes by using the dataset of parameters extracted from the ECG signals. This section briefly reviews the SVM algorithm formulation and the boot-

strap resampling method to estimate the performance of the SVM classifiers.

A. SVM Formulation

In recent years, SVM algorithms have been successfully used in a wide number of practical classification problems [30], due to their good generalization capability derived from the structural risk minimization principle [31]. SVM binary classifiers are sampled-based statistical learning algorithms that construct a maximum margin separating hyperplane. Given a training dataset $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$, where $\mathbf{x}_i \in \mathbb{R}^d$ and $y_i \in \{-1, +1\}$, SVM solves a quadratic optimization problem

$$\min_{\mathbf{x}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i,$$

$$\text{subject to } y_i (\langle \phi(\mathbf{x}_i), \mathbf{w} \rangle + b) - 1 + \xi_i \geq 0,$$

$$\xi_i \geq 0, i = 1, \dots, N \quad (1)$$

where $\phi(\mathbf{x}_i)$ is a nonlinear transformation that maps training data to a higher dimensional space, ξ_i represent the losses, and C is a regularization parameter that represents a trade-off between the margin and the losses. By using Lagrange multipliers, (1) can be rewritten into its dual form, and then, the problem consists of solving

$$\max_{\alpha_i} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i,j=1}^N \alpha_i y_i \alpha_j y_j K(\mathbf{x}_i, \mathbf{x}_j) \quad (2)$$

constrained to $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^N \alpha_i y_i = 0$, where α_i are the Lagrange multipliers corresponding to primal constraints, and $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ is the kernel function, which allows us to calculate the dot product of pairs of vectors transformed by $\phi(\cdot)$ without explicitly knowing neither the nonlinear mapping nor the higher dimensional space. We used the Gaussian kernel in our experiments

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2). \quad (3)$$

After obtaining the Lagrange multipliers, the SVM classification for a new sample \mathbf{x} is simply given by

$$y = \text{sgn} \left(\sum_{i=1}^N \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \right). \quad (4)$$

The free parameters of the SVM model γ and C have to be settled a priori. Methods such as cross validation can be used for this purpose.

B. Bootstrap Resampling

Bootstrap resampling is a computer-based method for nonparametric estimation of the distribution of statistical magnitudes, and it can be used to estimate the performance of SVM classifiers [18]. Let $\mathbf{V} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ be a set of data in a classification problem. A bootstrap resample $\mathbf{V}^* = \{(\mathbf{x}_1^*, y_1^*), \dots, (\mathbf{x}_N^*, y_N^*)\}$ is a new dataset drawn at random with replacement from sample \mathbf{V} . Let us consider a partition of \mathbf{V} in terms of the resample, given by $\mathbf{V} = (\mathbf{V}_{\text{in}}^*, \mathbf{V}_{\text{out}}^*)$,

being \mathbf{V}_{in}^* and $\mathbf{V}_{\text{out}}^*$ the subsets of samples included and excluded in the resample, respectively. For the resample r , the SVM classifier can be trained with \mathbf{V}_{in}^* and its performance $P^*(r)$ can be estimated by using $\mathbf{V}_{\text{out}}^*$, in terms of a predefined metric such as the accuracy, or the error probability. Then, given a collection of R independent resamples, $\{\mathbf{V}^*(1), \mathbf{V}^*(2), \dots, \mathbf{V}^*(R)\}$, the performance density function can be estimated by the histogram built from replicates $P^*(r)$, where $r = 1, \dots, R$. A typical choice for R is from 100 to 500 resamples.

IV. FEATURE SELECTION

Performance of supervised learning algorithms can be strongly affected by the number and relevance of input variables. FS techniques aim to find the best describing subset of the input variables, compared to the original set of features [32]. FS techniques can be divided into three major categories, namely, filter, wrapper, and embedded methods.

In this study, we estimated the relevance of the computed parameters by applying a combination of filter-type FS procedures. From here on, we will use the terms *parameters* and *features* indistinctly. Filter methods are general FS procedures that rank the features according to a predefined evaluation criterion, which is independent of the machine learning classifier. Examples of filter methods include correlation criteria, classical test statistics (χ^2 -test, F-test, t-test), principal/independent component analysis, mutual information techniques, classification trees, self-organizing trees, or fuzzy clustering.

A. Combined Filter Methods Procedure

Following a similar approach as in [33], we considered a combined strategy of filter methods, accounting for correlation-based methods (correlation criterion and the maximum separability Fisher criterion), and mutual information methods (minimal redundancy maximal relevance -mRMR- criterion [34]).

1) *Correlation criterion*: assesses the degree of dependence of individual parameters with the outcome. For the j th feature $\mathbf{x}^{(j)}$ with labels \mathbf{y} , the linear correlation coefficient is defined as

$$\rho(j) = \frac{\sum_{i=1}^N (x_i^{(j)} - \mu_j)(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i^{(j)} - \mu_j)^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (5)$$

where μ_j represents the mean value for samples of feature $\mathbf{x}^{(j)}$, and \bar{y} is the average of outcomes. Note that $-1 \leq \rho \leq 1$. Larger absolute values of ρ indicate higher linear correlation between $\mathbf{x}^{(j)}$ and \mathbf{y} , whereas they are uncorrelated if ρ approaches to zero.

2) *Fisher criterion*: measures the ability of the j th feature to separate between two sets of labeled data (positive and negatives instances) by computing the F-score as

$$F(j) = \frac{\mu(y_+) + \mu(y_-)}{\sigma^2(y_+) + \sigma^2(y_-)} \quad (6)$$

where $\mu(y_{\pm}) = \mu_{j,\pm} - \mu_j$ represents the difference between the average of the j th feature for the positive/negative classes $\mu_{j,\pm}$ and the whole set of samples μ_j . In the denominator, $\sigma^2(y_{\pm})$ is the sample variance of the positives/negative instances and can be calculated as $\sigma^2(y_{\pm}) = \frac{1}{n_{\pm}-1} \sum_{i=1}^{n_{\pm}} (x_{i,\pm}^{(j)} - \mu_{j,\pm})^2$, being n_{\pm} the number of positive/negative samples. The larger the value of $F(j)$, the more likely this feature is discriminative.

3) *mRMR Criterion*: both correlation and Fisher criteria are computationally easy and fast, but they do not reveal mutual information among features (apart from linear correlation). Therefore, we also applied the mRMR criterion, which aims at maximizing the mutual information between the outcomes and the feature distribution while minimizing the redundancy between features, according to the following expression:

$$\max_{\mathbf{x}^{(j)}} \left\{ \frac{1}{|S|} \sum_{\mathbf{x}^{(j)} \in S} \text{MI}(\mathbf{x}^{(j)}, \mathbf{y}) - \frac{1}{|S|^2} \sum_{\mathbf{x}^{(j)}, \mathbf{x}^{(k)} \in S} \text{MI}(\mathbf{x}^{(j)}, \mathbf{x}^{(k)}) \right\} \quad (7)$$

where $\text{MI}(x, y)$ accounts for the mutual information among variables x and y , and $|S|$ represents the size of the feature set.

Fig. 1(a) represents the normalized absolute value of the scores provided by the three filter methods under consideration, for both the VF versus nonVF (black bars), and the Shock versus nonShock (white bars) problems. The closer to one the score is, the more relevant the feature is considered. The correlation criterion showed a number of parameters as relevant, while F-score and mRMR are more conservative standing TCSC out as the most relevant feature. This result is emphasized in Fig. 1(b), in which the scores of individual filter methods are combined to provide a single score for each parameter. This combined score is calculated as the normalized absolute value of the element multiplication of the individual scores. According to the combined score, TCSC is clearly more relevant than the rest of parameters. Besides TCSC, SpEn and VFleak are also likely to be highly relevant features. The final score and ranking of the ECG parameters is presented in Table I.

B. FS With SVM Classifiers

Following a similar approach as in [35] and [36], we combined the above mentioned filter FS procedure and SVM algorithms in order to build a high-performance classifier. The list of ranked features presented in Table I provided with an estimate of how valuable an ECG feature is for the detection problems under analysis. We applied a backward selection procedure: starting from the completed dataset, we progressively eliminated the less relevant feature according to Table I, and then estimated the performance of the SVM classifier using bootstrap resampling (we set $R = 500$). The FS-SVM procedure is summarized in Algorithm 1. Using this algorithm, the performance of the SVM classifier for different subsets of ranked features can be estimated.

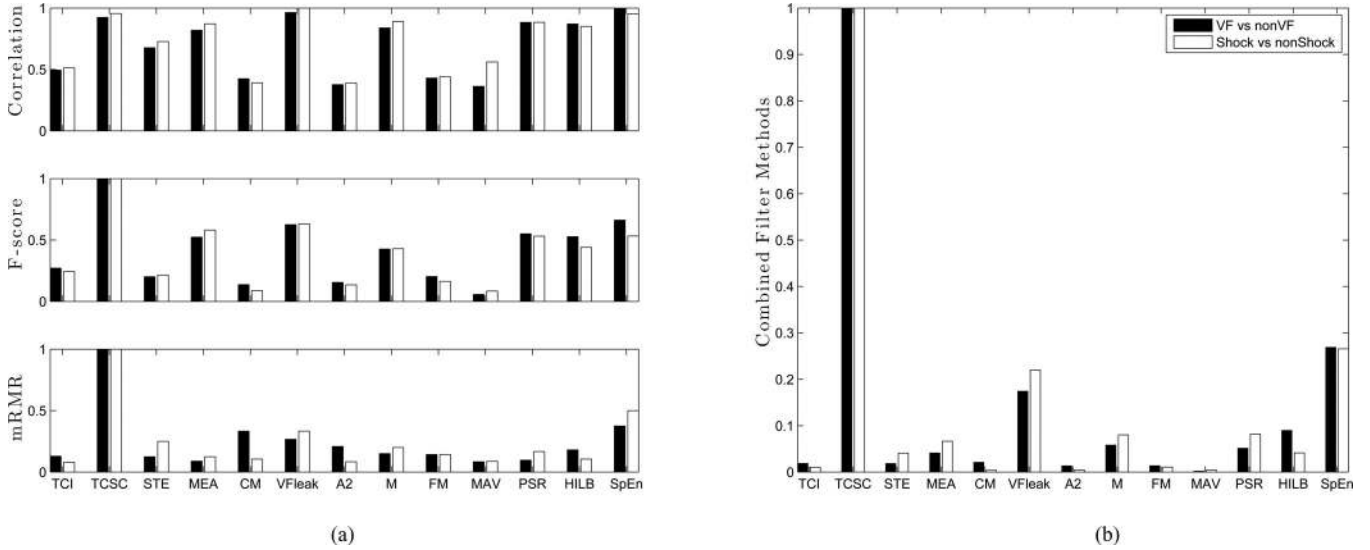


Fig. 1. Normalized feature ranking weights using (a) individual scores criteria: correlation, Fisher, mRMR; and (b) combination of individual scores. Black bars correspond to the VF detection problem, whereas white bars represent the shockable rhythms detection scenario.

TABLE I
SCORE AND RANKING ANALYSIS OF ECG PARAMETERS

VF vs nonVF		Shock vs nonShock	
Parameter	Score	Parameter	Score
TCSC	1.0	TCSC	1.0
SpEn	0.27	SpEn	0.26
VFleak	0.17	VFleak	0.22
HILB	0.09	PSR	0.08
M	0.06	M	0.08
PSR	0.05	MEA	0.07
MEA	0.04	HILB	0.04
CM	0.02	STE	0.04
TCI	0.02	FM	0.01
STE	0.02	TCI	0.01
FM	0.01	A2	0.00
A2	0.01	MAV	0.00
MAV	0.00	CM	0.00

Algorithm 1 : FS-SVM algorithm

- 1) Sort the features according to the combined-filter score \mathbf{s} , being $\mathbf{s}[1]$ the highest score variable and $\mathbf{s}[d]$ the lowest scored, respectively, and d the number of features of the input space.
- 2) Initialization: $\mathbf{V} = \{\mathbf{X}(:, \mathbf{s}), \mathbf{y}\}$, $j = d$.
- 3) **Repeat**:
 - a) Train the SVM classifier using $\{\mathbf{X}(:, \mathbf{s}[1 : j]), \mathbf{y}\}$, and calculate the free parameters (C_j, γ_j) .
 - b) $j = j - 1$.
- 4) **Until** $j = 0$.
- 5) Start with complete dataset $\mathbf{V} = \{\mathbf{X}(:, \mathbf{s}), \mathbf{y}\}$, $j = d$.
- 6) Build R paired bootstrap resamples $\mathbf{V}^* = \{\mathbf{X}^*(:, \mathbf{s}), \mathbf{y}^*\}$
- 7) **Repeat**:
 - a) For each bootstrap resample r , compute the performance of the SVM $P^*(r)$, $r = 1, \dots, R$ using the calculated free parameters (C_j, γ_j) .
 - b) Remove the feature with the lowest score.
 - c) $\mathbf{V}^* = \{\mathbf{X}^*(:, \mathbf{s}[1 : j]), \mathbf{y}^*\}$
 - d) $j = j - 1$.
- 8) **Until** $j = 0$.

V. RESULTS

A. Individual Parameters Performance

First, we studied the discrimination ability of the calculated ECG parameter by analyzing their corresponding receiver operating characteristics (ROC) curve obtained by using the entire dataset. The performances of the detection parameters were assessed in terms of the area under the ROC curve (AUC) and by evaluating the sensitivity (SE), i.e., the proportion of correctly detected VF/Shockable observations, and the specificity (SP), i.e., the proportion of correctly identified nonVF/nonShockable samples. SE and SP are calculated as

$$SE = \frac{TP}{TP + FN} \quad (8)$$

$$SP = \frac{TN}{TN + FP} \quad (9)$$

where TP represents the number of true-positive decisions, FN the number of false-negative decisions, TN the number of true-negative decisions, and FP the number of false-positive decisions. The results of the ROC analysis are presented in Table II. For both VF versus nonVF and Shock versus nonShock problems, the TCSC parameter obtained the best performance, hence supporting the original study [5]. Also, VFleak and SpEn showed high values of SE and SP. The performances of PSR and HILB are quite acceptable. However, this result differs from the original investigations [12], [13], but it is similar to other studies [5]. Note that the performance results presented in Table II are highly correlated with the scores and ranking shown in Table I, thus demonstrating that the proposed FS procedure is an interesting method for evaluating the individual performance of a set of ECG parameters.

B. SVM Performance

In this experiment, we aimed to analyze the performance of the SVM algorithm when using the complete set of ECG

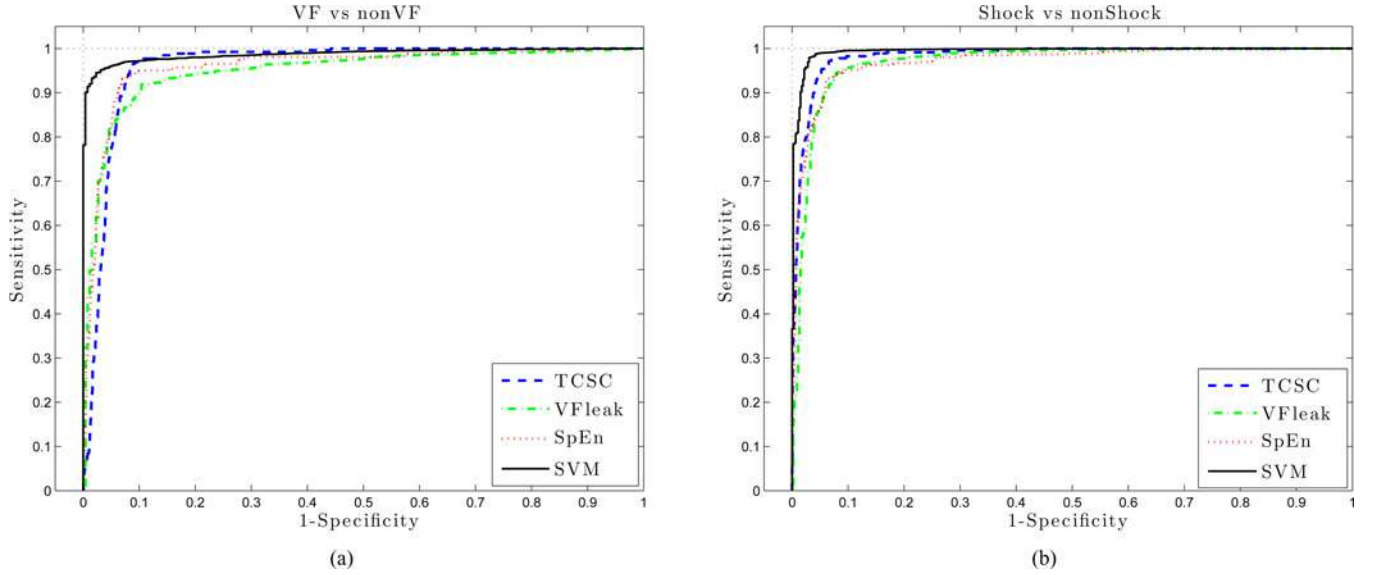


Fig. 2. ROC curves calculated on the out of sample test set for the (a) VF versus nonVF problem, and (b) shockable versus nonshockable scenario.

TABLE II
ROC ANALYSIS FOR THE COMPUTED PARAMETERS USING
THE COMPLETE DATASET

Param	VF vs nonVF			Shock vs nonShock		
	AUC	SE ^a	SP ^b	AUC	SE ^a	SP ^b
TCSC	0.96	75	92	0.98	92	96
SpEn	0.94	77	90	0.96	83	93
VFleak	0.95	73	89	0.97	82	93
PSR	0.92	74	85	0.95	85	92
HILB	0.92	75	80	0.94	76	86
MEA	0.92	70	83	0.95	80	91
M	0.91	71	72	0.95	81	82
A2	0.88	34	74	0.90	51	72
TCI	0.89	49	68	0.92	65	76
FM	0.84	56	41	0.85	56	53
STE	0.83	48	47	0.88	62	58
MAV	0.72	24	18	0.79	43	26
CM	0.80	23	47	0.78	25	37

^a Sensitivity(%) for a 95% specificity.

^b Specificity(%) for a 90% sensitivity.

parameters. Thus, the complete dataset was used as the input to the SVM detector. A random subset of the input space (70%) was used for training while the remaining data were used as test set. Given that the datasets associated with the two problems under analysis were unbalanced, weights were assigned to each class. In addition, we used the balanced error rate (BER) [32] as the metric to set the free parameters (C, γ) of the SVM by following a fivefold cross-validation strategy over the training set. The performance of the SVM detector was assessed using the ROC analysis in terms of SE, SP, AUC and of the positive predictivity (PP), the accuracy (ACC) and the BER calculated over the test set as

$$PP = \frac{TP}{TP + FP} \quad (10)$$

$$ACC = \frac{TN + TP}{PC + NC} \quad (11)$$

$$BER = \frac{1}{2} \left(\frac{FN}{PC} + \frac{FP}{NC} \right) \quad (12)$$

where $PC = TP + FN$ and $NC = TN + FP$.

The performance of the SVM was benchmarked against the TCSC, SpEn, and VFleak parameters, as shown in Table III. Both in the VF and the shockable rhythms scenarios, the SVM classifier outperformed individual detectors in all analyzed metrics. The McNemar's test showed that these differences in performance were statistically significant (p -value < 0.001). A complete perspective of the performance of the SVM algorithm can be seen in Fig. 2, which represents the ROC curves for the SVM algorithm, and the TCSC, SpEn, and VFleak parameters for both the VF versus nonVF [panel (a)] and the Shock versus nonShock scenarios [panel (b)].

Table IV shows the performance of the analyzed algorithms per arrhythmia type. Besides ventricular arrhythmias (VT, VFL, VF) and NSR, supraventricular rhythms have been included in the table since they might produce misclassification when discriminating ventricular rhythms. In almost all types of rhythms and in both detection problems, the SVM surpass (if possible) the individual detectors. Note that in the VF detection scenario, VT is poorly classified by all algorithms showing that the separability of VT and VF/VFL rhythms is difficult to achieve. With respect to individual detectors, TCSC behaves moderately well to detect shockable rhythms but fails to discriminate VT and VF. VFleak is able to detect shockable rhythms very accurately, but does not discriminate VT and VF and behaves poorly with supraventricular arrhythmias. SpEn discriminates VT and VF better than TCSC and VFleak, but shows inferior performance when classifying shockable rhythms.

TABLE III
COMPARATIVE ANALYSIS (METRICS IN %) OF THE SVM DETECTOR (TEST SET)

	VF vs nonVF						Shock vs nonShock					
	SE	SP	PP	ACC	BER	AUC	SE	SP	PP	ACC	BER	AUC
SVM	91.9	97.1	61.6	96.8	5.5	98.7	95.0	99.0	89.7	98.6	3.0	99.4
TCSC [†]	88.1	93.5	40.9	93.2	9.2	96.1	86.7	96.7	70.9	95.8	8.3	98.3
SpEn [†]	84.6	94.7	44.7	94.2	10.36	95.9	77.7	97.2	72.4	95.5	12.5	97.0
VFleak [†]	88.0	91.9	35.8	91.7	10.0	94.7	90.6	95.4	64.8	95.0	7.0	96.9

[†] McNemar's test p-value < 0.001.

TABLE IV
SVM PERFORMANCE (ACCURACY IN %) PER ARRHYTHMIA TYPE[†] (TEST SET)

Rhythm	VF vs nonVF				Shock vs nonShock			
	SVM	TCSC	SpEn	VFleak	SVM	TCSC	SpEn	VFleak
NSR	99.1	95.4	96.4	94.3	98.7	95.4	96.4	94.3
AFIB	100.0	100.0	100.0	99.3	100.0	100.0	100.0	99.3
AFL	100.0	100.0	73.1	73.1	100.0	100.0	73.1	73.1
SVTA	98.2	98.2	98.2	98.2	98.2	98.2	98.2	98.2
SBR	100.0	100.0	100.0	98.4	100.0	100.0	100.0	98.4
VT	41.9	15.1	31.3	6.0	96.0	84.8	68.7	93.9
VFL	100.0	100.0	65.2	91.3	100.0	100.0	65.2	91.3
VF	91.1	86.9	86.5	87.8	93.7	86.9	86.5	87.8

[†] Acronyms; NSR: normal sinus rhythm, AFIB: atrial fibrillation, AFL: atrial flutter, SVTA: supraventricular tachycardia, SBR: sinus bradycardia, VT: ventricular tachycardia, VFL: ventricular flutter, VF: ventricular fibrillation.

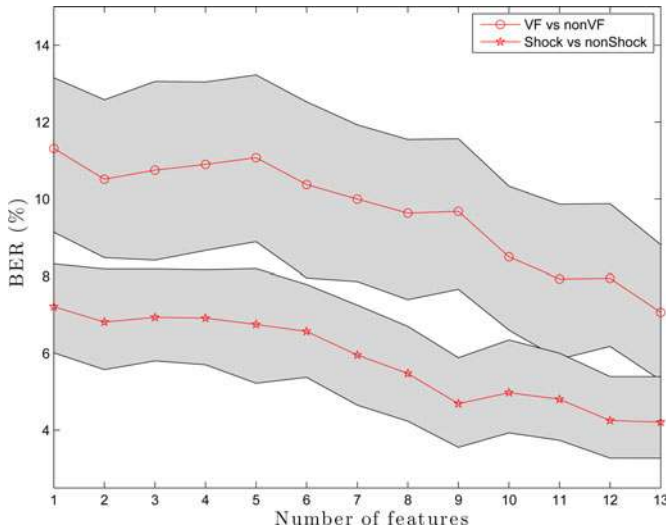


Fig. 3. BER metric (in %) analysis with respect to the number of features: mean (central line), and 95% confidence interval (gray area).

C. SVM Performance Using FS

Given that it might be possible that not all the computed parameters are relevant for detection purposes, we studied the usefulness of the calculated ECG features to construct a simplified and robust detector. For doing this, we combined FS techniques with SVM classifiers by following Algorithm 1. Fig. 3 shows the mean (red central line) and the 95% confidence interval (gray area) of the estimated SVM performance, in terms of the BER

TABLE V
ANALYSIS (METRICS IN %) OF THE FS+SVM DETECTOR (TEST SET) FOR THE SHOCK VERSUS NONSHOCK SCENARIO

	SE	SP	PP	ACC	BER	AUC
FS+SVM [†]	93.6	98.8	88.3	98.4	3.7	99.4
SVM	95.0	99.0	89.7	98.6	3.0	99.4

[†] McNemar's test p-value = 0.128.

metric (y-axis), using M ranked features (x -axis) for the two detection problems under consideration.

In the case of VF detection, the BER improves as the number of features increases, showing that the SVM classifier requires the information of all features due to the complexity of the problem. Thus, the classification model could not be simplified if the performance is to be maximized. On the other hand, in the shockable classification scenario, the mean BER value slightly increases from 13 to 9 features, from which the performance more rapidly decreases as the number of features reduces. Hence, we set $M = 9$ features to build a simplified SVM detection model, whose performance over the test set is summarized in Table V. Selected feature were: TCSC, SpEn, VFleak, PSR, M, MEA, HILB, STE, and FM. The McNemar's test showed no statistical differences in performance between the completed and the reduced models.

VI. DISCUSSION AND CONCLUSIONS

In this study, a novel detection algorithm that combines ECG parameters with SVM to identify VF/shockable arrhythmias has

been presented. Together with this algorithm, a FS procedure has been used to further analyze the discriminatory properties of the extracted ECG parameters. Given that, in general, parsimonious detection models provide better prediction, the proposed FS method has been combined to the SVM algorithm to provide a robust classifier using a reduced set of ECG parameters (for shockable rhythms). The detection performance of the developed methodology is remarkable, and it significantly outperforms previous proposed detection algorithms.

We used the complete records of the MITDB, CUDB, and the VFDB. No preselection of episodes was made. In the preprocessing task, noise, and asystole segments were removed from the classification procedure, as done in other studies [7], [8], [37]. For this purpose, we used the information contained in the annotation files. Nevertheless, usual signal processing algorithms could be applied instead. Noise can be detected by examining the slew rate of the ECG signal [8], while asystole intervals can be identified by amplitude and signal power analysis [37].

A total of 13 ECG parameters have been computed to characterize VF and shockable rhythms. These include widely analyzed parameters, such as TCI, CM, PSR, HILB, STE, A2, M, and VFLeak and relatively recent proposals, namely TCSC, SpEn, MEA, MAV, and FM (for detection purposes). The overall detection performances, when considering each parameter individually, are in agreement with previous studies [4], [5], [7], [13], [14], [38], demonstrating that TCSC, SpEn, and VFLeak provide the best diagnostic properties, followed by HILB and PSR, while TCI, CM, and STE perform poorly.

In this study, it has been shown that the use of SVM algorithms combining ECG features significantly improves the efficiency for the detection of life-threatening arrhythmias. SVM classifiers have been extensively used with the ECG signal in the context of wave delineation, beat detection, general arrhythmia discrimination, and in other application, such as heart rate variability or detection of ischemia (see [39] for a comprehensive review). However, the proposed utilization of SVM algorithms to detect VF/shockable episodes using a number of well-known ECG features has not been widely explored. In [40], a VF detection algorithm based on SVM algorithm combining the Hurst index and the PSR parameter was proposed, showing a better performance than VFLeak, TCI, CM, PSR, M, and A2 parameters, and thus emphasizing the efficacy of combining ECG parameters with robust machine learning algorithms. More recently, in [20] a total 14 ECG parameters were computed and a selection of nine were used with SVM classifiers. Compared to the present study, Li *et al.* [20] used a different window size (5 s), a different set parameters and different databases (they included the American Heart Association database). They reported values of SE = 90.2, SP = 99.6, and AUC = 99.7 on the test set. Differences in performance between this and the present study emphasize the need for building a public ECG signal database divided into a training and a test datasets, in order to compare machine learning strategies.

The use of machine learning algorithms requires to set the free parameters using a training set of examples. In the case of unbalanced datasets, this step is crucial to assess a good generalization performance. Under these circumstances, the BER

metric in SVM algorithms represents as a suitable figure of merit to jointly maximize the values of SE and SP. Besides the BER, other metrics taking into account the imbalanced nature of the detection problem [41], such as the F-measure or the AUC could be of interest to guide the classification model and set the free parameters.

Combining the information from a number of features to perform a given learning tasks requires FS methods to analyze the relevance of those features, in order to eliminate unnecessary or redundant information, and this way to construct a robust and well-performed machine learning algorithm. In this study, the relevance of 13 well-known ECG parameters to detect life-threatening arrhythmias has been studied using FS methods. We used a filter-type approach, combining the correlation, Fisher, and the mRMR criteria scores to take into account both linear and nonlinear relationship among features and the class label. The proposed FS methodology was consistent with the detection performance of the analyzed parameters, demonstrating that this procedure is an interesting method for evaluating the discrimination ability of a set of ECG parameters. Using this combined score ranking, parameters as TCI, A2, and CM, which have been extensively used in the literature, have shown not to be relevant features for arrhythmia discrimination.

It is also important to determine which parameters are relevant to the classification process, and how they affect the learning process. This has been evaluated by a backward selection procedure defined in Algorithm 1. Comparative analysis between the VF and the shockable detection scenarios (see Fig. 3), suggests that in the VF classification problem the SVM algorithm is suffering from high variance (wider confident interval area) and high bias (higher BER metric), thus indicating that there is still possibility to improve the performance if more training examples and features/parameters are incorporated into the SVM model. This again raises the need of having large public databases to evaluate VT/VF detection algorithms. Also, the FS analysis showed that features need to be combined in order to provide highly accurate results. The individual discriminative power of a variable is not sufficient to build a robust detector. The relationships with others and with the classifier have to be taken into account. In this context, more elaborated FS approaches (wrapper or embedded methods) and/or different classifiers could be of interest. However, this implementation exceeds the purpose of this study.

In conclusion, the present study has shown that the use of SVM learning algorithms can improve the efficiency for the detection of life-threatening arrhythmias. In this scenario, FS techniques might help to better understand the data and to provide valuable insights to build highly accurate detection algorithms. Also, in the case of SVM classifiers using unbalanced datasets, which constitute the standard case in arrhythmia detection problems, the BER metric is an interesting magnitude to set the free parameters of the algorithm.

ACKNOWLEDGMENT

The authors would like to thank to Óscar Barquero-Pérez and Rebeca Goya-Esteban for their kindly help in programming the complexity measurement and the sample entropy parameters.

REFERENCES

- [1] Y. Li, J. Bisera, M. Weil, and W. Tang, "An algorithm used for ventricular fibrillation detection without interrupting chest compression," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 78–86, Jan. 2012.
- [2] M. Harris, "A shocking truth," *IEEE Spectr.*, vol. 49, no. 3, pp. 30–58, Mar. 2012.
- [3] S. JS and M. WH. (2006). Recalls and safety alerts affecting automated external defibrillators. *J. Amer. Med. Assoc.* [Online]. 296(6), pp. 655–660. Available: + <http://dx.doi.org/10.1001/jama.296.6.655>
- [4] A. Amann, R. Tratnig, and K. Unterkofler, "Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators," *Biomed. Eng. Online*, vol. 4, no. 60, Oct. 2005.
- [5] M. Arafat, A. Chowdhury, and M. Hasan, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," *Signal, Image Video Process.*, vol. 5, pp. 1–10, 2011.
- [6] N. V. Thakor, Y. S. Zhu, and K. Y. Pan, "Ventricular tachycardia and fibrillation detection by a sequential hypothesis testing algorithm," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 9, pp. 837–43, Sep. 1990.
- [7] E. Anas, S. Lee, and M. Hasan, "Sequential algorithm for life threatening cardiac pathologies detection based on mean signal strength and emd functions," *BioMed. Eng. OnLine*, vol. 9, no. 1, 2010.
- [8] I. Jekova and V. Krasteva, "Real time detection of ventricular fibrillation and tachycardia," *Physiol. Meas.*, vol. 25, no. 5, pp. 1167–1178, 2004.
- [9] S. Barro, R. Ruiz, D. Cabello, and J. Mira, "Algorithmic sequential decision making in the frequency domain for life threatening centricular arrhythmias and aimitative artifacts: A diagnostic system," *J. Biomed. Eng.*, vol. 11, no. 4, pp. 320–328, Jul. 1989.
- [10] S. Kuo and R. Dillman, "Computer detection of ventricular fibrillation," in *Proc. Comput. Cardiol.*, 1978, pp. 2747–2750.
- [11] X. S. Zhang, Y. S. Zhu, N. V. Thakor, and Z. Z. Wang, "Detecting ventricular tachycardia and fibrillation by complexity measure," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 5, pp. 548–55, May 1999.
- [12] A. Amann, R. Tratnig, and K. Unterkofler, "Detecting ventricular fibrillation by time-delay methods," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 1, pp. 174–177, Jan. 2007.
- [13] A. Amann, R. Tratnig, and K. Unterkofler, "A new ventricular fibrillation detection algorithm for automated external defibrillators," in *Proc. Comput. Cardiol.*, Sep. 2005, pp. 559–562.
- [14] H. Li, W. Han, C. Hu, and M.-H. Meng, "Detecting ventricular fibrillation by fast algorithm of dynamic sample entropy," in *Proc. IEEE Int. Conf. Robot. Biomimet.*, Dec. 2009, pp. 1105–1110.
- [15] R. H. Clayton, A. Murray, and R. W. Campbell, "Recognition of ventricular fibrillation using neural networks," *Med. Biolog. Eng. Comput.*, vol. 32, no. 2, pp. 217–220, Mar. 1994.
- [16] A. Neurauter, T. Eftestol, J. Kramer-Johansen, B. Abella, K. Sunde, V. Wenzel, K. Lindner, J. Eilevsjonn, H. Myklebust, P. Steen, and H. Strohmeier, "Prediction of countershock success using single features from multiple ventricular fibrillation frequency bands and feature combinations using neural networks," *Resuscitation*, vol. 73, no. 2, pp. 253–263, May 2007.
- [17] J. Pardey, "Detection of ventricular fibrillation by sequential hypothesis testing of binary sequences," in *Proc. Comput. Cardiol.*, 2007, pp. 573–576.
- [18] F. Alonso-Atienza, J. L. Rojo-Álvarez, A. Rosado-Muñoz, J. J. Vinagre, and A. García-Alberola, G. Camps-Valls, "Feature selection using support vector machines and bootstrap methods for ventricular fibrillation detection," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 1956–1967, 2012.
- [19] F. Alonso-Atienza, E. Morgado, L. Fernandez-Martinez, A. Garcia-Alberola, and J. Rojo-Alvarez, "Combination of ECG parameters with support vector machines for the detection of life-threatening arrhythmias," in *Proc. Comput. Cardiol.*, 2012, pp. 385–388.
- [20] Q. Li, C. Rajagopalan, and G. Clifford, "Ventricular fibrillation and tachycardia classification using machine learning method," *IEEE Trans. Biomed. Eng.*, vol. PP, no. 99, pp. 1–1, 2013.
- [21] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, 2003.
- [22] Z. Zhang, S. Lee, and J. Lim, "Discrimination of ventricular arrhythmias using NEWFM," in *Proc. Fourth Asia Inf. Retrieval Conf. Inf. Retrieval Technol.*, 2008, pp. 176–183.
- [23] B. Ribeiro, A. Marques, J. Henriques, and M. Antunes, "Premature ventricular beat detection by using spectral clustering methods," in *Proc. Comput. Cardiol.*, Sep.30Oct. 3 20072007, pp. 149–152.
- [24] I. Jekova, "Shock advisory tool: Detection of life-threatening cardiac arrhythmias and shock success prediction by means of a common parameter set," *Biomed. Signal Process. Control*, vol. 2, no. 1, pp. 25–33, 2007.
- [25] G. Moody and R. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.
- [26] F. Nolle, F. Badura, J. Catlett, and B. R.W., and S. M.H. , "CREI-GARD, a new concept in computerized arrhythmia monitoring systems," in *Proc. Comput. Cardiol.*, 1986, vol. 13, pp. 515–518.
- [27] S. Greenwald, *The Development and Analysis of a Ventricular Fibrillation Detector*. Cambridge, MA, USA: Dept Electrical Eng. Comput. Sci., MIT, 1986.
- [28] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [29] R. Dzwonczyk, C. Brown, and H. Werman, "The median frequency of the ECG during ventricular fibrillation: Its use in an algorithm for estimating the duration of cardiac arrest," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 6, pp. 640–646, Jun. 1990.
- [30] G. Camps-Valls, J. L. Rojo-Álvarez, and M. Martínez-Ramón, *Kernel Methods in Bioengineering, Communications and Image Processing*. Hershey, PA, USA: Idea Group Inc., Jan 2007.
- [31] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag, 1999.
- [32] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature Extraction: Foundations and Applications*. Secaucus, NJ, USA: Springer-Verlag, 2006.
- [33] H. W. Cho, S. Baek, E. Youn, M. Jeong, and A. Taylor, "A two-stage classification procedure for near-infrared spectra based on multi-scale vertical energy wavelet thresholding and SVM-based gradient-recursive feature elimination," *J. Oper. Res. Soc.*, vol. 60, no. 8, pp. 1107–1115, Aug. 2009.
- [34] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.
- [35] Y. W. Chen and C. J. Lin, "Combining SVMs with various feature selection strategies," in *Feature extraction: Foundations and Applications*, I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Eds. Secaucus, NJ, USA: Springer-Verlag, 2006.
- [36] T. N. Lal, O. Chapelle, and B. Schölkopf, "Combining a filter method with SVMs," in *Feature extraction: Foundations and Applications*. I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Eds. Secaucus, NJ, USA: Springer-Verlag, 2006.
- [37] U. Irusta, J. Ruiz, E. Aramendi, S. R. de Gauna, U. Ayala, and E. Alonso, "A high-temporal resolution algorithm to discriminate shockable from nonshockable rhythms in adults and children," *Resuscitation*, vol. 83, no. 9, pp. 1090–1097, 2012.
- [38] E. Aramendi, U. Irusta, E. Pastor, A. Bodegas, and F. Benito, "ECG spectral and morphological parameters reviewed and updated to detect adult and paediatric life-threatening arrhythmia," *Physiol. Meas.*, vol. 31, no. 6, pp. 749–761, 2010.
- [39] J. L. Rojo-Alvarez, G. Camps-Valls, and A. J. Caamaño-Fernández, J. F. Guerrero-Martínez, "A review of kernel methods in ECG signal classification," in *ECG Signal Processing, Classification and Interpretation: A Comprehensive Framework of Computational Intelligence*. A. Gacek and W. Pedrycz, Eds. Berlin, Germany: Springer-Verlag, 2011.
- [40] Q. Li, J. Zhao, and Y.-N. Zhao, "Detection of ventricular fibrillation by support vector machine algorithm," in *Proc. Int. Asia Conf. Informat. Control, Autom. Robot.*, 2009, pp. 287–290.
- [41] H. He and E. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.

Authors' photographs and biographies not available at the time of publication.