I. Isasi et al., "A Machine Learning Shock Decision Algorithm for Use During Piston-Driven Chest Compressions," in IEEE Transactions on Biomedical Engineering, vol. 66, no. 6, pp. 1752-1760, June 2019, doi: 10.1109/TBME.2018.2878910.

A Machine Learning Shock Decision Algorithm for use during Piston-driven Chest Compressions

Iraia Isasi, Unai Irusta*, *Member, IEEE* Andoni Elola, Elisabete Aramendi, Unai Ayala, Erik Alonso, Jo Kramer-Johansen, and Trygve Eftestøl, *Member, IEEE*

Abstract*—*Goal: **Accurate shock decision methods during piston-driven cardiopulmonary resuscitation (CPR) would contribute to improve therapy and increase cardiac arrest survival rates. The best current methods are computationally demanding, and their accuracy could be improved. The objective of this work was to introduce a computationally efficient algorithm for shock decision during piston-driven CPR with increased accuracy.** ***Methods:*** **The study dataset contains 201 shockable and 844 nonshockable ECG segments from 230 cardiac arrest patients treated with the LUCAS-2 mechanical CPR device. Compression artifacts were removed using state of the art adaptive filters, and shock/no-shock discrimination features were extracted from the stationary wavelet transform analysis of the filtered ECG, and fed to a support vector machine (SVM) classifier. Quasi-stratified patient wise nested cross-validation was used for feature selection and SVM hyperparameter optimization. The procedure was repeated 50 times to statistically characterize the results.** ***Results:*** **Best results were obtained for a 6 feature classifier with mean (standard deviation) sensitivity, specificity, and total accuracy of 97.5 (0.4), 98.2 (0.4) and 98.1 (0.3), respectively. The algorithm presented a five-fold reduction in computational demands when compared to the best available methods, while improving their balanced accuracy by 3-points.** ***Conclusions:*** **The accuracy of the best available methods was improved while drastically reducing the computational demands.** ***Significance:*** **An efficient and accurate method for shock decisions during mechanical CPR is now available to improve therapy and contribute to increase cardiac arrest survival.**

***Index Terms*—Support Vector Machine (SVM), Machine Learning, Stationary Wavelet Transform (SWT), Cardiac arrest, cardiopulmonary resuscitation (CPR), electrocardiogram (ECG), mechanical chest compressions, piston-driven compressions, shock decision algorithm.**

Manuscript submitted X, 2018; accepted X, X. This work was supported by: The Spanish Ministerio de Economía y Competitividad, TEC2015-64678-R, jointly with the Fondo Europeo de Desarrollo Regional (FEDER), UPV/EHU via GIU17/031 and the Basque Government through grant pre-2017-2-0137.

Asterisk indicates corresponding author.

*I. Isasi is with the Department of Communications Engineering, University of the Basque Country UPV/EHU, Ingeniero Torres Quevedo Plaza, 1, 48013, Bilbao, Spain (e-mail: iraia.isasi@ehu.eus).

U. Irusta, E. Aramendi and A. Elola are with the Department of Communications Engineering, University of the Basque Country UPV/EHU, Ingeniero Torres Quevedo Plaza, 1, 48013, Bilbao, Spain.

U. Ayala is with the Department of Signal Processing and Communications, Mondragon University, Loramendi, 4, 20500 Arrasate, Spain.

E. Alonso is with the Department of Applied Mathematics, University of the Basque Country UPV/EHU, Rafael Moreno Pitxitxi 3, 48013 Bilbao, Spain.

J. Kramer-Johansen is with the Norwegian National Advisory Unit on Prehospital Emergency Medicine (NAKOS) and Department Of Anaesthesiology, Oslo University Hospital and University of Oslo, Pb 4956 Nydalen, 0424 Oslo, Norway.

T. Eftestøl is with the Department of Electrical Engineering and Computer Science, University of Stavanger, 4036 Stavanger, Norway.

I. INTRODUCTION

HIGH quality cardiopulmonary resuscitation (CPR) and early defibrillation are key for the survival of out-of-hospital cardiac arrest (OHCA) patients [1]. During CPR, chest compressions and ventilations should be delivered according to international guidelines [1]. Interruptions in chest compressions decrease coronary perfusion pressure [2], and may compromise the survival of the patient [3]. Chest compressions induce an artifact in the ECG, so current defibrillators instruct the rescuers to stop chest compressions for a reliable shock decision [4].

Many efforts have been made to allow a reliable shock decision during CPR, with solutions that go from analyzing the rhythm during ventilation pauses [5], [6] to ad-hoc algorithms designed for a reliable shock decision in the presence of chest compression artifacts [7], [8], [9]. The best known solutions are based on adaptive filters that remove the CPR artifact before using the shock decision algorithm of the defibrillator. These filters model the artifact using additional reference channels recorded by the defibrillator such as compression depth, thoracic impedance, chest acceleration, or chest force/pressure. Several solutions have been proposed including Wiener filters [10], Matching Pursuit algorithms [11], [12], Kalman filters [13], [14], Gabor filters [15], Least Mean Squares (LMS) filters [16], [17], [18] and Recursive Least Squares (RLS) filters [19]. Reference channels are not always available and may increase the cost of defibrillators, fortunately filters based only on the frequency of chest compressions are as effective as complex filters based on several reference channels [16], [20]. For manual CPR, solutions based on adaptive filters followed by the shock decision algorithms of commercial defibrillators do not meet the accuracy requirements of the American Heart Association (AHA) [4]. The sensitivity (Se) for shockable rhythms is above the minimum 90% recommendation, but the specificity (Sp) for nonshockable rhythms is below the minimum recommended value of 95%. Filtering residuals have been identified as the main confounding factor for the shock decision algorithms of commercial defibrillators [12], [21], which are designed to classify ECGs free of artifacts [22].

Mechanical CPR is becoming increasingly popular to treat OHCA patients, even if it has not shown benefits in survival [23], [24], [25]. Mechanical devices guarantee high quality chest compressions, and have become important in scenarios where manual CPR is impractical, such as during transport or invasive procedures [26], [27], [28],

[24]. There are two families of mechanical compressors available: pneumatically driven pistons and load distributing bands. According to the resuscitation guidelines the most popular/widespread devices are the LUCAS-2 (Physio-Control Inc/Jolife AB, Lund, Sweeden) piston-driven device and the Autopulse (Zoll Circulation, Chelmsford, Massachusetts, USA) load distributed band [29]. This study focuses on the LUCAS-2 device, whose impact on survival has been thoroughly studied on two of the three largest randomized controlled trials on mechanical chest compression devices [23], [25].

Mechanical chest compression artifacts have larger amplitudes and more harmonics than manual CPR artifacts [30], but their frequency is fixed and known [19]. So the methods to remove manual CPR artifacts have to be recast for piston-driven devices. In the last few years, methods based on comb filters [31], [30], LMS filters [30] and RLS filters [19] have been introduced. Unfortunately these filters followed by the shock decision algorithms of commercial defibrillators were strongly affected by filtering residuals and did not meet AHA goals [30]. Recently, a multi stage algorithm based on two RLS filters and three decision algorithms has been proven to meet the AHA Se/Sp goals [19], albeit with a complex solution and a high computational cost. There is a need to simplify the algorithms that allow an accurate shock decision during piston-driven chest compressions.

This study introduces a new method for shock decision during piston-driven compressions based on an adaptive filter followed by a machine learning algorithm designed to classify the filtered ECG. The machine learning algorithm learns the characteristics of the filtered ECG, including those of the filtering residuals that confound the shock decision algorithms designed for artifact free ECGs. This solution considerably simplifies the best current multistage solution, and improves its accuracy with a much lower computational cost. The paper is organized as follows: the study dataset is described in Section II; feature engineering including CPR artifact filtering, the Stationary Wavelet Transform (SWT) and feature extraction are described in Section III; Section IV describes the architecture used for feature selection and the optimization and evaluation of the classifier. Finally, results, conclusions and discussion are presented in Sections V to VI.

II. STUDY DATASET

The dataset used in this study was collected and annotated for a previous study, so further details on data collection and preparation are available in [30], [19]. In brief, data comes from 263 OHCA patients treated with the LUCAS-2 device by the Oslo and Akershus (Norway) emergency services between July 2012 and December 2013. Signals including ECG and thoracic impedance were recorded using the Lifepak 15 monitor-defibrillator (Physio Control, Redmond, WA, USA), exported to an open matlab format for processing, and resampled to 250 Hz. A 50 Hz notch filter was used to remove powerline interferences from the ECG.

The complete episodes were reviewed and 20-s segments were extracted for studies on mechanical CPR artifact removal. These segments, like the ones shown in Fig. 1, contain an initial 15-s interval during LUCAS-2 use, followed by a 5-s interval without compressions. Ground truth shock/no-shock decisions were adjudicated by consensus between two specialists on cardiac arrest data, a clinical researcher and a biomedical engineer, who inspected the 5-s artifact-free intervals. Nonshockable rhythms included organized rhythms (OR) and asystole (AS), and shockable rhythms were ventricular fibrillation (VF) and ventricular tachycardia (VT). The initial 15-s intervals were used to develop and test the shock decision methods during mechanical compressions. The final dataset contained 1045 20-s segments from 230 patients, whereof 201 were shockable (62 patients) and 844 were

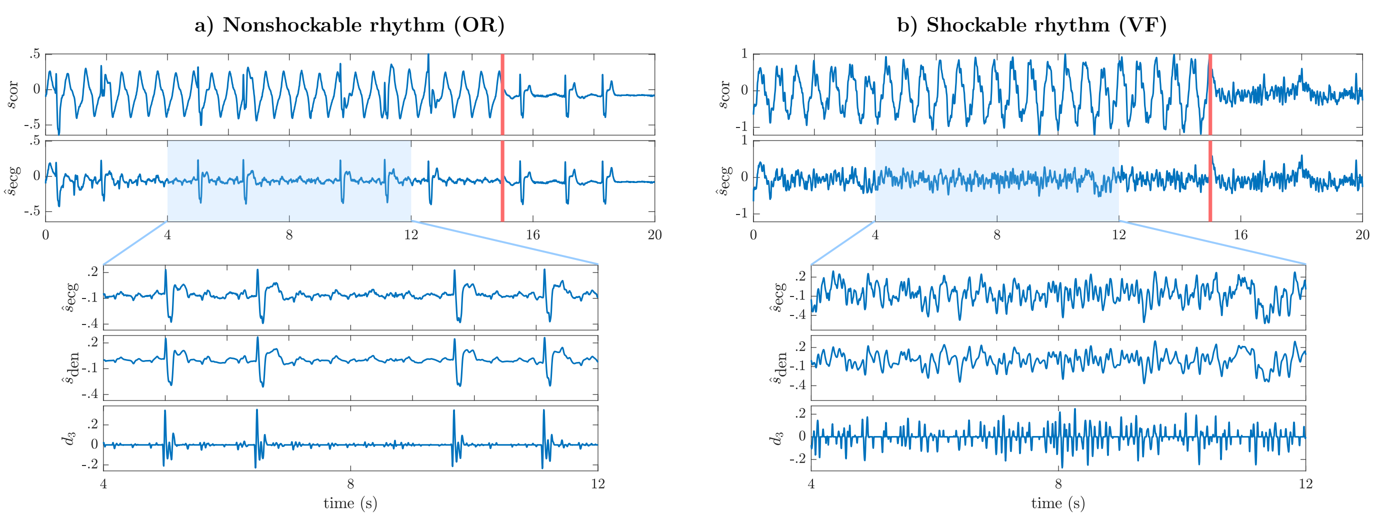

Fig. 1. Two examples of 20-s ECG segments corresponding to a patient presenting a nonshockable rhythm (example a) and to a patient presenting a shockable rhythm (example b). The top panel depicts the corrupt ECG, $s_{\text{cor}}(n)$, and the panel below the ECG after adaptive filtering. The top panel has two intervals, the initial 15-s in which the chest compression artifact is visible, and the last 5-s without artifact in which the underlying rhythm is visible. Finally, the three panels at the bottom zoom in on the 8-s interval used by the shock decision algorithm, and show the filtered ECG, and two significant components obtained from the wavelet analysis of the filtered ECG: the denoised ECG, $\hat{s}_{\text{den}}(n)$, and the detail 3 coefficient, d_3.

nonshockable (209 patients). For an extended description of the dataset and the annotation process consult [30], [19].

III. FEATURE ENGINEERING

Shock/no-shock decision features were extracted in three phases. First an adaptive CPR artifact filter was used to remove chest compression artifacts and obtain the filtered ECG, $\hat{s}_{\text{ecg}}(n)$, then a wavelet analysis provided the denoised signal, $\hat{s}_{\text{den}}(n)$, and the subband decomposition. Finally features were extracted from $\hat{s}_{\text{den}}(n)$ and the subband components. Filtering and wavelet analysis (denoising and the most relevant subband) are illustrated in Fig. 1 for a shockable and a nonshockable rhythm.

A. CPR artifact filtering

During compressions the corrupt ECG, $s_{\text{cor}}(n)$, was assumed to follow an additive artifact model [32], [10]:

$$s_{\text{cor}}(n) = s_{\text{ecg}}(n) + s_{\text{cc}}(n) \tag{1}$$

where $s_{\text{ecg}}(n)$ is the ECG containing the underlying rhythm and $s_{\text{cc}}(n)$ the chest compression artifact. Chest compressions given by the LUCAS-2 device have a constant rate of $100 \pm 2\,\text{min}^{-1}$ ($f_0 = 1.694\,\text{Hz}$), and a depth of 4.0-5.3 cm (depending on the chest height), with a 50% duty cycle at a fixed position on the chest. The pattern of the artifact is therefore quasi-periodic and can be represented as an N term Fourier series of fixed frequency and slowly time varying amplitudes:

$$s_{\text{cc}}(n) = A(n)\sum_{k=1}^{N} a_k(n)\cos(k\omega_0 n) + b_k(n)\sin(k\omega_0 n) \tag{2}$$

where $\omega_0 = 2\pi f_0/f_s$ is the fundamental frequency of the LUCAS-2 device and f_s the sampling frequency. The amplitude envelope $A(n)$ was introduced to differentiate intervals with ($A = 1$) and without ($A = 0$) compressions.

In this work two adaptive methods, LMS [16] and RLS [19] filters, were examined to estimate the time varying in-phase, $a_k(n)$, and quadrature, $b_k(n)$, amplitudes. For each filter two degrees of freedom were adjusted: N the number of harmonics of the artifact model and μ/λ the coarseness of the filter [16], [19]. N can also be interpreted as the order of the filter. It determines the number of filter coefficients, which is $2N$ since there are a cuadrature and in-phase coefficient per harmonic. The coarseness of the filter is either μ, the step size of the LMS filter, or λ the forgetting factor of the RLS filter. Both these values offer a compromise between tracking capabilities and misadjustment and stability of the filter. A small forgetting factor in the RLS filter or a large step size in the LMS filter mean that a bigger change can occur in the filter coefficients for each new sample, i.e. a more coarse filter [16], [19]. This produces adaptive filters that follow changes in the input signal better, but also that filter coefficients can increase without bound if changes accumulate, resulting in an unstable filter.

B. Stationary Wavelet Transform

Feature extraction was based on the wavelet decomposition of the filtered ECG. Previous studies on OHCA rhythm classification have successfully applied feature extraction based on the Discrete Wavelet Transform (DWT) [33]. We chose instead a Stationary Wavelet Transform (SWT) approach [34], [35]. Unlike the DWT, the SWT is shift-invariant and better suited for edge detection, fiducial point location or denoising [36], [37]. The SWT is based on the same dyadic decomposition as the DWT, a typical architecture is shown in Fig. 2. Shift invariance is achieved by upsampling the filters instead of sub-sampling the signal at each level of decomposition. The DWT scaling and wavelet filters for signal decomposition, $g_0(n)$ and $h_0(n)$, are a pair of quadrature mirror lowpass and highpass filters. The filters at stage j are obtained by upsampling the original filters by a factor of 2^j, that is:

$$h_j(n) = (h_0 \uparrow 2^j)(n) = \begin{cases} h_0\left(\frac{n}{2^j}\right) & n = k \cdot 2^j \\ 0 & n \neq k \cdot 2^j \end{cases} \tag{3}$$

The detail, $d_j(n)$, and approximation, $a_j(n)$, coefficients at all levels from $j = 1, \ldots, J$ are then recursively obtained:

$$a_0(n) = \hat{s}_{\text{ecg}}(n) \tag{4}$$
$$a_{j+1}(n) = g_j(n) * a_j(n) \tag{5}$$
$$d_{j+1}(n) = h_j(n) * a_j(n) \tag{6}$$

where $*$ stands for convolution. The filter coefficients depend on the mother wavelet used. In this work a Daubechies-2 mother wavelet was adopted because it produced the best results (see supplementary materials). The filters for reconstruction are obtained by time reversion: $\overline{g}_j(n) = g_j(-n)$ and $\overline{h}_j(n) = h_j(-n)$. Therefore, the original signal can be reconstructed from the level J coefficients (ISWT) by recursively applying [35]:

$$a_{j-1}(n) = \tfrac{1}{2}\big(\overline{g}_j(n) * a_j(n) + \overline{h}_j(n) * d_j(n)\big) \tag{7}$$

from $j = J, \ldots, 1$.

Eight decomposition levels ($J = 8$) were used to generate nine sets of coefficients, a_8 and $d_8, \ldots d_1$. A signal interval of $M = 2048$ samples was analyzed, for a sampling frequency of $f_s = 250\,\text{Hz}$ it included the 8-s interval of the filtered ECG highlighted in Fig. 1. Since the analysis is based on a dyadic decomposition in which the available bandwidth is split in two at each successive decomposition level, and considering that the bandwidth of interest in defibrillators is commonly between 0.5-30 Hz, only detail coefficients d_3-d_8 were kept and d_1, d_2 and a_8 were set to zero [33]. A soft denoising was then applied to d_3-d_8 using a fixed treshold, ρ, and single estimation of level noise based on first-level detail coefficients [38]:

$$\rho = 1.483 \cdot \text{MAD}(d_1)\sqrt{2\ln M} \tag{8}$$

where $\mathrm{MAD}(d_1)$ is the median absolute deviation of d_1. Finally, the denoised d_3-d_8 coefficients were used in equation (7) to reconstruct $\hat{s}_{\mathrm{den}}(n)$ in the $0.5 - 31.25\,\mathrm{Hz}$ frequency range.

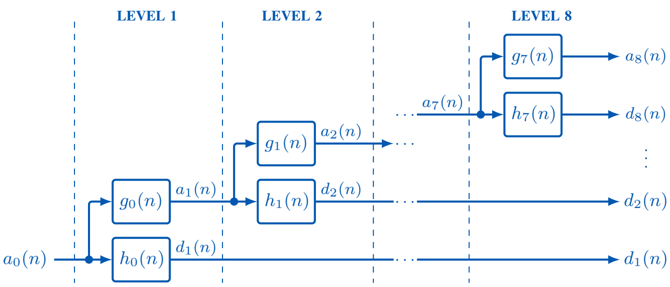

Fig. 2. SWT implementation for eight levels of decomposition.

C. Feature extraction

The denoised signal, $\hat{s}_{\mathrm{den}}(n)$, and the detail coefficients, d_3-d_8, were used to obtain a set of 38 features for the shock decision algorithm, selected from the literature on the topic [33], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], [49], [50], [51].

The first 18 features were the interquartile range (IQR), first quartile (FQR) and the sample entropy (SampEn) of the detail coefficients d_3-d_8 [33]. The remaining 20 features were computed from $\hat{s}_{\mathrm{den}}(n)$, and constitute a comprehensive set of features from the available methods on shock decision algorithms that included time domain, frequency domain and signal complexity characterizations of the ECG. The extracted features were TCSC [39], Expmod [40], MAV [41], count1-count3 [42], x1-x2 [43], bCP and bWT [44], A1-A3 [45], VFleak [46], SampEn [47], [48], the number of peaks in the 8-s interval (Np) [33], HILB [51], CM [50], Kurt and Frqbin [49]. A detailed description can be found in the references given above, and a Matlab implementation of the features derived from the denoised ECG is available in: https://github.com/FelipeURJC/ohca-vs-public-dbs/tree/master/ecg parameters computation/parameters.

IV. Architecture of the Model and Evaluation

A nested cross-validation (CV) architecture was used for feature selection, and classifier hyperparameter optimization, and model assessment, as shown in Fig. 3. In the inner loop features were selected using a wrapper approach in a 5-fold CV [52]. In the outer loop, 10-fold CV was used for hyperparameter optimization and model assessment. Both inner and outer folds were partitioned patient-wise in a quasi-stratified way, by ensuring that the shock/no-shock case prevalences matched to at least 85% those of the whole dataset. The performance of the method was evaluated by comparing the shock/no-shock decisions of the classifier with ground truth labels in the outer test set. The following metrics were computed: Se, Sp, accuracy (Acc) and the Balanced Accuracy (BAC), i.e. the mean value of Se and Sp.

A. Feature selection

In the inner loop, a PTA$(4,3)$ (plus 4, take away 3) feature selection algorithm was used [53], [54]. The criterion to include or exclude a feature within each inner loop was the maximization of the BAC of a Linear Discriminant Analysis (LDA) classifier [33], see inner loop in Fig. 3. BAC values were obtained by comparing the shock/no-shock decisions obtained through the LDA classifier with ground truth labels of the inner test set. At each step of the PTA(4,3) four features were included in the model using Sequential Forward Selection, and then three were removed from the model using Sequential Backward Selection. The feature selection method was run until K features were included, several values of K were tested in the experiments. A wrapper-based approach was adopted in order to address feature dependencies and hence select K features that altogether are the most discriminative ones. Finally, we chose the PTA algorithm to avoid the nesting effects of sequential feature selection [53].

B. Shock Decision Algorithm

The decision algorithm was designed in the outer loop, deploying a Support Vector Machine (SVM) classifier with a Gaussian kernel [55]. Features were standardized to zero mean and unit variance using the data in the training set, and the K features from the inner feature selection loop were used. This resulted in a training set of instance-label pairs $\{(\boldsymbol{x}_1, y_1), ..., (\boldsymbol{x}_n, y_n)\} \in \mathbb{R}^K \times \{\pm 1\}$, where $y_i = 1$ for shockable and $y_i = -1$ for nonshockable rhythms. The decision function of the SVM is found by solving the following maximization problem [55]:

$$W(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j \exp(-\gamma \|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2) \quad (9)$$

$$\text{s.t.}: \; 0 \le \alpha_i \le C \quad \forall i, \quad \text{and} \quad \sum_{i}^{N} \alpha_i y_i = 0 \quad (10)$$

where the α_i Lagrange multipliers are non-zero only for N_s support vectors, C is the soft margin parameter and γ the

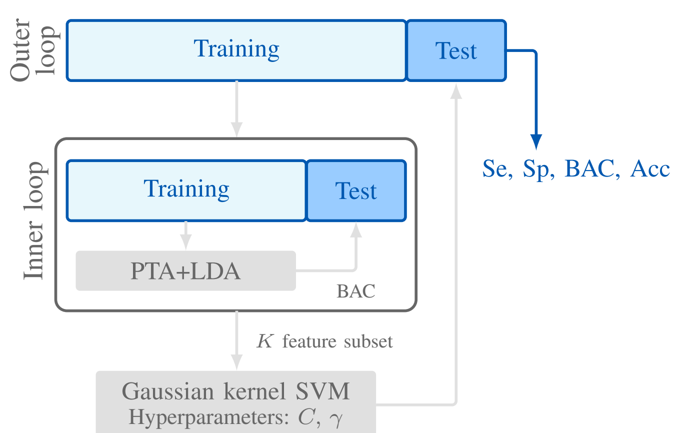

Fig. 3. Nested cross-validation architecture used for feature selection and for model optimization and evaluation.

width of the gaussian kernel. Once the support vectors are determined the decision function is:

$$f(\boldsymbol{x}) = \text{sign}\left[\sum_{i=1}^{N_s} \alpha_i y_i \exp(-\gamma\|\boldsymbol{x}-\boldsymbol{x}_i\|^2) + b\right] \qquad (11)$$

where the threshold b is determined in the optimization phase. A rhythm will be classified as shockable for $f(\boldsymbol{x}) = 1$ or nonshockable for $f(\boldsymbol{x}) = -1$.

Hyperparameter optimization for a gaussian kernel SVM involves selecting γ and C, and was carried out using the *libsvm* library [56]. The soft margin parameter C represents a trade-off between maximizing the margin and minimizing errors in the training data, and γ controls the flexibility of the decision boundary [57]. The values of C and γ that maximized the BAC were determined in the outer loop doing a 25×25 logarithmic grid search in the ranges $10^{-1} \leq C \leq 10^2$ and $10^{-3} \leq \gamma \leq 10^1$, respectively. The nested CV procedure was repeated 50 times to estimate the statistical distributions of the performance metrics that will be reported as mean (standard deviation).

V. Results and Discussion

This section provides the main results for the shock decision algorithm; additional results are given in the supplementary materials and referenced in the manuscript. First the LMS/RLS filter was optimized; then the effect of two variables were analyzed, the number of features used by the classifier (K), and the length of the analysis segment used for the shock/no-shock decision (L). Finally the results are compared to all available solutions for shock decisions during piston-driven chest compressions. The results are reported for the C/γ pair with best average BAC in the 50 repetitions of the outer CV loop.

A. *CPR artifact filter configuration and processing times*

Fig. 4 shows the mean values of the BAC obtained in the 50 random repetitions of the nested CV procedure for different configurations of the LMS and RLS filters, using an interval of $L = 8\,\text{s}$ for feature extraction and an SVM classifier with $K = 6$ features. Both filters showed near-optimal performance with a BAC above 96.5% for a wide range of configurations, that is, for different filter orders (N) and coarseness levels (μ, λ): $N \geq 10$ and $\mu \sim$ 3-12 $\cdot 10^{-3}$ for the LMS filter and $N \geq 10$ and $\lambda \sim$ 0.970-0.990 for the RLS filter. The accuracy of the solution is not very sensitive to the CPR artifact filter, so filters can be considerably simplified by decreasing their order N to reduce the computational cost. Table I shows the distribution of the performance metrics and the average computation time for different filter orders. The filters were configured at their optimal coarsness, $\mu = 8 \cdot 10^{-3}$ and $\lambda = 0.99$, as shown in Fig. 4. The computation time t_1 is the time required to suppress the CPR artifact and t_2 includes the wavelet decomposition, feature calculations ($K = 6$), and the decision of the SVM classifier obtained through Eq (11). All calculations were done in Matlab on an i7 3.2 GHz single-core processor and 16 GB of memory.

AHA performance goals were met with the RLS and LMS filters with as few as $N = 5$ harmonics, but best results were obtained with $N = 20$, as shown in Table I. For $N = 5$ the computational demands of the complete algorithm were very low, 16 ms for the LMS or 38 ms for the RLS filter. Feature extraction including SWT/ISWT analysis and denoising consumed on average 6 ms, so the LMS filter is computationally very cheap and its computational cost negligible regardless of its order, it uses up 10 ms for $N = 5$, and 18 ms for $N = 30$. The RLS filter has a greater computational cost that increases considerably with its order, from 30 ms for $N = 5$ to over 140 ms for $N = 30$. This excessive computational cost is caused by the RLS recursion

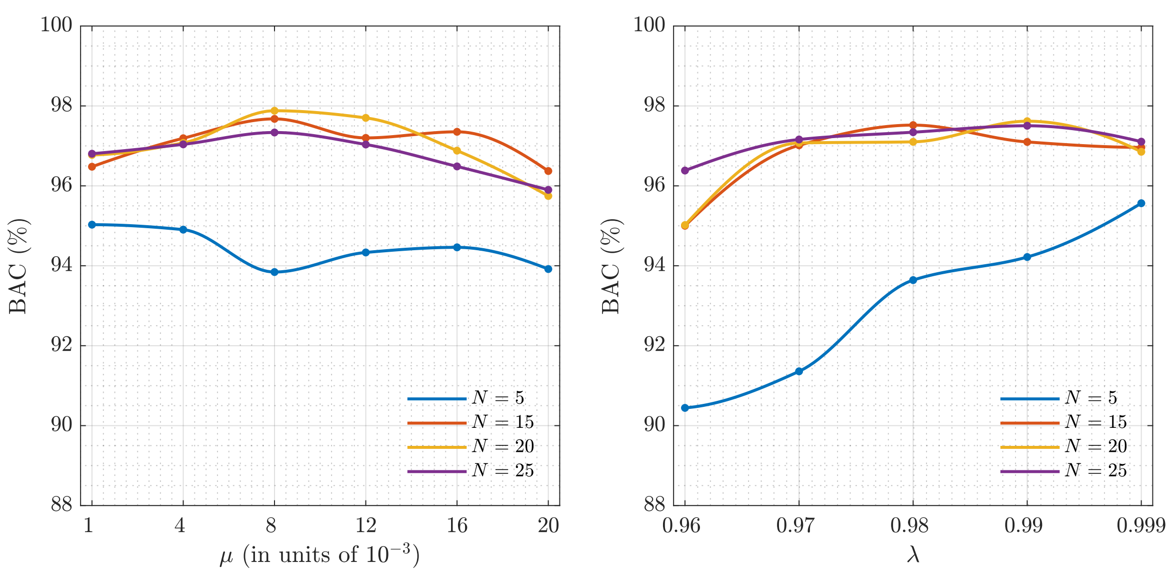

Fig. 4. The mean values of BAC obtained in the 50 repetitions of the nested CV procedure when a LMS (left) or a RLS (right) filter is used to remove the CPR artifact. The performance is given as a function of the coarseness (λ, μ). of the filter for 4 significant values of the filter order, N

TABLE I
SHOCK DECISION ACCURACIES AND PROCESSING TIMES FOR FILTERING (t_1) AND SHOCK DECISION (t_2) FOR DIFFERENT FILTER ORDERS.

	RLS, $\lambda = 0.99$					LMS, $\mu = 8 \cdot 10^{-3}$				
N	Se (%)	Sp (%)	BAC (%)	Acc (%)	t_1/t_2 (ms)	Se (%)	Sp (%)	BAC (%)	Acc (%)	t_1/t_2 (ms)
4	90.6 (1.1)	94.3 (0.7)	92.5 (0.7)	93.6 (0.7)	30/5	92.3 (0.8)	94.6 (0.6)	93.5 (0.5)	94.2 (0.5)	10/6
5	92.8 (1.2)	95.6 (0.6)	94.2 (0.7)	95.1 (0.5)	32/6	91.8 (1.2)	95.9 (0.3)	93.8 (0.7)	95.1 (0.4)	10/6
10	95.4 (0.7)	97.9 (0.4)	96.7 (0.4)	97.4 (0.4)	37/5	96.0 (0.4)	98.6 (0.3)	97.3 (0.3)	98.1 (0.3)	14/7
15	95.8 (0.7)	98.4 (0.3)	97.1 (0.4)	97.9 (0.3)	50/5	96.7 (0.4)	98.6 (0.4)	97.7 (0.3)	98.3 (0.3)	15/7
20	**97.0 (0.5)**	**98.3 (0.2)**	**97.6 (0.2)**	**98.0 (0.2)**	**72/6**	**97.5 (0.4)**	**98.2 (0.4)**	**97.9 (0.3)**	**98.1 (0.3)**	**16/5**
25	96.6 (0.5)	98.5 (0.3)	97.5 (0.2)	98.1 (0.3)	96/4	96.8 (0.4)	97.9 (0.3)	97.3 (0.3)	97.7 (0.3)	17/5
30	96.9 (0.6)	98.0 (0.4)	97.4 (0.4)	97.8 (0.3)	147/6	97.5 (0.4)	97.9 (0.3)	97.7 (0.2)	97.8 (0.3)	18/7

formula for the gain matrix which involves $2N \times 2N$ matrix multiplications for each signal sample [19]. The RLS filter has been shown to be more effective than the LMS filter to remove piston-driven compression artifacts when shock decision algorithms from commercial defibrillators are used in the classification stage [58], [19] (see also Table III). Shock decision algorithms in commercial defibrillators are designed to classify artifact free ECGs, so an effective suppression of the CPR artifact is critical. This is also important if the filtered ECG ($\hat{s}_{\text{ecg}}$ in Fig. 1 and Fig. 7) is shown in the screen of the monitor-defibrillator to serve as a decision support signal for the emergency clinician. However, our results show that the design of CPR artifact filters can be relaxed when a properly designed machine learning algorithm trained with the filtered ECG is used for classification. This is probably because the classification algorithm now learns the characteristics of filtering residuals that confound the shock decision algorithms of commercial defibrillators.

For all the analyses hereafter an LMS filter with $\mu = 8 \cdot 10^{-3}$ and $N = 20$ was used.

B. Classification features and feature ranking

One of the pivotal aspects of a machine learning algorithm is the design of the classification features. The method proposed includes features extracted from the d_3-d_8 denoised SWT components and their reconstructed signals. Table II shows the ranking of the features by the number of times they were selected using the PTA(4,3) feature selection scheme in the inner loop and 50 random repetitions of the outer CV loop ($50 \times 10 = 500$ feature selection loops). This ranking was obtained for a solution with $K = 6$ features. The features with the best ranking are a mixture of those derived from the detail coefficients and from the denoised signal, and represent a variety of signal analysis approaches that comprise signal regularity/complexity (SampEn, CM, Frqbin) [59], [50], [49], spectral analysis (VFleak, A1-3, bWT) [60], [45], [44], time domain features (MAV, Np, count2) [41], [33], [42], or the sample distributions of the denoised signal (Kurt) and its detail coefficients FQR/IQR [33]. Additional results for the discriminative power of the features using ROC curve analysis are available in the supplementary materials.

Fig. 5 shows the accuracies (balanced and absolute) of the shock decision system as a function of features allowed in the SVM. For a good accuracy the number of features in the classifier must be between 3 and 7, which gives an Acc and BAC above 97.8%. A classifier with fewer features presented lower BAC and Acc, with a more negative impact on Acc. This

TABLE II
FEATURES RANKED BY N_f, THE NUMBER OF TIMES THEY WERE SELECTED IN THE 500 INNER LOOPS.

RLS filter		LMS filter	
Feature	N_f	Feature	N_f
SampEn, d_3	500	SampEn, d_3	500
FQR, d_7	397	VFleak	321
VFleak	337	FQR, d_7	236
A1	275	IQR, d_7	217
CM	255	A2	183
Kurt	248	Kurt	157
A2	207	A3	148
bWT	146	FQR, d_6	119
A3	86	Np	102
IQR, d_7	65	FQR, d_8	85
MAV	60	CM	73
Frqbin	52	count2	67

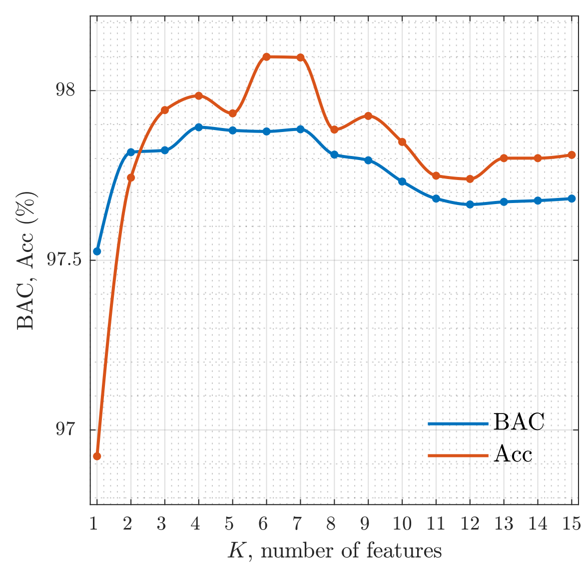

Fig. 5. Mean values of BAC and Acc as a function of the number of features, K, used in the classifier.

means that the most prevalent class, the Sp for nonshockable rhythms, is negatively affected by using a simpler classifer. Adding more than 7 features sligthly reduces both accuracies, and makes the classifier more complex.

C. Duration of the analysis segment

Fig. 6 shows how the performance metrics change as the analysis segment is shortened. The perfomance of the algorithm stabilizes at near-optimal values for analysis segments longer than 4 s, and drops if shorter segments are used. However, for segments as short as 2 s the algorithm still meets the minimum AHA recommendations for Se and Sp, with values of 96.5 (94.9 – 97.6) and 96.0 (95.1 – 96.7), respectively. Studies that have developed ad-hoc algorithms for cardiac arrest data have reported minimum segment lengths for an accurate analysis around 3-4 s, both for the analysis of the ECG without CPR artifacts [61], [47] or after suppression of manual CPR artifacts [43]. Previous studies on shock decision during piston-driven chest compressions relied on shock decision algorithms of commercial defibrillators. These algorithms require analysis segments in excess of 5 s in most devices [62]. For instance, in two previous studies on shock decision during mechanical CPR the analysis segment was either 6 s or 9 s long, because the algorithm applied a majority vote to three consecutive 3-s analysis subsegments [30], [19]. Reducing the length of the analysis segments is not critical during compressions, since CPR therapy is not interrupted for the analysis. However, if a rhythm transition analysis is to be performed during CPR [63] short intervals would permit a more accurate time-location of transitions between shockable and nonshockable rhythms, and a reduction of computational burden.

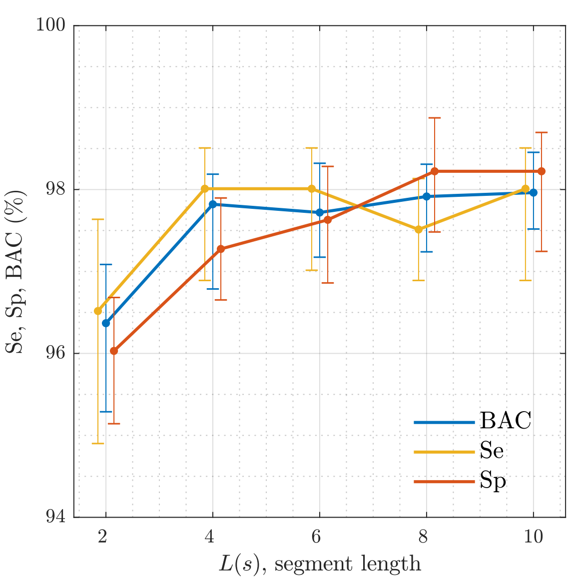

Fig. 6. Distribution of the performance metrics as a function of the length of the analysis segment (L). The graph shows the median values and the 2.5-97-5 percentile range for Se, Sp and BAC.

D. Discussion on the near-optimal solution

The accuracy for the (near)-optimal solutions using an RLS and an LMS filter (see Table I) are compared in Table III to the available methods for shock decision during piston-driven compressions. Feature extraction was done with $L = 8\,\mathrm{s}$ and an SVM with $K = 6$ features was used. The optimal (C, γ) pairs for the SVM were $(17.8 \cdot 10^{-2}, 6.8 \cdot 10^{-2})$ and $(3.162, 1 \cdot 10^{-2})$ for the LMS and RLS filter based solutions, respectively.

The multistage solution introduced in [19] was the most accurate shock decision algorithm for mechanical devices proposed to date. As shown in Table III, the machine learning approach proposed in this study increases the BAC of single filter solutions by over 5-points, and that of the multistage solution by 3-points, and increases the sensitivity substantially, making the solution very reliable for the detection of shockable rhythms. The overall accuracy is also increased by around 1-point, which is a considerable increase because the multistage solution had an overall accuracy of 96.9 %. A 1-point increase from that baseline means that around 30 % of the errors are now correctly classified. Very importantly, this improvement was achieved together with a drastic reduction of the computation demands of the algorithms. For a solution based on the LMS filter the mean processing time per 8-s segments was 21 ms, an over five fold improvement when compared to the 110 ms required by the multistage solution. This reduction is very important in defibrillators with scarce computation resources.

Finally, Fig. 7 shows three illustrative examples of misclassified segments both for shockable and nonshockable rhythms. In the two examples of nonshockable rhythms the denoised signal and the d_3 detail coefficient (best features) show a disorganized signal, fast in the case of the AS example (middle) and slower for the OR (top). These disorganized signals are interpreted as shockable rhythm by the SVM classifier. In the example of the missed VF, the filter is unable to remove the spiky artifact introduced by the mechanical device at each compression, and these spikes confound the decision algorithm. In any case misclassifications were very

TABLE III
COMPARISON TO PREVIOUS METHODS USING THE SAME DATA.

	Performance metric			
Method	Se (%)	Sp (%)	BAC (%)	Acc (%)
1-Stg, Dfb†				
LMS [30]	98.6 (1.0)	84.0 (1.8)	91.3 (1.2)	86.8 (1.6)
RLS [19]	98.1 (1.0)	87.0 (1.8)	92.5 (1.1)	89.1 (1.5)
Comb [30]	97.1 (2.0)	84.3 (1.8)	90.7 (1.3)	86.8 (1.6)
M-Stg, Dfb‡				
LMS [19]	94.4 (3.0)	93.2 (1.2)	93.8 (1.6)	93.4 (1.1)
RLS [19]	91.7 (6.0)	98.1 (1.1)	94.9 (2.6)	96.9 (0.9)
Comb [19]	88.8 (6.0)	96.6 (1.7)	92.7 (2.4)	95.1 (1.1)
1-Stg, SVM				
LMS	97.5 (0.4)	98.2 (0.4)	97.9 (0.3)	98.1 (0.3)
RLS	97.0 (0.5)	98.3 (0.2)	97.6 (0.2)	98.0 (0.2)

† Single stage filtering, shock decision of a commercial defibrillator
‡ Multistage filtering, shock decision of a commercial defibrillator

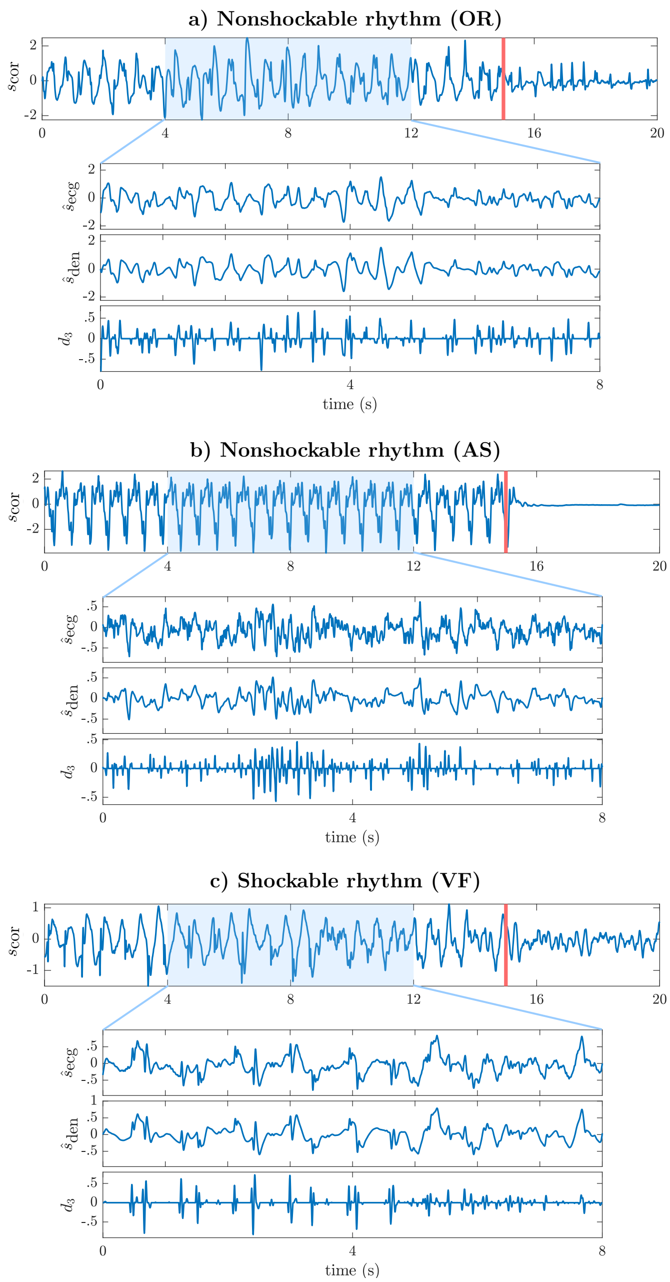

Fig. 7. Three examples of misclassified segments. Panels a and b depict nonshockable rhythms, OR and AS, respectively, while panel c represents a shockable (VF) rhythm. From top to bottom, each panel shows the 20-s ECG segment, and the filtered ECG, the denoised ECG and the detail 3 coefficient of the 8-s interval used by the shock decision algorithm.

few, around 15 for nonshockable rhythms (Sp $\sim$ 98.2 %), and around 5 for VF (Se $\sim$ 97.5 %).

VI. Conclusions

This study introduces a machine learning algorithm for shock decisions during piston-driven chest compressions. The algorithm improves the accuracy of the best known solutions to date by 3 points in BAC with an additional 5-fold reduction in computational cost. This makes this solution very accurate and efficient. There are two main reasons for these advances. First, the feature extraction phase based on the stationary wavelet analysis resulted in new and improved discriminating features. Second, extracting the features after removing the CPR artifact and feeding those features to the SVM improves the accuracy considerably, because the machine learning algorithm is able to learn the characteristics of filtering residuals. Our results show that this approach allows relaxing the characteristics of the compression artifact filters.

The main limitations of this study are associated with the data. The dataset came from a single type of monitor-defibrillator, so the methods may need adjusting to encompass data from other devices with different ECG acquisition characteristics like bandwidth, sampling rates or A/D resolution. Furthermore, the data were compiled from a single emergency service and the LUCAS-2 device may be used differently across emergency services, that may also enforce different resuscitation protocols. Those differences may result in chest compression artifacts with different characteristics. Finally, the (near)-optimal solutions presented in Table I were obtained following a training/validation data partition given the amount of samples available. If more data were available the results should be confirmed using an independent test set.

References

[1] G. Perkins *et al.*, "European resuscitation council guidelines for resuscitation 2015: Section 2. Adult basic life support and automated external defibrillation," *Resuscitation*, vol. 95, pp. 81–99, 2015.

[2] K. B. Kern *et al.*, "Importance of continuous chest compressions during cardiopulmonary resuscitation: improved outcome during a simulated single lay-rescuer scenario," *Circulation*, vol. 105, no. 5, pp. 645–649, 2002.

[3] C. Vaillancourt *et al.*, "The impact of increased chest compression fraction on return of spontaneous circulation for out-of-hospital cardiac arrest patients not in ventricular fibrillation," *Resuscitation*, vol. 82, no. 12, pp. 1501–1507, 2011.

[4] S. Ruiz de Gauna *et al.*, "Rhythm analysis during cardiopulmonary resuscitation: past, present, and future," *Biomed. Res. In.*, vol. 2014, 2014.

[5] J. Ruiz *et al.*, "Feasibility of automated rhythm assessment in chest compression pauses during cardiopulmonary resuscitation," *Resuscitation*, vol. 84, no. 9, pp. 1223–1228, 2013.

[6] U. Ayala *et al.*, "Fully automatic rhythm analysis during chest compression pauses," *Resuscitation*, vol. 89, pp. 25–30, 2015.

[7] Y. Li *et al.*, "Identifying potentially shockable rhythms without interrupting cardiopulmonary resuscitation," *J. Crit. Care Med.*, vol. 36, no. 1, pp. 198–203, 2008.

[8] Y. Li *et al.*, "An algorithm used for ventricular fibrillation detection without interrupting chest compression," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 1, pp. 78–86, 2012.

[9] H. Kwok *et al.*, "Adaptive rhythm sequencing: A method for dynamic rhythm classification during CPR," *Resuscitation*, vol. 91, pp. 26–31, Jun 2015.

[10] S. O. Aase *et al.*, "CPR artifact removal from human ECG using optimal multichannel filtering," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 11, pp. 1440–1449, 2000.

[11] J. Husoy *et al.*, "Removal of cardiopulmonary resuscitation artifacts from human ECG using an efficient matching pursuit-like algorithm," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1287–1298, 2002.

[12] J. Eilevstjønn *et al.*, "Feasibility of shock advice analysis during CPR through removal of CPR artefacts from the human ECG," *Resuscitation*, vol. 61, no. 2, pp. 131–141, 2004.

[13] K. Rheinberger *et al.*, "Removal of CPR artifacts from the ventricular fibrillation ECG by adaptive regression on lagged reference signals," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 130–137, 2008.

[14] J. Ruiz *et al.*, "Cardiopulmonary resuscitation artefact suppression using a Kalman filter and the frequency of chest compressions as the reference signal," *Resuscitation*, vol. 81, no. 9, pp. 1087–1094, 2010.

[15] T. Werther *et al.*, "CPR artifact removal in ventricular fibrillation ECG signals using Gabor multipliers," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 2, pp. 320–327, 2009.
[16] U. Irusta *et al.*, "A least mean-square filter for the estimation of the cardiopulmonary resuscitation artifact based on the frequency of the compressions," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 4, pp. 1052–1062, 2009.
[17] E. Aramendi *et al.*, "Suppression of the cardiopulmonary resuscitation artefacts using the instantaneous chest compression rate extracted from the thoracic impedance," *Resuscitation*, vol. 83, no. 6, pp. 692–698, 2012.
[18] Y. Gong *et al.*, "An enhanced adaptive filtering method for suppressing cardiopulmonary resuscitation artifact," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 2, pp. 471–478, 2017.
[19] I. Isasi *et al.*, "A multistage algorithm for ECG rhythm analysis during piston driven mechanical chest compressions," *IEEE Transactions on Biomedical Engineering, Online version, DOI: 10.1109/TBME.2018.2827304*, 2018.
[20] U. Ayala *et al.*, "Are dual-channel methods as accurate as multi-channel methods to suppress the CPR artifact?" in *Computing in Cardiology, 2011*. IEEE, 2011, pp. 509–512.
[21] S. R. de Gauna *et al.*, "A method to remove cpr artefacts from human ecg using only the recorded ecg," *Resuscitation*, vol. 76, no. 2, pp. 271–278, 2008.
[22] R. E. Kerber *et al.*, "Automatic external defibrillators for public access defibrillation: recommendations for specifying and reporting arrhythmia analysis algorithm performance, incorporating new waveforms, and enhancing safety," *Circulation*, vol. 95, no. 6, pp. 1677–1682, 1997.
[23] S. Rubertsson *et al.*, "Mechanical chest compressions and simultaneous defibrillation vs conventional cardiopulmonary resuscitation in out-of-hospital cardiac arrest: the linc randomized trial," *Jama*, vol. 311, no. 1, pp. 53–61, 2014.
[24] L. Wik *et al.*, "Manual vs. integrated automatic load-distributing band cpr with equal survival after out of hospital cardiac arrest. the randomized circ trial," *Resuscitation*, vol. 85, no. 6, pp. 741–748, 2014.
[25] G. D. Perkins *et al.*, "Mechanical versus manual chest compression for out-of-hospital cardiac arrest (paramedic): a pragmatic, cluster randomised controlled trial," *The Lancet*, vol. 385, no. 9972, pp. 947–955, 2015.
[26] G. Putzer *et al.*, "LUCAS compared to manual cardiopulmonary resuscitation is more effective during helicopter rescue. A prospective, randomized, cross-over manikin study," *Am. J. Emerg. Med.*, vol. 31, no. 2, pp. 384–389, 2013.
[27] A. I. Larsen *et al.*, "Cardiac arrest with continuous mechanical chest compressions during percutaneous coronary intervention: a report on the use of the LUCAS device," *Resuscitation*, vol. 75, no. 3, pp. 454–459, 2007.
[28] K. Sunde *et al.*, "Quality of mechanical, manual standard and active compression–decompression CPR on the arrest site and during transport in a manikin model," *Resuscitation*, vol. 34, no. 3, pp. 235–242, 1997.
[29] J. Soar *et al.*, "European resuscitation council guidelines for resuscitation 2015: Section 3. Adult advanced life support," *Resuscitation*, vol. 95, pp. 100–147, 2015.
[30] E. Aramendi *et al.*, "Filtering mechanical chest compression artefacts from out-of-hospital cardiac arrest data," *Resuscitation*, vol. 98, pp. 41–47, 2016.
[31] J. Sullivan *et al.*, "A digital filter can effectively remove mechanical chest compression artifact," *Resuscitation*, vol. 85, p. S41, 2014.
[32] A. Langhelle *et al.*, "Reducing CPR artefacts in ventricular fibrillation in vitro," *Resuscitation*, vol. 48, no. 3, pp. 279–291, 2001.
[33] A. B. Rad *et al.*, "Ecg-based classification of resuscitation cardiac rhythms for retrospective data analysis," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 10, pp. 2411–2418, 2017.
[34] M. Holschneider *et al.*, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets*. Springer, 1990, pp. 286–297.
[35] J. E. Fowler, "The redundant discrete wavelet transform and additive noise," *IEEE Signal Processing Letters*, vol. 12, no. 9, pp. 629–632, Sep. 2005.
[36] M. Merah *et al.*, "R-peaks detection based on stationary wavelet transform." *Computer methods and programs in biomedicine*, vol. 121, pp. 149–160, Oct. 2015.
[37] M. Cesari *et al.*, "A new wavelet-based ecg delineator for the evaluation of the ventricular innervation." *IEEE journal of translational engineering in health and medicine*, vol. 5, p. 2000215, 2017.
[38] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, May 1995.
[39] M. A. Arafat *et al.*, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," *Signal, Image and Video Processing*, vol. 5, no. 1, pp. 1–10, 2011.
[40] A. Amann *et al.*, "Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators," *Biomedical engineering online*, vol. 4, no. 1, p. 60, 2005.
[41] E. M. A. Anas *et al.*, "Sequential algorithm for life threatening cardiac pathologies detection based on mean signal strength and emd functions," *Biomedical engineering online*, vol. 9, no. 1, p. 43, 2010.
[42] I. Jekova and V. Krasteva, "Real time detection of ventricular fibrillation and tachycardia," *Physiological measurement*, vol. 25, no. 5, p. 1167, 2004.
[43] U. Ayala *et al.*, "A reliable method for rhythm analysis during cardiopulmonary resuscitation," *Biomed. Res. In.*, vol. 2014, 2014.
[44] U. Irusta *et al.*, "A high-temporal resolution algorithm to discriminate shockable from nonshockable rhythms in adults and children," *Resuscitation*, vol. 83, no. 9, pp. 1090–1097, 2012.
[45] S. Barro *et al.*, "Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: a diagnostic system," *Journal of biomedical engineering*, vol. 11, no. 4, pp. 320–328, 1989.
[46] S. Kuo, "Computer detection of ventricular fibrillation," *Proc. of Computers in Cardiology, IEEE Comupter Society*, pp. 347–349, 1978.
[47] C. Figuera *et al.*, "Machine learning techniques for the detection of shockable rhythms in automated external defibrillators." *PloS one*, vol. 11, p. e0159654, 2016.
[48] F. Alonso-Atienza *et al.*, "Detection of life-threatening arrhythmias using feature selection and support vector machines," *IEEE Trans. Biomed. Eng*, vol. 61, no. 3, pp. 832–840, 2014.
[49] I. Jekova, "Shock advisory tool: Detection of life-threatening cardiac arrhythmias and shock success prediction by means of a common parameter set," *Biomedical Signal Processing and Control*, vol. 2, no. 1, pp. 25–33, 2007.
[50] X.-S. Zhang *et al.*, "Detecting ventricular tachycardia and fibrillation by complexity measure," *IEEE Transactions on biomedical engineering*, vol. 46, no. 5, pp. 548–555, 1999.
[51] A. Amann *et al.*, "Detecting ventricular fibrillation by time-delay methods," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 1, pp. 174–177, 2007.
[52] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.
[53] S. Strearns, "On selecting features for pattern classifiers," *Proc. ICPR, 1976*, 1976.
[54] A. R. Webb, *Statistical pattern recognition*. John Wiley & Sons, 2003.
[55] B. Scholkopf *et al.*, "Comparing support vector machines with gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2758–2765, Nov. 1997.
[56] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, p. 27, 2011.
[57] A. Ben-Hur and J. Weston, "A user's guide to support vector machines," in *Data mining techniques for the life sciences*. Springer, 2010, pp. 223–239.
[58] I. Isasi *et al.*, "Removing piston-driven mechanical chest compression artefacts from the ECG," in *Proc. Comput, Cardiol.*, vol. 44, 2017.
[59] B. Chicote *et al.*, "Application of entropy-based features to predict defibrillation outcome in cardiac arrest," *Entropy*, vol. 18, no. 9, p. 313, 2016.
[60] S. Kuo and R. Dillman, "Computer detection of ventricular fibrillation," in *Computers in Cardiology*, 1978, pp. 2747–2750.
[61] J.-P. Didon *et al.*, "Shock advisory system with minimal delay triggering after end of chest compressions: accuracy and gained hands-off time," *Resuscitation*, vol. 82, pp. S8–S15, 2011.
[62] D. Snyder and C. Morgan, "Wide variation in cardiopulmonary resuscitation interruption intervals among commercially available automated external defibrillators may affect survival despite high defibrillation efficacy." *Crit Care Med*, vol. 32, no. 9 Suppl, pp. S421–S424, Sep 2004.
[63] A. B. Rad *et al.*, "An automatic system for the comprehensive retrospective analysis of cardiac rhythms in resuscitation episodes." *Resuscitation*, vol. 122, pp. 6–12, Jan. 2018.