Life-Threatening Ventricular Arrhythmia Detection With Personalized Features

Ping Cheng and Xiaodai Dong

2017

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 6, 2017, accepted June 23, 2017, date of publication July 4, 2017, date of current version August 8, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2723258

Life-Threatening Ventricular Arrhythmia Detection With Personalized Features

PING CHENG AND XIAODAI DONG, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC V8P 5C2, Canada
Corresponding author: Xiaodai Dong (xdong@ece.uvic.ca)

ABSTRACT The timely detection of life-threatening ventricular arrhythmias (VAs) is critical for saving a patient's life. General features that characterize ECG waveforms are extracted for VA detection. To take into account the subtle differences in the QRS complexes among different people, new personalized features are proposed in this paper based on the correlation coefficient between a patient-specific regular QRS-complex template and his/her real-time ECG data. Small sets of the most effective features are chosen with support vector machines from 11 newly extracted and 15 previously existing features, for efficient performance and real-time operation. Our proposed new features aveCC and medianCC are verified to be effective in enhancing the performance of existing features under both the record-based and database-based data divisions. Through 50-time random record-based data divisions, all combinations of two features and three features are tested. The top two-feature combination is VFleak and aveCC, which achieves an area under curve (AUC) value of $98.56\% \pm 0.89\%$, a specificity (SP) of $94.80\% \pm 2.15\%$, and an accuracy (ACC) of $94.66\% \pm 1.97\%$; the top three-feature combination is VFleak, MEA, and aveCC, which obtains an AUC of $98.98\% \pm 0.58\%$, an SP of $95.56\% \pm 1.45\%$, and an ACC of $95.46\% \pm 1.36\%$; these results outperform the previous top-two and top-three feature combinations. Similar results are obtained on the database-based data division.

INDEX TERMS Correlation coefficient, ventricular arrhythmia detection, ECG, feature selection, personalized feature, patient-specific template matching, real-time classification, machine learning.

I. INTRODUCTION

Ventricular fibrillation (VF), ventricular flutter (VFL) and rapid ventricular tachycardia (VT), are life-threatening ventricular arrhythmias (VAs), which may cause sudden cardiac arrest and even death if timely therapy is not conducted within a few minutes [1]–[3]. A high quality, easily implementable, fast ventricular arrhythmia (VA) detection algorithm will help achieve a high probability of survival from out-of-hospital heart attack incidents. Henceforth, many research efforts have been dedicated to developing effective VA detection algorithms, aiming to achieve a trade-off between classification performance and real-time performance.

A large number of algorithms for VA detection have been proposed and evaluated in the literature. Detection methods using a single effective feature are proposed, of which features are extracted in temporal/morphologic domains [4]–[6], spectral domain [7]–[9] or other domains [2], [10]–[12]. Jekova [13] conduct comparative assessment of five previously-existing VA detection algorithms. Amann *et al.* [2] evaluate multiple algorithms for VA detection, to verify the proposed algorithm using a single feature. However, the detection performance by using one single feature is limited [14]. One of the key reasons is that the ECG signals vary from one person to another [15], and also change according to different body movements or emotional status even for the same person [16].

To improve VA detection performance, multi-feature classification is investigated, with the aim to obtain the most effective feature set by feature selection techniques [14], [17]–[20] or classifier-based methods [21]–[23]. Recent studies about multi-feature VA detection are presented in [14], [17], and [22] using a large number of ECG data from the public databases and achieving superior performance over that of a single feature. Li *et al.* [22] select a subset of nine features from 14 features using genetic algorithm, further two most effective features using Support Vector Machines (SVMs). Alonso-Atienza *et al.* [14] propose a new filter-type feature selection technique, obtaining nine features to build a simplified high-performance SVM detector for VA detection. Figuera *et al.* [17] explore the difference in the detection of shockable rhythms involving public and out-of-hospital cardiac arrest data. Thirty previously-defined ECG features and five state-of-the-art machine learning classifiers are investigated. Papers [14], [17], and [22] obtain desired

feature sets from the previously existing features, and show generalized classification results, however, without considering any patient-specific information.

In recent years, personalized medicine has received an increasing attention, especially as the Internet based wearable technology allows a significant amount of personal data to be collected. Aramendi *et al.* [24] assess the performance of two spectral and two morphological features for adult and paediatric VA detection and the result shows that the morphological parameters present significant differences between the adult and paediatric patients because of the faster heart rates of the paediatric rhythms. Irusta *et al.* [25] propose a high-temporal resolution algorithm to discriminate shockable from nonshockable rhythms in adults and children. Both [24] and [25] show the individual differences in two distinct populations of adults and children.

Some research on personalized ECG classification consider training with a patient's own ECG records [26], [27], using known general existing features in the literature. One drawback of these methods is that a patient's data cannot include all kinds of arrhythmia events, and hence is limited in arrhythmia training varieties and data size. Furthermore, these methods do not examine deeper the characteristics of individual ECG waveforms which are unique to each person and can be used as a personal identification signature [28], missing potential effective personalized ECG features.

In this paper, we are motivated to 1) propose new personalized ECG features using the patient-specific ECG heart beat waveform; 2) select a small set of the most effective features from the newly-extracted and previously-existing features, aiming to meet the requirement of real-time application. The newly-extracted personalized features are based on the correlation between a patient's regular QRS-complex template extracted from the pre-selected regular/normal sinus rhythms (NSR) and the incoming ECG signal for VA detection.

Since our proposed VA detection method is correlation based, the applications of correlation in ECG signal processing are firstly reviewed here. Correlation based techniques have been widely used in the past, mainly in three types of applications: 1) heart rate detection [29]; 2) alignment method for heart beats [30]; 3) ECG classification [3], [20], [31], [32]. For ECG classification, Dutta *et al.* [31] present a cross-correlation based three-class ECG classification algorithm, to separate normal beats, PVC beats and other beats. They extract 20 features from the magnitude and the phase of the cross-spectral density which is calculated from the Fourier transform of the cross-correlation sequences between each beat signal and one normal reference beat signal. For VA detection, Chen *et al.* divide the incoming ECG signal into short fixed-length segments and calculate the autocorrelation function (ACF) of each segment [2], [20]. If the peak magnitudes of the ACF as a function of time lags do not pass a linear regression test, it is then determined that the test rhythm is subject to ventricular fibrillation. Chin *et al.* [3] classify ECG segments based on the correlation coefficients between the testing segment and the pre-extracted templates respectively for VT and VF, as well as NSR. However, VF is a random-like signal and a fixed template of VF cannot be very accurate. The classification performance is therefore not very good. Hammed and Owis [32] use a hard correlation threshold at 0.85 to distinguish the normal and abnormal beats. Such fixed threshold cannot easily adapt to personalized ECG data, noise levels and measurement platforms.

In general, correlation based approaches have two drawbacks: 1) the correlation result is directly affected by noise and interference in the ECG signal; 2) calculation of the correlation coefficients/function can be time-consuming because of the sliding operation in template matching.

Keeping these in mind, we explore new ways to obtain correlation coefficients (CCs) between the normal ECG template and testing ECG signal with reduced complexity, and search for effective CC based features for VA classification using public ECG databases. The result has potential usage in the surface wireless ECG (wECG) monitoring [33] and automated external defibrillator (AED). In particular, the contributions of this paper are summarized as follows.

- This paper proposes to extract a range of new personalized, simple temporal features originating from the correlation coefficients between a patient-specific QRS-complex template and the heart beats of the same patient.
- This paper studies classification performances using different feature combinations, and the best two-feature combination and the best three-feature combination which include the newly-proposed feature aveCC and medianCC respectively, outperform those mentioned in current reported methods [14], [22].

The rest of this paper is organized as follows. Section II introduces the ECG databases, the data preprocessing and feature extraction. Section III presents personalized feature extraction. Section IV introduces the classification algorithm SVM. Section V conducts simulation and shows the superiority of the proposed method. Section VI concludes this paper.

II. DATA PREPARATION

A. DATABASE INFORMATION

Three commonly-used ECG databases are used in this paper: MIT-BIH arrhythmia database (MITDB) [34], Creighton University Ventricular Tachyarrhythmia Database (CUDB) [35], and MIT-BIH Malignant Ventricular Arrhythmia Database (VFDB) [36]. MITDB is composed of 48 records from different patients and each record contains 30-minute 2-channel ECG data with the sampling rate of 360 Hz. CUDB includes 35 records of 8-minute single-channel ECG data, of which the sampling frequency is 250 Hz. VFDB includes 22 records of 30-minute 2-channel ECG data with the sampling rate of 250 Hz. In this study, only the first channels of records in MITDB and VFDB are used. Moreover, the four paced records in MITDB have been kept.

TABLE 1. Database introduction (segment length: 8 seconds).

Database	Record index	Number	f_s (Hz)	VA segments	Non-VA segments	Total segments
VFDB	1-22	22	250	1027	4737	5764
MITDB	23-70	48	360	20	10780	10800
CUDB	71-105	35	250	464	1710	2174
Total	1-105	105	N/A	1511	17227	18738

The specifications about these three databases are presented in Table 1.

B. DATA PREPROCESSING

The ECG data records are inevitably contaminated by external noises [37] and the ECG signals of interest fall in a specific frequency range. Henceforth, it is necessary to process the raw ECG data before feature extraction is conducted. To this end, the databases downloaded from the online sources [34]–[36] are preprocessed in the same way as in [14]: 1) the mean value is subtracted from the measured ECG signal; 2) the signal is filtered using a five-order moving average filter; 3) the baseline wander is removed using a high-pass filter with the 1 Hz cut-off frequency; 4) the high-frequency noise is eliminated using a second-order Butterworth low-pass filter with the cut-off frequency at 30 Hz.

The preprocessed data are segmented and labeled as VAs or non-VAs by the following labeling rule [14]. VAs include VF, VFL and VT, whereas non-VAs consist of all other rhythms. The rule of segment labeling is: one segment is labeled as $+1$ if no less than 50% of the data inside this segment are VAs; otherwise, this segment is labeled as -1. The segment length is 8 seconds by default. As can be observed in Table 1, MITDB contains very few VA rhythms whereas CUDB includes a lot. The reason that MITDB is still included in the dataset, is to verify general classification performance of the proposed method when up to 15 other rhythms of MITDB are present at the same time.

C. FEATURE EXTRACTION

Each feature characterizes the corresponding segment, distinguishing a VA segment from a non-VA segment. In the literature, many different types of features extracted from an ECG data segment have been studied, and some of them are summarized in Table 2. Basically, these features can be divided into three types, temporal/morphological features, spectral features and complexity features.

In this paper, we propose five correlation coefficient related features and six R-peak related features for VA detection. These new features are highlighted in bold in Table 2. aveCC, devCC, minCC, maxCC and medianCC are correspondingly the average, the standard deviation, the minimum, the maximum and the median of CCs calculated in one segment; aveRR, devRR, minRR, maxRR, and medianRR respectively represent the average, the standard deviation, the minimum, the maximum, and the median of RR intervals in one segment; numPeaks is the number of R-peaks in one ECG segment. Details about these features are elaborated in Section III.

III. PERSONALIZED FEATURE EXTRACTION

It is well known that each person has a unique QRS-complex [28]. If we use a person's regular QRS-complex as a normal template, correlating the person's ECG data samples with the template would give us a subtle indicator how similar the measured beat and the regular beat template are. In the case of severe arrhythmia events such as VAs, the similarity, in other words, the correlation coefficient will distribute randomly. Therefore, the CC related features are potentially useful incorporating the patient-specific ECG signature for VA classification. As mentioned in the introduction, traditional sliding operation for template matching in the CC calculation is computational complex.

Here we propose to simplify the CC calculation and extract new features based on the CC set from one segment. Considering the fact that QRS-complex detection is required in most ECG applications, R-peak positions would already be known after QRS detection. Therefore, we propose to simplify the CC calculation by circumventing sample sliding. Instead, align the detected R-peak with the template R-peak and compute the correlation coefficient between the QRS template and the beat corresponding to each R-peak in the segment. For segment based feature extraction, there are multiple R-peaks and hence multiple CCs in one segment. Next we try to obtain effective features based on these coefficients. There are several ways to derive a CC feature from the CC set of one segment. For example, the median, the average, the standard deviation, the minimum or the maximum of the set are all tested. aveCC and medianCC will be shown later to achieve the superior performances.

To be more specific, the CC related feature extraction is implemented by two successive stages: the QRS-complex detection and the CC related feature extraction. Besides, six R-peak related features are extracted at the same time as a comparison. The details of the two-stage feature extraction are presented next.

A. QRS DETECTION

QRS detection is the first stage for the personalized feature extraction. The objective of QRS detection is to identify the R-peaks. The QRS-complexes are detected by the well-known Pan and Tompkins (P & T) QRS-complex detection algorithm [39]. The preprocessed ECG signal goes through operations of bandpass filtering, derivative, squaring and moving-window integration. The QRS detection identifies a windowed ECG waveform as the QRS-complex, with the window length approximately the same as the maximum possible width of a QRS-complex. Then a peak detector finds the maximum point within

TABLE 2. Feature types.

Types	Features
Temporal/Morphological features	Threshold crossing interval (TCI) [5], threshold crossing sample count (TCSC) [4], Auxiliary counts (Count2) [9] standard exponential (STE) [2], modified exponential (MEA) [2], mean absolute value (MAV) [6], **aveCC, devCC, minCC, maxCC, medianCC, aveRR, devRR, minRR, maxRR, medianRR, and numPeaks**
Spectral features	VF filter (VFleak) [9], spectral algorithm (M, A2 and A3) [7] and median Frequency (FM) [8]
Complexity features	Complexity measurement (CM) [10], phase space reconstruction (PSR) [11], Hilbert transform (HILB) [38] and sample entropy (SpEn) [12]

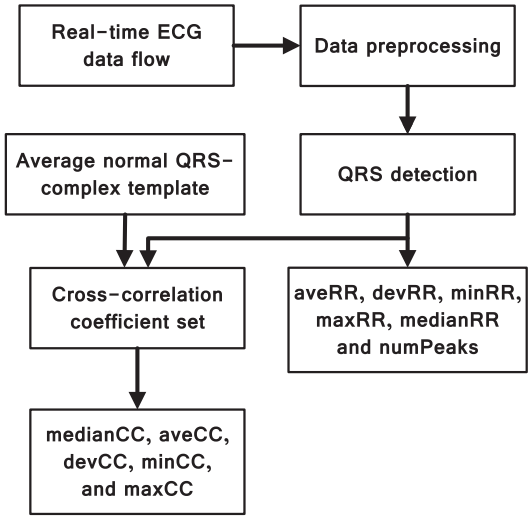

FIGURE 1. General flowchart for extracting the RR-related features and the CC-related features.

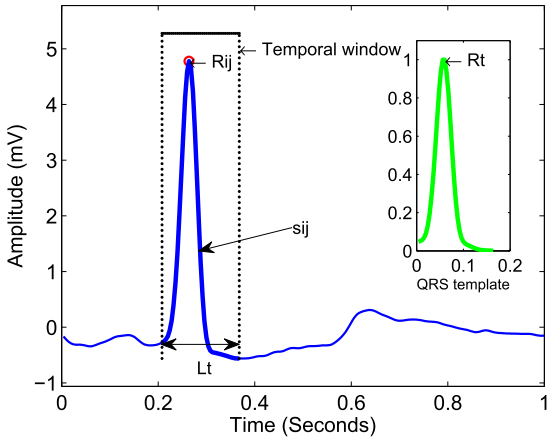

FIGURE 2. CC-related feature extraction.

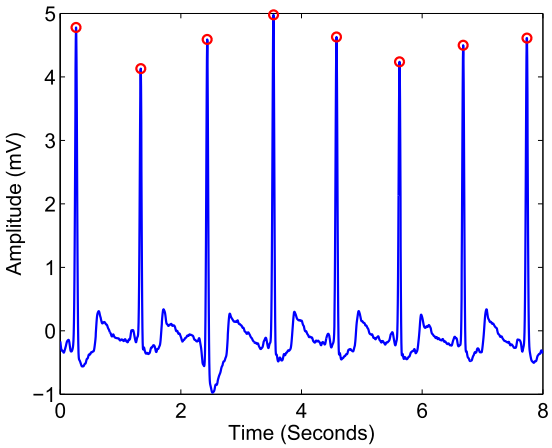

FIGURE 3. One segment of NSR from the first record in CUDB.

this window as the R peak. The details are described in [39].

In this stage, six RR related features are extracted for one segment, namely, aveRR, devRR, minRR, maxRR, medianRR and numPeaks.

B. NEW FEATURE EXTRACTION AFTER QRS DETECTION

Based on the QRS detection, the second stage about the CC related feature extraction is introduced. The average normal QRS-complex template for each person is obtained with QRS detection. The feature extraction procedure and the qualitative analysis of the CC related features are described as follows.

The general flow-chart of the proposed feature extraction algorithm is shown in Fig. 1: first, data preprocessing such as filtering and data segmentation is conducted when raw real-time ECG data come in; then, for each segment, the QRS detection are applied and a set of R peaks is located; based on the RR set, six R-peak related features are calculated; meanwhile, with each detected R peak as a fiducial point, each value of a CC set is obtained from aligning the prepared QRS-complex template with the corresponding detected QRS-complex and calculating the correlation coefficient of the two time series (Fig. 2); based on the CC set, five CC related features are extracted. The mathematical description of the feature extraction is introduced in Algorithm 1. The ECG segment to determine the template was manually selected for each record, and in practice this can be done in an initialization phase in a portable ECG monitor or holter. Thus the QRS-complex template is established after the QRS detection in our simulation as described by the step 3 in Algorithm 1. Automatic determination and update of the template can be designed for real-time ECG as future work.

As shown in Figs. 3-5, the detected R-peaks by the P & T algorithm are marked using red circles, including real R-peaks and mistakenly detected R-peaks. For the NSR segments, the detected R-peaks are almost the real R-peaks based on the fact that QRS detection rate is high [39]. However, the random property of a VA segment leads to the randomness for detected R-peaks. Given the fact that the detected R-peak positions in VA segments are random, the CC calculated from

Algorithm 1 Personalized feature extraction.

1: Identify the R-peaks in each segment. Real-time ECG data flow is firstly divided into segments of a fixed length of L_s seconds ($L_s = 8$ s by default). For the i^{th} segment, J_i R-peaks are detected by the P & T QRS-complex algorithm. Thus, the R-peak set is denoted as $\mathbf{R}_i = \{R_{i1}, R_{i2}, \ldots, R_{ij}, \ldots, R_{iJ_i}\}$, $j = 1, 2, \cdots, J_i$. According to the fiducial point R_{ij}, the j^{th} QRS-complex time series under the temporal window of L_t milliseconds (L_t is set as 160 ms according to the statistical QRS-complex duration of normal beats), is expressed as $\mathbf{s}_j = \{s_{j1}, s_{j2}, \ldots, s_{jk}, \ldots, s_{jK}\}$, s_{jk} represents the k^{th} ECG data sample, $k = 1, 2, \cdots, K$, and $K = L_t * f_s$, as shown in Fig. 2.

2: Calculate the RR set for the i^{th} segment and extract the RR related features based on the RR set. The RR set is calculated by

$$\begin{aligned} \mathbf{r}_i &= r_1, r_2, \ldots, r_j, \ldots, r_{J_i-1}, \\ r_j &= R_{i(j+1)} - R_{ij}. \end{aligned} \tag{1}$$

The RR related features are obtained by

$$\begin{aligned} aveRR_i &= \frac{1}{J_i - 1} \sum_{j=1}^{J_i-1} r_j; \\ devRR_i &= \frac{1}{\sqrt{J_i - 2}} \|\mathbf{r}_i - aveRR_i\|_2 \ (\|\cdot\|_2 - L_2 \text{ norm}); \\ minRR_i &= \mathbf{min}(\mathbf{r}_i); \ maxRR_i = \mathbf{max}(\mathbf{r}_i); \\ medianRR_i &= \mathbf{median}(\mathbf{r}_i); \quad numPeaks_i = J_i. \end{aligned} \tag{2}$$

3: Extract the normalized average normal QRS-complex template for each patient. A pre-selected normal segment from each record is denoted as the t^{th} segment. Then in the template segment, a series of J_t R-peaks is detected and denoted as $\mathbf{R}_t$. The corresponding J_t QRS-complex time series are aligned by R-peaks, averaged, and normalized into the range of [0,1]. The normalized average QRS-complex template is expressed by

$$\bar{\mathbf{s}} = \{\bar{s}_1, \bar{s}_2, \cdots, \bar{s}_k, \cdots, \bar{s}_K\}. \tag{3}$$

4: Calculate the CC set for the i^{th} segment and extract the CC related features based on the CC set. The CC set is calculated by

$$\mathbf{c}_i = \{c_1, c_2, \ldots, c_j, \ldots, c_{J_i}\}, \ c_j = \frac{\sum_{k=1}^{K} \bar{s}_k s_{jk}}{\|\bar{\mathbf{s}}\|_2 \left\|\mathbf{s}_j\right\|_2}. \tag{4}$$

The CC related features are obtained by

$$\begin{aligned} medianCC_i &= \mathbf{median}(\mathbf{c}_i); \ aveCC_i = \frac{1}{J_i} \sum_{j=1}^{J_i} c_j; \\ devCC_i &= \frac{1}{\sqrt{J_i - 1}} \|\mathbf{c}_i - aveCC_i\|_2; \\ minCC_i &= \mathbf{min}(\mathbf{c}_i); \quad maxCC_i = \mathbf{max}(\mathbf{c}_i). \end{aligned} \tag{5}$$

5: Repeat from Step 2 to Step 4 to get CC related values for all the segments.

6: End.

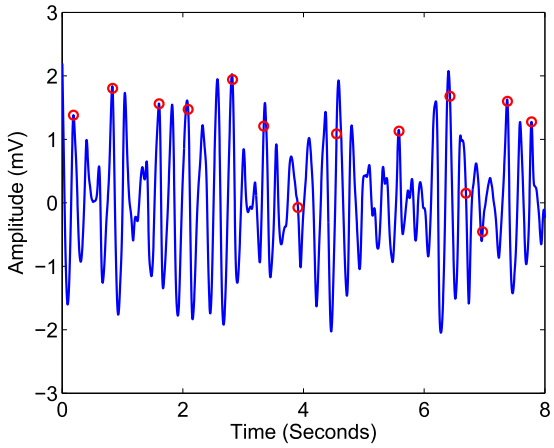

FIGURE 4. One segment of VF from the first record in CUDB.

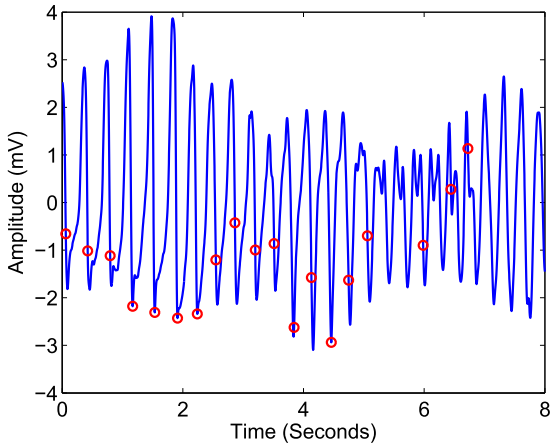

FIGURE 5. One segment of VT from the third record in VFDB.

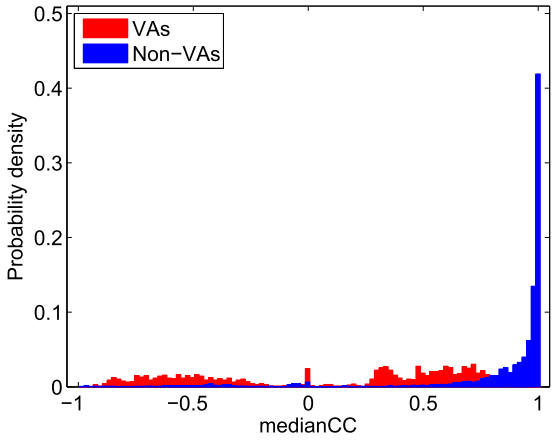

FIGURE 6. Probability histogram of medianCC on the complete dataset.

the R-peak alignment will be low because the VA waveform has dissimilar shapes to the average normal template. As shown in Fig. 6, the values of medianCC are generally high for non-VA segments while they are in the low range for VA segments on the complete dataset.

TABLE 3. Performance of single feature on test sets.

feature 1	SE (%)	SP (%)	PP (%)	ACC (%)	BER (%)	AUC (%)
TCSC	95.64 ± 2.77	92.23 ± 3.25	51.94 ± 12.02	92.54 ± 2.98	**6.06 ± 2.18**	**97.45 ± 1.34**
VFleak	90.00 ± 4.80	**95.65 ± 2.47**	**64.89 ± 14.39**	**95.25 ± 2.35**	7.17 ± 2.83	97.17 ± 1.84
MAV	95.57 ± 2.54	91.62 ± 3.10	49.81 ± 11.59	91.97 ± 2.84	6.40 ± 2.02	97.03 ± 1.71
PSR	93.13 ± 3.13	91.60 ± 2.83	48.84 ± 11.71	91.74 ± 2.57	7.63 ± 1.93	96.32 ± 1.38
HILB	87.95 ± 10.33	90.94 ± 3.55	46.09 ± 12.03	90.64 ± 2.96	10.56 ± 4.61	95.85 ± 1.49
SpEn	88.55 ± 9.78	89.43 ± 3.98	42.44 ± 11.91	89.29 ± 3.37	11.01 ± 4.45	95.23 ± 2.01
A2	80.41 ± 6.44	95.07 ± 3.03	60.41 ± 16.66	93.94 ± 2.77	12.26 ± 3.36	93.50 ± 2.67
medianCC	**97.28 ± 1.58**	84.51 ± 3.47	34.94 ± 9.02	85.51 ± 3.18	9.11 ± 1.73	92.54 ± 2.32
MEA	84.85 ± 9.74	85.25 ± 4.24	33.33 ± 9.59	85.14 ± 3.57	14.95 ± 4.39	92.26 ± 2.24
TCI	87.53 ± 4.58	83.71 ± 4.15	31.88 ± 10.09	84.03 ± 3.79	14.38 ± 2.76	92.16 ± 2.33
aveCC	93.28 ± 4.87	83.87 ± 3.59	33.13 ± 8.51	84.59 ± 3.20	11.42 ± 2.36	92.14 ± 2.31
Count2	85.36 ± 5.24	84.72 ± 4.19	33.09 ± 11.08	84.76 ± 3.84	14.96 ± 3.00	91.11 ± 2.52
A3	78.31 ± 8.30	88.17 ± 3.38	36.67 ± 11.17	87.34 ± 2.99	16.76 ± 4.12	90.32 ± 3.48
M	83.79 ± 6.08	84.90 ± 3.58	32.59 ± 9.90	84.80 ± 3.34	15.65 ± 3.49	89.94 ± 3.36
numPeaks	75.42 ± 6.89	89.43 ± 4.13	39.19 ± 12.19	88.38 ± 3.71	17.58 ± 3.45	88.69 ± 3.74
FM	80.02 ± 8.63	70.54 ± 6.06	19.04 ± 5.72	71.21 ± 5.42	24.72 ± 4.01	83.52 ± 4.29
STE	58.12 ± 6.18	91.86 ± 3.14	39.25 ± 13.96	89.19 ± 2.68	25.01 ± 3.15	80.33 ± 3.46
CM	68.67 ± 12.92	73.36 ± 10.67	19.79 ± 8.29	72.72 ± 9.31	28.99 ± 5.47	78.56 ± 6.86
maxRR	17.34 ± 20.82	87.24 ± 16.89	NaN ± NaN	81.11 ± 13.55	47.71 ± 3.52	61.78 ± 5.26

IV. CLASSIFICATION ALGORITHM

Support vector machine (SVM) is a widely used and effective algorithm in the literature [40], [41]. Among the numerous variants of SVMs, the soft-margin SVM using a Gaussian kernel function is widely adopted in practical classification problems. This method can classify data having non-linear relationship with features, and also non-separable data with a designed or minimum error rate [42]. The SVM model is confirmed through two-stage operations: the first stage is to train this model on the training set; the second stage is to evaluate the classification performance of the model on the test set. The model with desired performance is finally determined. The basic SVM operation is described as follows.

Given any feature vectors $\tilde{\mathbf{x}}_i \in \mathbb{R}^{M\times 1}$ and the corresponding labels $y_i \in \{+1, -1\}$ on the training set, the following optimization problem is solved

$$\begin{aligned} \min_{\mathbf{w},b,s_i} \quad & \frac{1}{2}\|\mathbf{w}\|^2 + \tau\sum_{i=1}^{N} s_i \\ \text{subject to} \quad & y_i(\mathbf{w}^T\phi(\tilde{\mathbf{x}}_i) + b) \geq 1 - s_i, \\ & s_i \geq 0 \text{ for } i = 1, \ldots, N, \end{aligned} \tag{6}$$

where the weight vector $\mathbf{w} \in \mathbb{R}^{M\times 1}$, $\tilde{\mathbf{x}}_i \in \mathbb{R}^{M\times 1}$, $\phi(\tilde{\mathbf{x}}_i)$ is a linear or nonlinear transformation of $\tilde{\mathbf{x}}_i$, $\phi(\tilde{\mathbf{x}}_i) \in \mathbb{R}^{M\times 1}$, s_i represents the violation value of data pair $(\tilde{\mathbf{x}}_i, y_i)$ to the classification boundaries, τ is the cost parameter for the violation chosen by users, and b is an unknown constant.

By using Lagrange multipliers, the Lagrange dual problem of (6) is expressed as

$$\begin{aligned} \min_{\mu_1,\mu_2,\cdots,\mu_N} \quad & \frac{1}{2}\sum_{i,j=1}^{N}\mu_i y_i \mu_j y_j K_G(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j) - \sum_{i=1}^{N}\mu_i \\ \text{subject to} \quad & \sum_{i=1}^{N}\mu_i y_i = 0,\ 0 \leq \mu_i \leq \tau, \end{aligned} \tag{7}$$

where μ_i is a lagrange multiplier corresponding to the constraints of (6), $K_G(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j) = e^{-\|\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j\|^2/2\sigma^2}$, is called a Gaussian kernel and σ is a user-defined parameter. After solving (7) and obtaining the lagrange multiplier set, the first stage of SVM classification is completed.

The second stage for SVM classification is to evaluate the performance of the classification model. For any known feature vector $\tilde{\mathbf{x}}$ on the test set, the predicted label $\hat{y}$ is obtained as

$$\hat{y} = \text{sign}(\sum_{i=1}^{N}\mu_i y_i K_G(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}) + b). \tag{8}$$

By comparing the predicted labels with the true labels on the whole test set, the classification performance is analysed by calculating performance indices. There are several commonly-used statistical indices to characterize the classification performance, such as the sensitivity (SE), the specificity (SP), the accuracy (ACC), the area under the curve (AUC), the false positive rate (FPR), the positive prediction (PP) and the balanced error rate (BER) [14].

V. SIMULATION

In this section, the effectiveness of the proposed personalized features for VA classification using SVM are evaluated through simulation. A total number of 105 records are considered in this paper. In order to guarantee the data independence between the training dataset and the test dataset [43], a record-based data division and a database-based data division are both employed. As seen in Table 1, the classification problem we deal with is a binary classification of unbalanced data. To solve the unbalanced classification problem, τ in the SVM optimization problem (Eq. (6)) is assigned different values for the VA (positive) class and the non-VA (negative) class, according to the practical proportion of these two classes in the training set. (Note: Features or feature combinations in the following tables are sorted by AUC.)

For the record-based data division [22], the whole dataset is divided by randomly choosing 70% records (74 records) as the training set and the left 30% records (31 records) as

TABLE 4. Performance of combinations of two features on test sets.

feature 1	feature 2	SE (%)	SP (%)	PP (%)	ACC (%)	BER (%)	AUC (%)
VFleak	**aveCC**	92.60 ± 5.27	**94.80 ± 2.15**	60.62 ± 11.27	**94.66 ± 1.97**	6.30 ± 2.67	**98.56 ± 0.89**
VFleak	**medianCC**	93.18 ± 4.92	94.45 ± 2.34	59.33 ± 12.00	94.38 ±2.15	6.18 ± 2.60	98.55 ± 0.94
TCSC	TCI	94.80 ± 3.33	92.79 ± 3.36	54.00 ± 13.04	92.98 ± 3.07	6.20 ± 2.30	98.38 ± 0.79
TCSC	MEA	**95.06 ± 2.90**	92.59 ± 3.32	53.31 ± 12.59	92.82 ± 3.03	6.17 ± 2.14	98.37 ± 0.74
TCSC	SpEn	94.82 ± 3.37	92.99 ± 3.27	54.67 ± 12.94	93.17 ± 3.00	6.10 ± 2.28	98.36 ± 0.82
MAV	SpEn	94.40 ± 3.41	93.12 ± 3.35	55.23 ± 13.64	93.26 ± 3.08	6.24 ± 2.33	98.35 ± 0.87
TCSC	VFleak	94.51 ± 4.51	93.47 ± 2.95	56.13 ± 12.58	93.60 ± 2.73	**6.01 ± 2.69**	98.24 ± 1.05
VFleak	SpEn	91.46 ± 5.26	94.77 ± 2.84	**60.90 ± 14.79**	94.54 ± 2.71	6.89 ± 3.16	98.24 ± 1.14
TCSC	Count2	93.59 ± 4.11	92.90 ± 3.36	54.07 ± 12.98	92.99 ± 2.99	6.76 ± 2.24	98.19 ± 1.01
MAV	MEA	94.32 ± 3.07	92.74 ± 3.35	53.73 ± 13.05	92.90 ± 3.05	6.47 ± 2.10	98.19 ± 0.95
TCSC	SpEn [14]	94.82 ± 3.37	92.99 ± 3.27	54.67 ± 12.94	93.17 ± 3.00	6.10 ± 2.28	98.36 ± 0.82
VFleak	Count2 [22]	88.96 ± 5.85	95.98 ± 2.29	66.33 ± 14.45	95.46 ± 2.12	7.53 ± 3.04	97.65 ± 1.38

TABLE 5. Performance of combinations of three features on test sets.

feature 1	feature 2	feature 3	SE (%)	SP (%)	PP (%)	ACC (%)	BER (%)	AUC (%)
VFleak	MEA	**aveCC**	93.87 ± 3.80	**95.56 ± 1.45**	**63.81 ± 10.53**	**95.46 ± 1.36**	**5.28 ± 2.00**	**98.98 ± 0.58**
VFleak	**medianCC**	MEA	93.93 ± 3.94	95.50 ± 1.58	63.59 ± 11.21	95.40 ± 1.47	5.29 ± 2.07	98.96 ± 0.62
VFleak	**medianCC**	SpEn	93.44 ± 4.51	95.08 ± 2.19	62.12 ± 12.09	94.98 ± 2.05	5.74 ± 2.45	98.94 ± 0.68
VFleak	**aveCC**	TCI	93.48 ± 4.16	95.44 ± 1.61	63.20 ± 10.58	95.31 ± 1.53	5.54 ± 2.23	98.92 ± 0.66
VFleak	SpEn	**aveCC**	93.07 ± 4.97	95.06 ± 2.17	61.90 ± 11.79	94.93 ± 2.00	5.93 ± 2.61	98.91 ± 0.68
VFleak	**medianCC**	TCI	93.85 ± 4.25	95.36 ± 1.68	63.00 ±10.89	95.27 ± 1.58	5.39 ± 2.28	98.90 ±0.69
VFleak	**medianCC**	Count2	92.86 ± 4.59	95.16 ± 1.85	61.97 ± 11.62	95.02 ± 1.64	5.99 ± 2.22	98.87 ± 0.65
VFleak	Count2	**aveCC**	92.85 ± 4.51	95.27 ± 1.77	62.49 ± 11.18	95.11 ± 1.57	5.94 ± 2.15	98.87 ± 0.64
VFleak	**aveCC**	CM	93.59 ± 4.49	95.06 ± 2.20	62.31 ± 11.59	94.98 ± 2.01	5.68 ± 2.34	98.83 ± 0.78
TCSC	VFleak	**medianCC**	**95.22 ± 3.78**	94.10 ± 2.69	58.84 ± 12.42	94.23 ± 2.48	5.34 ± 2.25	98.80 ± 0.66
TCSC	VFleak	SpEn [14]	93.98 ± 4.61	93.68 ± 3.07	57.09 ± 13.48	93.75 ± 2.85	6.17 ± 2.77	98.63 ± 0.86
Count2	VFleak	A3 [22]	88.69 ± 5.88	96.00 ± 2.31	66.49 ± 14.49	95.47 ± 2.14	7.65 ± 3.05	97.65 ± 1.38

TABLE 6. Evaluation performance on CUDB with VFDB and MITDB as the training set.

feature 1	feature 2	feature 3	SE (%)	SP (%)	PP (%)	ACC (%)	BER (%)	AUC (%)
medianCC	MAV	SpEn	89.66	89.47	69.80	89.51	10.44	**96.17**
MAV	Count2	**aveCC**	93.10	84.97	62.70	86.71	10.96	96.08
MAV	SpEn	**aveCC**	88.15	89.59	69.68	89.28	11.13	96.05
medianCC	MAV	Count2	92.67	85.32	63.14	86.89	11.00	96.03
medianCC	Count2	**aveCC**	**94.18**	87.43	67.02	88.87	**9.20**	95.92
Count2	**aveCC**	FM	92.46	87.49	66.72	88.55	10.03	95.90
medianCC	Count2	FM	**94.18**	87.31	66.82	88.78	9.25	95.78
TCSC	**medianCC**	SpEn	89.22	85.79	63.01	86.52	12.49	95.78
VFleak	Count2	**aveCC**	83.84	**92.22**	**74.52**	90.43	11.97	95.77
medianCC	Count2	CM	86.64	91.52	73.49	**90.48**	10.92	95.70

TABLE 7. Evaluation performance on VFDB with CUDB and MITDB as the training set.

feature 1	feature 2	feature 3	SE (%)	SP (%)	PP (%)	ACC (%)	BER (%)	AUC (%)
VFleak	**medianCC**	MEA	**97.47**	90.63	69.27	91.85	5.95	**98.46**
VFleak	**medianCC**	TCI	**97.47**	90.42	68.80	91.67	6.06	98.40
VFleak	**medianCC**	Count2	96.88	90.12	68.01	91.33	6.50	98.34
VFleak	MEA	**aveCC**	97.18	90.12	68.08	91.38	6.35	98.31
VFleak	**aveCC**	TCI	97.18	89.70	67.16	91.03	6.56	98.28
VFleak	Count2	MEA	95.62	92.65	73.83	93.18	**5.86**	98.23
A2	Count2	MEA	92.60	**93.01**	**74.18**	92.94	7.19	98.22
VFleak	Count2	**aveCC**	95.91	90.35	68.31	91.34	6.87	98.16
VFleak	**medianCC**	SpEn	**97.47**	84.99	58.47	87.21	8.77	98.10
VFleak	Count2	TCI	95.33	92.67	73.83	**93.15**	6.00	98.09

the test set. This data division procedure is repeated 50 times. The SVM classifier is trained on the training set and validated on the test set. The mean and the standard deviation of the 50-time classification performances on the test set are calculated and presented in tables. The performances are sorted descendingly according to AUC values for different feature combinations.

First, Table 3 shows the classification performance using SVM with a single feature. TCSC performs well in terms of BER and AUC. VFleak ranks the best in SP, PP and ACC. Our proposed feature medianCC achieves the highest SE among all the features.

Table 4 shows the classification performance using two features chosen from the feature set. There are 171 two-feature combinations in total. The top-ten feature combinations with the highest AUC values are presented. Our proposed feature aveCC and medianCC, combined with VFleak respectively, achieve the best two AUC values, whereas the

combination of aveCC and VFleak also obtains the highest SP and ACC.

Table 5 shows the classification performance with three-feature combinations. Among 969 three-feature combinations, the new features, aveCC and medianCC, separately working with VFleak and MEA, achieve the highest two AUC values, which is consistent with the results presented in Table 4. The combination of VFleak, MEA and aveCC performs the best in terms of SP, PP, ACC, BER and AUC, with an acceptable SE, compared with two top-three combinations mentioned in the previously existing papers, i.e., the combination of TCSC, VFleak and SpEn [14], and the combination of Count2, VFleak and A3 [22]. Furthermore, in the top-ten combinations with the highest AUC values, aveCC appears four times and medianCC appears five times, only after VFleak.

For the database-based data division, we do simulations to test three-feature combinations, namely, any two databases as the training set and the third one as the test set. As there are only a few VA segments in the MITDB database (shown in Table 1), MITDB database is only used in the training sets, combined with CUDB or VFDB, involving up to 15 other rhythms as kind of interference.

For data from CUDB as the test set, among 969 three-feature combinations, the best three-feature combination are (medianCC, MAV, SpEn); medianCC and aveCC appear respectively 6 times and 5 times in the top-ten three-feature combinations, as shown in Table 6. For data from VFDB as the test set, the best ones are (VFleak, medianCC, MEA); medianCC and aveCC appear respectively 4 times and 3 times, as shown in Table 7.

The limitation of the proposed method is that our used QRS-complex template is not designed to be adaptive in this paper. This is not a problem when using the open source databases, as the regular normal beat signals do not vary much in each record. Besides, the template we used is the QRS-complex (R complex), which is known as a function of distance, other than the heart rate. So the R complex remains fairly constant with changes in heart rate, other than the P complex or the T complex [28]. In the real time online application of this template, the patient-specific fixed template should be further verified at different heart rates.

VI. CONCLUSION

This paper has proposed two new simple, effective, personalized temporal features named aveCC and medianCC for VA detection. These two features have been selected and verified among 11 newly-extracted features and 15 previously-validated existing features. Combined with one or two other features, aveCC or medianCC works well with SVM classifiers under both the record-based data division and the database-based data division of well-known public ECG databases.

Through 50-times record-based data division, the effectiveness of our proposed features have been validated by the statistic results on the test sets. The top two two-feature combinations are VFleak with aveCC, and VFleak with medianCC. The top two three-feature combinations are respectively aveCC and medianCC combined with VFleak and MEA.

For database-based data division, the three-feature combination of medianCC, MAV, and SpEn ranks atop when testing on CUDB with VFDB and MITDB as the training set. And the combination of VFleak, medianCC, and MEA ranks first when testing on VFDB with CUDB and MITDB as the training set.

In conclusion, these simple two or three features involving aveCC or medianCC achieve the best classification performances using SVMs, compared to other feature combinations. These features, especially aveCC and medianCC, have explicit physical meanings and low implementation complexity. The top two-feature and three-feature combinations enable accurate, low complexity, fast and personalized VA detection, leading to potential usage in real time ECG detection applications to improve the prediction of sudden cardiac death.

Acknowledgment

The authors would like to thank Felipe Alonso-Atienza for his help in discussion and feature extraction.

REFERENCES

[1] *Resuscitation Guidelines 2000*, vol. 102. London, U.K., Resuscitation Council (UK) Ltd., Aug. 2000.

[2] A. Amann, R. Tratnig, and K. Unterkofler, "Reliability of old and new ventricular fibrillation detection algorithms for automated external defibrillators," *Biomed. Eng. Online*, vol. 60, no. 4, pp. 1–15, Oct. 2005.

[3] F. J. Chin, Q. Fang, T. Zhang, and I. Cosic, "A fast critical arrhythmic ECG waveform identification method using cross-correlation and multiple template matching," in *Proc. Eng. Med. Biol. Soc.*, Aug. 2010, pp. 1922–1925.

[4] M. A. Arafat, A. W. Chowdhury, and M. K. Hasan, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," *Signal Image Video Process.*, vol. 5, no. 1, pp. 1–10, Sep. 2011.

[5] 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. 837–843, Sep. 1990.

[6] E. M. A. Anas, S. Y. Lee, and M. K. Hasan, "Sequential algorithm for life threatening cardiac pathologies detection based on mean signal strength and EMD functions," *Biomed. Eng. Online*, vol. 9, no. 1, p. 43, 2010.

[7] S. Barro, R. Ruiz, D. Cabello, and J. Mira, "Algorithmic sequential decision-making in the frequency domain for life threatening ventricular arrhythmias and imitative artefacts: A diagnostic system," *J. Biomed. Eng.*, vol. 11, no. 4, pp. 320–328, Jul. 1989.

[8] R. Dzwonczyk, C. G. Brown, and H. A. 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.

[9] I. Jekova and V. Krasteva, "Real time detection of ventricular fibrillation and tachycardia," *Physiol. Meas.*, vol. 25, no. 5, pp. 1167–1178, 2004.

[10] 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. 548–555, May 1999.

[11] 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.

[12] H. Li, W. Han, C. Hu, and Q. Meng, "Detecting ventricular fibrillation by fast algorithm of dynamic sample entropy," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2009, pp. 1105–1110.

[13] I. Jekova, "Comparison of five algorithms for the detection of ventricular fibrillation from the surface ECG," *Physiol. Meas.*, vol. 21, pp. 429–439, Aug. 2000.

[14] F. Alonso-Atienza, E. Morgado, L. Fernandez-Martinez, A. Garcia-Alberola, and J. L. Rojo-Alvarez, "Detection of life-threatening arrhythmias using feature selection and support vector machines," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 832–840, Mar. 2014.
[15] M. S. Islam, N. Ammour, and M. Abdullah-Al-Wadud, "Selection of heartbiometric templates for fusion," *IEEE Access*, vol. 5, pp. 1753–1761, Feb. 2017.
[16] S. H. Jambukia, V. K. Dabhi, and H. B. Prajapati, "Classification of ECG signals using machine learning techniques: A survey," in *Proc. Int. Conf. Adv. Comput. Eng. Appl.*, Mar. 2015, pp. 714–721.
[17] C. Figuera *et al.*, "Machine learning techniques for the detection of shockable rhythms in automated external defibrillators," *PLoS ONE*, vol. 11, no. 7, pp. 1–17, Jul. 2016.
[18] 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.
[19] Q. Gu, Z. Li, and J. Han, "Generalized fisher score for feature selection," in *Proc. Conf. Uncertain. Artif. Intell.*, Jul. 2011, pp. 266–273.
[20] S. Chen, N. Thakor, and M. Mover, "Ventricular fibrillation detection by a regression test on the autocorrelation function," *Med. Biol. Eng. Comput.*, vol. 25, no. 3, pp. 241–249, May 1987.
[21] K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 2, pp. 179–185, Feb. 1999.
[22] Q. Li, C. Rajagopalan, and G. D. Clifford, "Ventricular fibrillation and tachycardia classification using a machine learning approach," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1607–1613, Jun. 2014.
[23] R. Clayton, A. Murray, and R. W. Campbell, "Recognition of ventricular fibrillation using neural networks," *Med. Biol. Eng. Comput.*, vol. 32, no. 2, pp. 217–220, Mar. 1994.
[24] 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, Apr. 2010.
[25] 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, Jan. 2012.
[26] M. Faezipour, A. Saeed, S. C. Bulusu, A. H. M. M. Nourani, and L. Tamil, "A patient-adaptive profiling scheme for ECG beat classification," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1153–1165, Sep. 2010.
[27] J. Park and K. Kang, "PcHD: Personalized classification of heartbeat types using a decision tree," *Comput. Biol. Med.*, vol. 54, pp. 79–88, Aug. 2014.
[28] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern Recognit.*, vol. 38, no. 1, pp. 133–142, Jan. 2005.
[29] T. Last, C. D. Nugent, and F. J. Owens, "Multi-component based cross correlation beat detection in electrocardiogram analysis," *Biomed. Eng. Online*, vol. 3, no. 1, p. 26, Jul. 2004.
[30] E. Laciar and R. Jané, and D. H. Brooks, "Improved alignment method for noisy high-resolution ECG and Holter records using multiscale cross-correlation," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 3, pp. 344–353, Mar. 2003.
[31] S. Dutta, A. Chatterjee, and S. Munshi, "Correlation technique and least square support vector machine combine for frequency domain based ECG beat classification," *Med. Eng. Phys.*, vol. 32, no. 10, pp. 1161–1169, Dec. 2010.
[32] N. S. Hammed and M. I. Owis, "Patient adaptable ventricular arrhythmia classifier using template matching," in *Proc. IEEE Conf. Biomed. Circuits Syst.*, Oct. 2015, pp. 1–4.
[33] A. Page, T. Soyata, J. Couderc, and M. Aktas, "An open source ECG clock generator for visualization of long-term cardiac monitoring data," *IEEE Access*, vol. 3, pp. 2704–2714, 2015.
[34] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May/Jun. 2001.
[35] F. M. Nolle, F. K. Badura, J. M. Catlett, R. W. Bowser, and M. H. Sketch, "Crei-gard, a new concept in computerized arrhythmia monitoring systems," in *Proc. Comput. Cardiol.*, vol. 13. 1986, pp. 515–518.
[36] S. Greenwald, "The development and analysis of a ventricular fibrillation detector," M.S. thesis, Dept. Elect. Eng. Comput. Sci., MIT, Cambridge, MA, USA, 1986.
[37] J. Li, G. Deng, W. Wei, H. Wang, and Z. Ming, "Design of a real-time ECG filter for portable mobile medical systems," *IEEE Access*, vol. 5, pp. 696–704, 2017.
[38] 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.
[39] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, Mar. 1985.
[40] G. Camps-Balls, J. L. Rojo-Alvarez, and M. Martinez-Ramon, *Kernel Methods in Bioengineering, Signal and Image Processing*. Hershey, PA, USA: Idea Group Inc., Jan. 2007.
[41] V. N. Vapnik, *The Nature of Statistical Learning Theory.* New York, NY, USA: Springer-Verlag, Mar. 1995.
[42] W. Lu, "Support vector machines," in *Lecture Notes of Machine Learning for Signal Processing*. Victoria, BC, Canada: Univ. Victoria, Jan. 2015, ch. 5, pp. 140–161.
[43] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, "Training versus testing," in *Learning From Data*. Berlin, Germany: AMLBook, Mar. 2012, ch. 2, pp. 39–76.

PING CHENG received the B.Sc. degree in electronics and communication engineering and the M.Sc. degree in electrical engineering from Northwestern Polytechnical University, Xi'an, China, in 2009 and 2012, respectively. She is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Victoria, Victoria, BC, Canada. Her research interests include eHealth, biomedical engineering, ECG signal processing, heart disease classification, data analysis, health care monitoring systems, and machine learning applications.

XIAODAI DONG (S'97–M'00–SM'09) received the B.Sc. degree in information and control engineering from Xi'an Jiaotong University, China, in 1992, the M.Sc. degree in electrical engineering from the National University of Singapore in 1995, and the Ph.D. degree in electrical and computer engineering from Queen's University, Kingston, ON, Canada, in 2000.

From 1999 to 2002, she was with Nortel Networks, Ottawa, ON, Canada, where she was involved in the base transceiver design of the third-generation (3G) mobile communication systems. From 2002 to 2004, she was an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. She was a Canada Research Chair (Tier II) from 2005 to 2015. Since 2005, she has been with the University of Victoria, Victoria, Canada, where she is currently a Professor with the Department of Electrical and Computer Engineering. Her research interests include wireless communications, radio propagation, ultra-wideband radio, machine-to-machine communications, wireless security, eHealth, smart grid, nano-communications, and signal processing for communication applications.

Prof. Dong served as an Editor of the IEEE TRANSACTIONS on WIRELESS COMMUNICATIONS from 2009 to 2014, the IEEE TRANSACTIONS on COMMUNICATIONS from 2001 to 2007, and the *Journal of Communications and Networks* from 2006 to 2015. She is currently an Editor of the IEEE TRANSACTIONS on VEHICULAR TECHNOLOGY.