Novel Density Poincaré Plot Based Machine Learning Method to Detect Atrial Fibrillation From Premature Atrial/Ventricular Contractions

Syed Khairul Bashar, *Student Member, IEEE*, Dong Han, *Student Member, IEEE*, Fearass Zieneddin, Eric Ding, Timothy P. Fitzgibbons, Allan J. Walkey, David D. McManus, Bahram Javidi, *Fellow, IEEE*, and Ki H. Chon, *Senior Member, IEEE*

Abstract*—*Objective: **Detection of Atrial fibrillation (AF) from premature atrial contraction (PAC) and premature ventricular contraction (PVC) is difficult as frequent occurrences of these ectopic beats can mimic the typical irregular patterns of AF. In this paper, we present a novel density Poincaré plot-based machine learning method to detect AF from PAC/PVCs using electrocardiogram (ECG) recordings. *Methods:* First, we propose the generation of this new density Poincaré plot which is derived from the difference of the heart rate (DHR) and provides the overlapping phase-space trajectory information of the DHR. Next, from this density Poincaré plot, several image processing domain-based approaches including statistical central moments, template correlation, Zernike moment, discrete wavelet transform and Hough transform features are used to extract suitable features. Subsequently, the infinite latent feature selection algorithm is implemented to rank the features. Finally, classification of AF vs. PAC/PVC is performed using K-Nearest Neighbor, Support vector machine (SVM) and Random Forest (RF) classifiers. Our method is developed and validated using a subset of Medical Information Mart for Intensive Care (MIMIC) III database containing 10 AF and 10 PAC/PVC subjects. Results- During the segment-wise 10-fold cross-validation, SVM achieved the best performance with 98.99% sensitivity, 95.18% specificity and 97.45% accuracy with the extracted features. In subject-wise scenario, RF achieved the highest accuracy of 91.93%. Moreover, we further validated the proposed method using two other databases: wearable armband ECG data and the Physionet AFPDB. 100% PAC detection accuracy was obtained for both databases without any further training. Conclusion: Our proposed density Poincaré plot-based method showed superior performance when compared with four existing algorithms; thus showing the efficacy of the extracted image domain-based features. Significance: From intensive care unit's ECG to wearable armband ECGs, the proposed method is shown to discriminate PAC/PVCs from AF with high accuracy.**

***Index Terms*—Premature atrial contraction, premature ventricular contraction, atrial fibrillation, density Poincaré plot, correlation, wavelet transform, random forest, SVM.**

I. Introduction

ATRIAL fibrillation (AF) is the most common arrhythmia and has become a global health interest as the prevalence and the associated mortality have grown exponentially in the past decade [1]. AF affects about 33.5 million people worldwide including 3 million Americans [2]. However, the true prevalence of AF is unknown because in many patients the condition is asymptomatic and remains undetected [3]. AF is also known to increase the risk for stroke, heart failure and other comorbidities [4]. Since AF is characterized by highly variable ventricular beat intervals, several algorithms for the automatic detection of AF is based on QRS beat interval variability [5]–[11].

However, Premature atrial and ventricular contractions (PAC and PVC, respectively), which are common benign causes of rhythm irregularity, can mimic the irregular beat pattern typical of AF [12]. PACs occur when an ectopic focus originating in the atrium leads to premature activation of the atria prior to typical sinoatrial node activation whereas a PVC occurs when a similar process occurs in the ventricle [13]. Rhythm-based AF detectors are prone to produce false AF alarms in the presence of premature beats [8], [14]. As a result, the effect of these ectopic beats needs to be carefully considered as these are known to confound the automated AF decision [15].

There have been many studies for beat detection for different types of cardiac arrhythmias using the MIT-BIH arrhythmia database. In [16] multiresolution wavelet transform is used to extract QRS complex features and based on those features, support vector machine (SVM) along with neural network is implemented to classify normal, left bundle branch block, right bundle branch block and paced beats. Shannon entropy, log energy entropy, Renyi entropy and Tsallis entropy, computed on wavelet packet decomposition; are used with random forest to classify five different types of beats [17]. Morphological features extracted from ECG signal components are used with linear

Manuscript received March 11, 2020; revised June 1, 2020; accepted June 16, 2020. Date of publication June 23, 2020; date of current version January 20, 2021. This work was supported by NIH under Grant R01 HL136660. *(Corresponding author: Ki H. Chon.)*

Syed Khairul Bashar, Dong Han, and Fearass Zieneddin are with the Biomedical Engineering Department, University of Connecticut.

Eric Ding, Timothy P. Fitzgibbons, and David D. McManus are with the Division of Cardiology, University of Massachusetts Medical School.

Allan J. Walkey is with the Department of Medicine, Boston University School of Medicine.

Bahram Javidi is with the Department of Electrical and Computer Engineering, University of Connecticut.

Ki H. Chon is with the Biomedical Engineering Department, University of Connecticut, Storrs, CT 06269 USA (e-mail: ki.chon@uconn.edu).

Digital Object Identifier 10.1109/TBME.2020.3004310

0018-9294 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

discriminant analysis in [18]. Wavelet transform, independent component analysis along with RR interval information are used as extracted features to classify different heartbeats with SVM in [19]. Higher order cumulants of the ECG beats are modeled by linear combination of Hermitian basis function in [20] and the model parameters are used as features to classify five different hear beats. In [21], principal component principal component analysis (PCA) of the segmented ECG beats are performed while PCA on the higher order statistics coefficients is applied for arrhythmia detection in [22]. In [23], combination of linear and nonlinear features along with SVM and neural network classifiers are used to recognize five types of arrhythmia beats. A diverse set of features including higher order statistics, morphological features, Fourier coefficients and higher order statistics of the wavelet packet coefficients are used in [24] to identify different arrhythmia beats.

However, none of these above-mentioned methods considered arrhythmia beat detection in the presence of AF. As a result, it is not clear how these methods will perform when AF is present and whether they can reduce false positives for AF detection even with PAC/PVC. Arrhythmia beat detection (i.e., PAC/PVC detection) when there is AF rhythm can become challenging. So the accuracy of the algorithms in the presence of ectopic beats needs to be established [12].

In order to address this problem, we have presented in this paper a novel image domain-based method to detect PAC/PVCs from AF. To the best of our knowledge, we are one of the first to propose this density Poincaré plot approach to discriminate AF from PAC/PVCs. The ECG recordings obtained from three distinct different data sets: critically ill MIMIC III database, the wearable armband data and the Physionet AFPDB, were used in this study. From the proposed density Poincaré plots, several image domain-based approaches were used to extract suitable features followed by infinite latent feature selection method to rank the individual features. Next, the extracted features were fed into several machine learning classifiers to detect PAC/PVC. Finally, we compared our method with four state-of-the-art AF detection methods.

II. Dataset Description

Three different data sets are used in this study:

A. MIMIC III Waveform Subset

ECG recordings obtained from 20 subjects of the Medical Information Mart for Intensive Care (MIMIC) III data set were used. MIMIC III is a large open source medical record database publicly available in PhysioNet [25]. MIMIC III contains de-identified health-related data from patients who stayed in critical care units of the Beth Israel Deaconess Medical Center between 2001 and 2012 [26]. However, no annotations were provided for this data set. As a result, ECG signals were annotated by board certified physician specializing in AF management (DDM). According to the physician's annotations, 10 subjects had AF whereas the rest 10 had PAC/PVCs. For this study, we do not discriminate PACs from PVCs; hence they are referred as PAC/PVC here. Moreover, these 20 subjects were identified to have sepsis according to the international classification of diseases, ninth revision (ICD-9) codes [27]. The ECG recordings were sampled at 125 Hz and the measuring unit was millivolt (mV).

B. Wearable Armband Database

The armband ECG data were collected using a recently developed novel wearable armband device [28]. It consists of three pairs of carbon black dry electrodes for recording three different ECG channels. The study data consisted of seven participants (over 21 years of age) recruited at the University of Massachusetts Memorial Health Care (UMMHC) system. Among the subjects, two subjects had persistent AF whereas five had PAC/PVC rhythms. Informed consents were obtained from the participants prior to their consented procedures. Trained research staff instructed the participants on the appropriate use of the wearable armband device and discussed study procedures. Next, the armband was placed on the left upper arm of the participants while they were simultaneously monitored on ECG telemetry prior to their procedure. Research staff returned to the participant's bedside after the procedure and prior to hospital discharge, they placed the ECG armband and an FDA cleared mobile cardiac telemetry patch (Cardiac Insight, Inc) on the participant, and asked each participant to wear both devices for 14 days. The ECG recordings from the patch were used as the reference. This protocol was approved by the institutional review board of the University of Massachusetts Memorial Hospital (H00013799). The armband had sampling frequency of 1000 Hz; ECG signals were down-sampled to 250 Hz.

C. AFPDB From Physionet

AFPDB data set has been collected from the publicly available paroxysmal atrial fibrillation prediction challenge, which was part of the Computers in Cardiology in 2001 [25]. The learning set consists of 30-minutes of two-channel ECG recordings from 25 subjects. ECG tracings were obtained from the Holter recordings with sampling frequency of 128 Hz. The ECG segments were collected just prior to the onset of paroxysmal AF (PAF). It is shown that for most of the cases, the frequency of PAC events increases before the onset of PAF. Details about this data set can be found in [29]. From the ECG recordings prior to PAF onset, we divide the recordings into two-minutes of non-overlapping segments and observed whether or not the individual two-minute segment contained PACs. By following this procedure, we obtained 124 segments from 13 different subjects and used these two-minute segments to evaluate the specificity (PAC detection accuracy) of our proposed method.

III. Methods

Our method consists of several major steps: preprocessing and generation of the density Poincaré plot, feature extraction, feature selection and classification. From the density Poincaré images, 79 features are extracted using 5 different approaches. Next, infinite latent feature selection algorithm is applied to rank the features. Finally, different classification algorithms are used

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

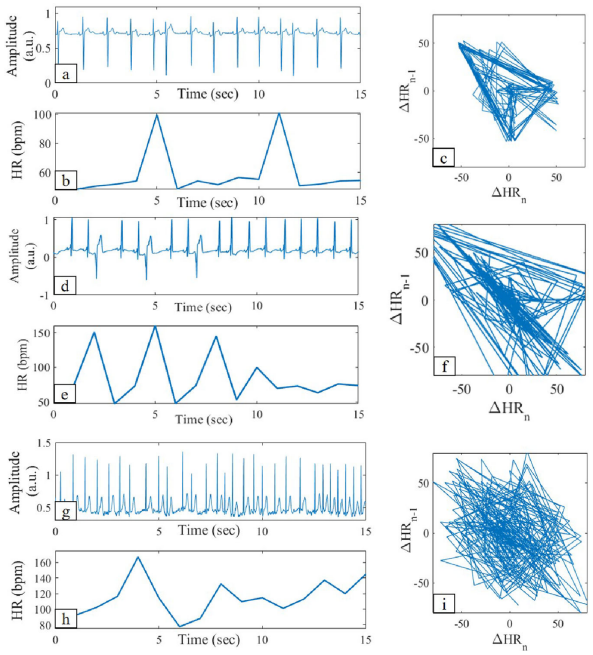

Fig. 1. (Left) Representative examples of 15-second ECG segments and the corresponding HR. (Right) Poincaré plots obtained from the HR. (a-c): from PAC, (d-f): from PVC and (g-i): from AF segments.

to discriminate the AF vs. PAC/PVC segments using both the selected and all features.

A. Preprocessing

ECG recordings are divided into 2 minutes non-overlapping segments and all input ECG signals are checked for noise artifacts using an automated motion noise detection algorithm [30]. After the clean segments are detected, QRS peaks are obtained by using the VERB algorithm, which performs ECG reconstruction to enhance the R-peak detection [31]. From the detected peaks, heart rate (HR) is calculated and finally difference of the HR (DHR) is obtained.

B. Density Poincaré Plot Generation

First, Poincaré plot is generated from the 2-minute input ECG segments. From the DHR, Poincaré plot is generated by connecting the two coordinates. The starting point coordinates consist of current DHR beat as the X-axis coordinate and the previous DHR beat as the Y-axis coordinate, and the ending point coordinates consist of the next DHR beat as the X-axis coordinate and the current DHR beat as the Y-axis coordinate. The range of X and Y axis for the Poincaré plots are arbitrarily set to [-80, 80] beats per minute (bpm). Details about the Poincaré plot can be found in [13], [32].

Representative examples of ECG segments and the corresponding HR along with Poincaré plots (obtained from 2-minute ECG) are shown in Fig. 1 for PAC (a-c), PVC (d-f) and AF (g-i) segments. From the figure, it is evident that for both of the PAC and PVC segments, we have repeated triangular patterns (i.e., kite shape in Fig. 1 (c) and Fig. 1 (f)) whereas for the AF segments, the HR varies in a random manner (Fig. 1 (i)).

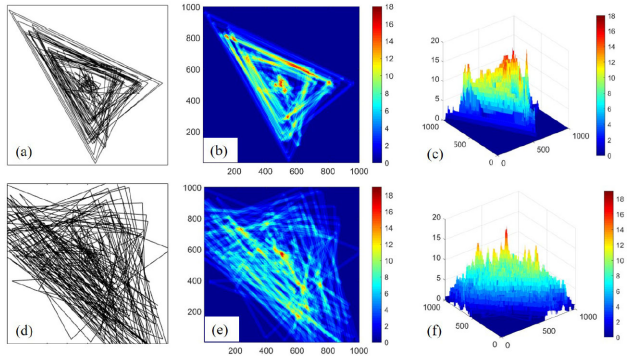

Fig. 2. (a-c) Binary Poincaré plot, density Poincaré plot and 3-D density plots, respectively, for a sample PAC segment. (d-f) Similar plots for an AF segment.

However, this pattern of well-behaved Poincaré plot will misdetect PAC/PVCs if 'kite shape is not obvious due to heart rate trajectories that do not perfectly overlap with each other, or due to other irregular heartbeats which may distort the repeated triangular/kite shape. As a result, a new "Density" Poincaré plot is proposed to better capture the overlapped trajectory information of the binary Poincaré plot. Note that AF rhythms are not likely to exhibit repeatable overlapped patterns whereas PAC/PVC rhythms do. Thus, we use this density approach to extract this additional information in addition to the kite shape for better discrimination between AF and PAC/PVC rhythms.

To generate the density Poincaré plot, we first increase the resolution of the plot to a predefined value of 1000 although any high resolution can be used. Similar to the binary Poincaré plot, a new line with pixel value 1 is drawn using the DHR coordinates and stored in a temporary resolution buffer matrix. The width of the line is arbitrarily set to 20. For each pair of coordinates, a new line is generated and added to the buffer matrix. At the end of this procedure, a matrix which has the summation of all the lines generated from every DHR coordinate pair is obtained. This matrix informs how many times the lines are overlapped, which indicates the density of the overlap, hence, this matrix is defined as the density Poincaré plot. Figs. 2 (b) and (e) show the density Poincaré plots obtained from both the AF and PAC Poincaré plots shown in Figs. 2 (a) and (d). From the figure, it can be seen that the density plots now have varying colors which correspond to different densities. The density distribution is more visible in the 3-D density plot shown in Figs. 2 (c) and (f).

C. Feature Extraction

From the density Poincaré images, different features are extracted using several image processing methods including statistical moments, correlation with image templates, Zernike moments, discrete wavelet transform and Hough transform.

1) Moment Based Features: We have used moment-based features as our first statistical descriptive measurement. The central moment-based features can be easily calculated and widely used in bio-signal and image processing [33], [34]. The central moments are defined as follows:

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

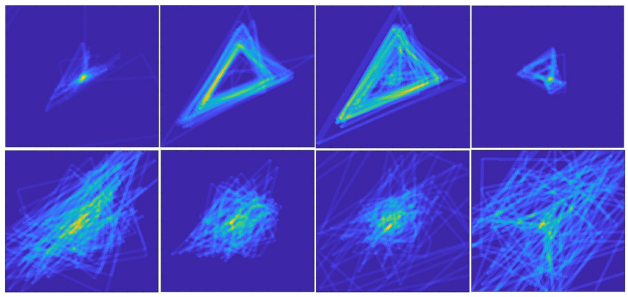

Fig. 3. (Top row) 4 template images for PAC/PVCs. (Bottom row) 4 template images for AF.

The p^{th} central moment m_p of a random variable X is determined as [35]:

$$m_p = E[(x - E[x])^p] \tag{1}$$

where E is the expected value and p is the order.

The "zeroth" central moment is 1 and the first central moment is 0. From the density Poincaré images, the second, third and fourth order central moments are calculated which are also known as variance, skewness and kurtosis, respectively.

Since the input is a $2D$ image, we first calculate different order moments (n = 2, 3 and 4 i.e., 2nd, 3rd and 4th) based on the row direction of the $2D$ image matrix and obtain a one-dimensional vector of moments. Next, from this $1D$ vector, we calculate the mean and variance to obtain one feature value for each moment and obtain 6 features. By following similar procedure for the column direction, 6 more features are calculated. Moreover, the mean (i.e., the first non-central moment) of the input $2D$ image is calculated and as a result, 13 features are obtained from the statistical moment-based approach.

2) Template Correlation Based Features: Image template correlation is popular in pattern recognition for image processing [36], [37]. For this study, the idea is to calculate the $2D$ correlation with some template images and use that correlation value as the AF vs. PAC/PVC discriminating feature. The $2D$ correlation coefficient between an input image W and a template image T, both having $m \times n$ pixels, is calculated as the following [38]:

$$r = \frac{\sum_m \sum_n (W_{mn} - \bar{W})(T_{mn} - \bar{T})}{\sqrt{\sum_m \sum_n (W_{mn} - \bar{W})^2 \times \sum_m \sum_n (T_{mn} - \bar{T})^2}} \tag{2}$$

Where $\bar{W}$ and $\bar{T}$ are the mean of the images W and T, respectively.

From the density Poincaré images, 14 representative images are visually selected as template images. Among these templates, 7 are selected from 4 PAC/PVC subjects and 7 are selected from 3 AF subjects. Fig. 3 shows some of the representative templates. These 14 template images are considered as the training data for the correlation coefficient feature and excluded from the testing data set.

For each input density Poincaré image, the $2D$ correlation coefficient values with all the 14 template images are calculated according to (2) and these correlation values are considered as features; thus, resulting in 14 features for each input image. It is to be noted that before calculating the $2D$ correlation coefficient, both images are normalized in [0, 1] range.

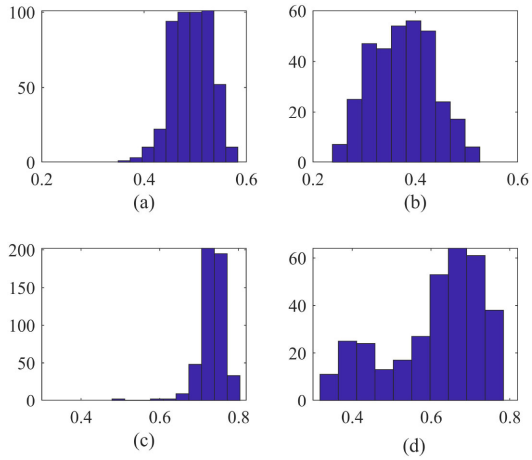

Fig. 4. Histograms of correlation coefficient with template #14 (a-b) and template #9 (c-d). (a) and (c) are from AF segments whereas (b) and (d) are from PAC/PVCs.

Fig. 4 shows the histogram distributions of the correlation coefficients obtained from two different template images. Figs. 4 (a-b) show the correlation coefficient distribution with template #14 for AF and PAC/PVC segments, respectively. Figs. 4 (c-d) show similar distributions obtained from correlation with template #9 for AF and PAC/PVC segments, respectively. From the histogram distributions, it is clear that the correlation coefficient distributions tend to span different ranges for AF and PAC/PVC segments although there are some overlaps.

3) Zernike Moments: Zernike moments are the mapping of an image onto a set of complex Zernike polynomials [39] and are widely used in image retrieval [40], face recognition [41], extracting shape properties of masses [42] etc.

From an input image, Zernike moments are computed using three steps: computation of radial polynomials, computation of Zernike basis functions and finally, computation of Zernike moments by projecting the image onto the basis functions [39], [42].

First, Zernike radial polynomials are calculated. The real-valued $1D$ radial polynomial $R_{nm}(\rho)$ is defined as:

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \times \frac{(n-s)!}{s!((n+|m|)/2 - s)!((n-|m|)/2 - s)!} \rho^{n-2s} \tag{3}$$

Here n is a non-negative integer and called the order. M denotes repetition which is an integer satisfying constrains $n - |m| =$ even and $|m| <= n$.

Using the radial polynomial equation, the complex-valued $2D$ Zernike basis functions are formed as the following which are defined within a unit circle:

$$V_{nm}(\rho, \theta) = R_{nm}(\rho) e^{jm\theta}, \quad |\rho| \leq 1 \tag{4}$$

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

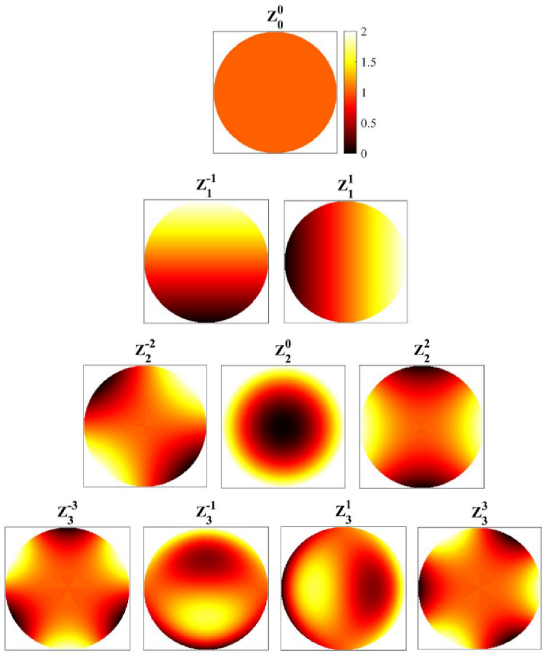

Fig. 5. Magnitude plots of some low-order Zernike moments in the unit disk.

Complex Zernike moments of order n and repetition m are defined as

$$Z_{nm} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho,\theta) V_{nm}^*(\rho,\theta) \rho \, d\rho \, d\theta \qquad (5)$$

Here $f(x,y)$ is the image function. Zernike moments can be regarded as an inner product between the image function and the Zernike basis function.

For digital images, the integrals are replaced by summations and the image coordinates are normalized into [0,1] by a mapping transform. The discrete form of the Zernike moments of an image size $N \times N$ is given as follows:

$$\begin{aligned} Z_{nm} &= \frac{n+1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) V_{nm}^*(x,y) \\ &= \frac{n+1}{\lambda_N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) R_{nm} \rho_{xy} e^{-jm\theta_{xy}} \end{aligned} \qquad (6)$$

Where λ_N is a normalization factor and $0 \leq \rho_{xy} \leq 1$.

The transformed distance ρ_{xy} and the phase θ_{xy} at the pixel (x,y) are given by

$$\begin{aligned} \rho_{xy} &= \frac{\sqrt{(2x-N+1)^2 + (N-1-2y)^2}}{N} \\ \theta_{xy} &= \tan^{-1}\left(\frac{N-1-2y}{2x-N+1}\right) \end{aligned} \qquad (7)$$

In this paper, the magnitudes of the Zernike moments are used as features as they are proper descriptors of shape characteristics [42]. The magnitude plots of some low order Zernike moments in the unit disk are illustrated in Fig. 5. Details about Zernike moments can be found in [39], [42].

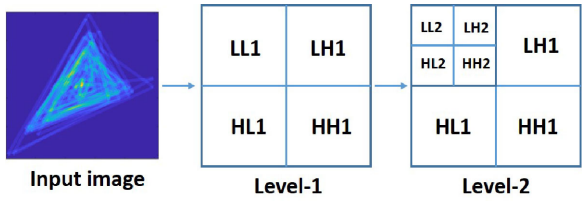

Fig. 6. Two-level wavelet decomposition structure for a sample density Poincaré image.

4) Discrete Wavelet Transform Features: Discrete wavelet transform (DWT) has become a popular mathematical tool for feature extraction from signals and images [43], [44]. DWT analyzes the signal or images at different scales and provides time-frequency spectral information [45]. The two-dimensional ($2D$) DWT for images consists of two general stages. First, a $1D$ DWT is applied to the input image followed by vertical subsampling to obtain the low-pass subband (L) and the high-pass subband (H). Next, another $1D$ DWT is applied to L and H followed by horizontal subsampling and four subbands namely LL, LH, HL and HH are obtained, respectively [46]. These 4 subbands matrices are called approximation, horizontal, vertical and diagonal coefficients, respectively. The entire frequency spectrum of the original image is covered by those four sub-bands. In this study, the $2D$ DWT of the density Poincaré image has been performed using the Daubechies 4 (DB4) wavelets and two-levels of wavelet decomposition has been performed (Fig. 6).

Fig. 7 shows the four coefficients at two levels for a sample PAC density Poincaré image while Fig. 8 shows the same for a sample AF image.

After the DWT, from each coefficient matrix, the spectral energy and entropy are calculated. Entropy is a statistical measure of randomness and can be used to characterize the texture of the input image. Entropy is defined as $-\sum p \times log_2(p)$, where "p" contains the normalized histogram counts which represents the probability. Our assumption is that by calculating entropy and energy from different DWT coefficients, we can obtain features which can be used to classify AF vs. PAC/PVCs. Moreover, normalized entropy is calculated which is the entropy after normalizing the DWT coefficient matrices from [0, 1] range. As a result, from 2 levels of decomposition, 8 energy features and 16 entropy features are obtained.

5) Hough Transform Features: Hough transform is a popular method for extracting global features such as straight lines, circles, etc. from an image [47], [48]. Standard Hough transform (SHT) is widely used in computer vision and pattern recognition [49], [50]. It is a voting process where each point belonging to the patterns votes for all the possible patterns which pass through that point. Finally, these votes are accumulated and the pattern receiving the maximum vote is recognized as the desired pattern [51].

For a $N \times N$ binary image, the straight lines are defined as

$$\rho = x\cos\theta + y\sin\theta \qquad (8)$$

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

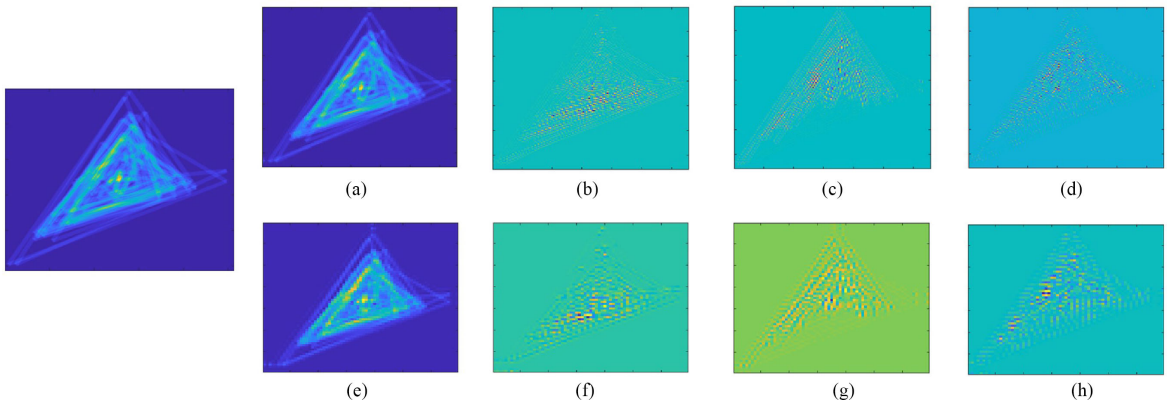

Fig. 7. (Left) The input PAC density image. (a-d) Four coefficients obtained from level 1 of the DWT. (e-h) Four coefficients obtained from level 2 of the DWT. The four coefficients are approximation, horizontal, vertical and diagonal, respectively.

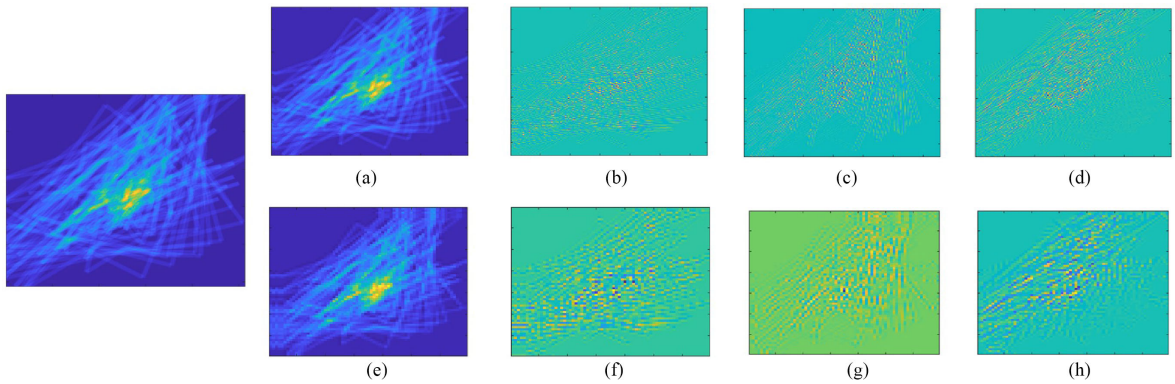

Fig. 8. (Left) The input AF density image. (a-d) Four coefficients obtained from level 1 of the DWT. (e-h) Four coefficients obtained from level 2 of the DWT. The four coefficients are approximation, horizontal, vertical and diagonal, respectively.

where (x, y) is the coordinates in the XY plane, θ denotes the angle the normal line makes with X-axis and ρ is the perpendicular distance from the origin.

In the SHT, (ρ, θ) is defined as the parameter space. Computation of the SHT consists of three major parts [51]: (1) determining the parameter values and accumulating the bins in the parameter space, (2) finding the local maxima which represent line segments and (3) extracting the line segments. The SHT provides a parameter space matrix (called Hough transform matrix) whose rows and columns correspond to ρ and θ respectively [52]. Fig. 9 shows the Hough transform matrix obtained from both PAC and AF density images.

From this Hough transform matrix, 13 statistical features are extracted using the process described in "section III-C-1: *Moment Based Features*. Moreover, entropy of the Hough transform matrix is also used; resulting in a total of 14 features from the SHT.

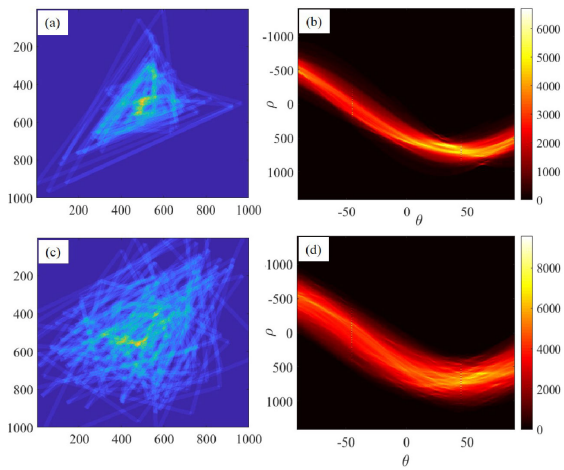

Fig. 9. (b) Hough transform matrix obtained from a PAC image shown in (a). (d) Hough transform matrix obtained from an AF image shown in (c).

D. Feature Selection

Feature selection plays an important role in machine learning as the performance of machine learning methods heavily depends on the choice of features. Different features can entangle and hide the different explanatory factors of variation behind the data [53]. There are several feature selection methods available in the literature. In this study, we have chosen a recently developed feature selection method called infinite latent feature selection (ILFS). This method is based on probabilistic

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

latent graph-based feature selection algorithm that performs the ranking by considering all the possible subsets of features. Relevancy is modeled as a latent variable in this method. ILFS consists of three major steps. At first, a discriminative quantization process is performed which maps the raw feature values to a countable nominal set of tokens. Next, an undirected fully connected graph is created where each node represents one feature. Weights are learnt automatically by a learning framework which is based on a variation of the probabilistic latent semantic analysis. Finally, a ranking is done considering all possible paths among the nodes while investigating the redundancy of the features. Details of the ILFS method can be found in [53].

E. Classification

From the five different feature extraction approaches, a total of 79 features are calculated. These features include 13 from statistical moments, 14 from correlation with templates, 14 from Zernike moments, 24 from DWT and 14 from standard Hough transform. These features are next fed into different machine learning classifier models. Moreover, classifiers with only the selected features from the ILFS algorithm are also studied. The studied classifiers include support vector machine (SVM), K-nearest neighbor (KNN) and random forest (RF).

1) K-Nearest Neighbor Classifier: K-NN classifier is an supervised machine learning method with desirable computational speed along with acceptable classification accuracy [54]. Here "K" refers to the number of samples to be used for the classification task. For any new sample, distances from "K" closest neighbors are calculated and based on the majority voting, that new test sample is classified.

Let for a training point $x_i \epsilon D$, the class label is denoted as $c^{x_i} \epsilon\, C$ where $i = 1, 2, ., n$. The set of all training points belonging to class c is denoted by C_c and $|C_c| = n_c$. The K-NNs of a data point $x \epsilon\, X$ from a set S form the set $\Gamma_K^S(x)$. $I(.)$ is an indicator function defined as follows:

$$I(cond) = \begin{cases} 1, & if\ cond\ is\ true \\ 0, & if\ cond\ is\ false \end{cases} \tag{9}$$

Here $cond$ is some condition to be satisfied [55]. The basic K-NN classification rule can be described as: For a test point y_i if the K-NNs of y_i in the training set D form the set $\Gamma_K^D(y_i)$ then the predicted class label of y_i can be denoted as:

$$\hat{c}^{y_i} = \underset{c \epsilon 1,2,....,C}{\arg\max} \sum_{x_i \epsilon \Gamma_K^D(y_i)} I(c^{(x_i)} = c) \tag{10}$$

For calculating the similarity condition, most popular "Euclidean" distance along with "city-block" distances are considered here. "City-block" is also known as Manhattan distance or absolute value distance. Moreover, the number of nearest neighbors "K" is also varied to obtain best performance.

2) Support Vector Machine Classifier: Support vector machine (SVM) is a very popular and well-established classifier for binary problems where a maximum margin between the training and test data is constructed [56]. The samples which are closest to the decision boundary are called the support vectors.

Let vector $\mathbf{x}$ denotes a test sample to be classified and its label is denoted by c, where $c \epsilon \{+1, -1\}$. Now, for a given set of training data, $\{(\mathbf{x}_i, c_i), i = 1, 2, ..., l\}$, the separating hyperplanes can be obtained by maximizing the margin, which is the minimization of the following function [56]

$$\mathbf{J}(\mathbf{w}, \xi) = \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_i \xi_i \tag{11}$$

with the following constraints:

$$c_i(\mathbf{w}^T\mathbf{x} + b) \gg 1 - \xi_i \quad where\ \ \xi_i \geq 0 \tag{12}$$

Here $\mathbf{w}$ is the weight vector, b is a constant, C is a positive regularization parameter and ξ_i is the slack variable.

Applying the Lagrange multipliers α_i, for vector $\mathbf{x}$, the solution of the decision function can be expressed as

$$f(\mathbf{x}) = \sum_i \alpha_i c_i \mathbf{x}_i^T\mathbf{x} + b \tag{13}$$

For the nonlinear SVM, a nonlinear mapping function $\phi(\mathbf{x})$ is used to map the input data into a higher dimensional feature space, thus making the samples more separable:

$$f(\mathbf{x}) = \sum_{j=1}^{N} \alpha_j c_j K(\mathbf{x}_j, \mathbf{x}) + b \tag{14}$$

where $\mathbf{x}_j$ are the support vectors and $K(\mathbf{x}_j, \mathbf{x})$ is the kernel function. For radial basis function, $K(\mathbf{x}_j, \mathbf{x}) = \exp(-\gamma\|\mathbf{x}_j - \mathbf{x}\|^2)$. Details of the SVM can be found in [57]. For this study, both the linear and RBF SVM were used.

3) Random Forest Classifier: Bootstrap-aggregated (bagged) decision trees can reduce the effects of overfitting and improves generalization by combining the results of many decision trees [58].

In the bagging learning concept, T weak learners are trained over subsets drawn with replacement from the training set and their outputs are voted to produce a predictive estimate of the model. By doing this bagging and voting, it is shown to decrease the variance of the model without increasing the bias as the weak learners are provided with different training sets; thus, resulting in diverse ensemble [59].

From the original set of M training instances, a bootstrap of M' instances is drawn uniformly at random with replacement for every tree $t \epsilon \{1, 2, ..., T\}$. During the training process, at each node of every tree, $D' < D$ features are randomly selected from D available features and the best split is decided using those D' features. For testing, a new sample is run down all the T trees of the forest; thus producing T predictions for the test sample. At the end, these predictions are aggregated through voting to make the final prediction.

IV. Experimental Results

In this section, we first present the feature selection results which is followed by AF vs. PAC/PVC classification performed for both segment-wise and subject-wise scenarios. Our MIMIC III waveform subset consists of 20 subjects with 10 AF and 10 PAC/PVC; resulting in a total of 500 AF and 340 PAC/PVC segments. However, after the 14 templates are discarded from

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

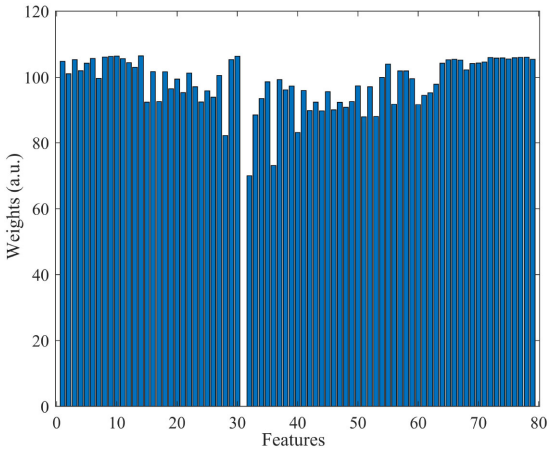

Fig. 10. Bar chart of all the feature weights determined by the ILFS.

the data set, our final database now consists of 493 AF and 333 PAC/PVC segments.

For segment-wise cross validation, widely used k-fold cross validation is used in this study. In k-fold, data is divided into k number of non-overlapping segments. Every time $(k-1)$ folds are used for training and the rest fold is used for testing. The entire process is repeated over k times. $k = 10$ is used in this study. For subject-wise cross validation, leave-one-subject-out procedure is followed as this is the best approach to account for subject variability.

In order to evaluate the classification performance, the following binary classification accuracy measures are calculated:

$$\begin{aligned} &Sensitivity = TP/(TP+TN) \\ &Specificity = TN/(TN+FP) \\ &Accuracy = (TP+TN)/(TP+FN+TN+FP) \\ &Positive\ predictive\ value\ (PPV) = TP/(TP+FP) \\ &Negative\ predictive\ value\ (NPV) = TN/(TN+FN) \end{aligned} \tag{15}$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

A. Feature Selection Results

During each step of the k-fold cross validation, ILFS based feature selection is performed using only the training data. The ILFS algorithm ranks every feature from the total pool of 79 features and assigns each feature a weight. Fig. 10 shows the bar chart with the ILFS weights for individual feature for a particular fold. From these ranked features, top 10 features for each validation stage are selected automatically. For the particular fold shown in Fig. 10, 4 correlation-based features, 1 Zernike moments-based feature and 5 wavelet transform-based features are ranked as the top 10 features.

Next, the top 10 ranked features for each of the 10-folds of the training data are analyzed and it is found that 9 features are common in all those 10 folds. As a result, these 9 features are selected later for the subject-wise classification. Among these 9 features, 4 comes from the correlation features, 1 from the Zernike moment feature and rest 4 are from the DWT features. These are: correlation coefficient values with templates #8, #9, #10 and #14, amplitude of zernike moment (at $n = 2$, $m = 0$), normalized entropy of level-2 vertical, approximation, horizontal and level-1 approximation coefficients of the DWT.

B. Classification Results

Classification result is divided into two parts: (1) segment-wise k-fold cross validation and (2) subject-wise classification. For both scenarios, AF vs. PAC/PVC detection performance using the abovementioned machine learning classifiers are presented for both the automatically selected (via ILFS) and all 79 features.

1) Segment-Wise Classification Results: Segment-wise classification results include 10-fold cross validation results for all 79 features as well as the automatically selected 10 features by the ILFS algorithm.

a) KNN classifier: For the KNN classifier, the hyperparameter value (K) was varied for optimal decision. It was found that "Manhattan" distance (i.e., city-block distance) with K = 7 provided the best results. During 10-fold cross-validation, with 10 selected features, KNN classifier achieved sensitivity, specificity, accuracy, PPV and NPV of 97.57%, 73.56%, 87.89%, 84.92% and 95.41%, respectively. When all the 79 features were used, KNN achieved sensitivity, specificity, accuracy, PPV and NPV of 84.01%, 64.55%, 76.16%, 78.01% and 74.16%, respectively.

b) SVM classifier: For the SVM, both the linear and radial basis kernels were used. Linear SVM resulted in 94.93% sensitivity, 90.35% specificity and 93.09% accuracy for the 10 features while for 79 features, we found 98.18% sensitivity, 93.98% specificity and 96.48% accuracy.

For the RBF SVM, with 10 features and sigma = 1, 97.97% sensitivity, 91.26% specificity and 95.27% accuracy was obtained. When all the 79 features were included, RBF SVM achieved sensitivity, specificity and accuracy value of 98.99%, 95.18% and 97.45%, respectively with sigma = 3.50.

c) Random forest: For the random forest classifier, the number of trees and number of predictors to sample were varied. With the selected 10 features, RF obtained 97.78% sensitivity, 90.38% specificity and 94.79% accuracy. When all the 79 features were used, the accuracy increased to 97.09%. 100 and 200 tree were used for 10 and 79 features, respectively. Number of trees is an important parameter in the RF model. To determine the optimal number of trees, the out-of-bag (OOB) classification error was calculated. Fig. 11 shows the OOB error for 10-fold cross-validation when all 79 features are used. From the figure it can be seen that OOB error decreases with the number of grown trees and near 200 trees, the OOB error becomes flat. Table I compares the classification accuracy, sensitivity, specificity, PPV and NPV for RF, SVM and KNN for both cases: 10 best selected features (from ILFS) and all 79 features. From the table it is clear that RBF SVM achieved the best accuracy.

2) Subject-Wise Classification Results: The motivation to perform subject-wise cross-validation is that AF/PAC/PVC

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

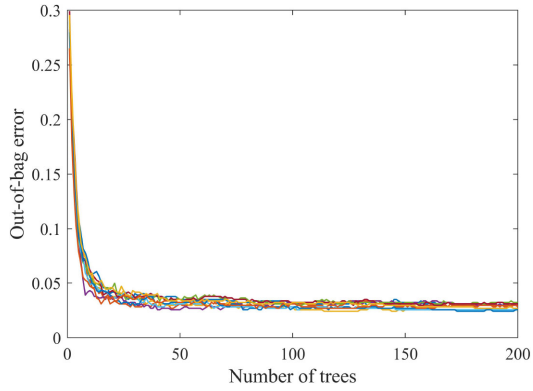

Fig. 11. Out-of-bag error of RF during 10-fold cross-validation with all 79 features.

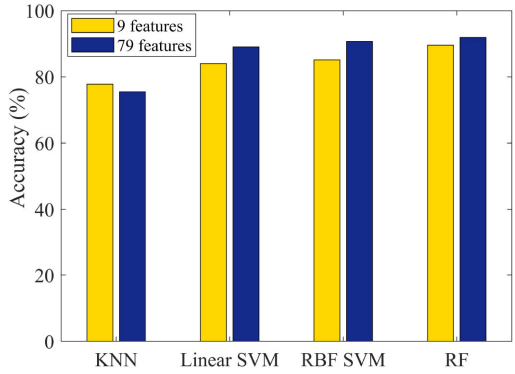

Fig. 12. Bar chart comparing subject-wise classification accuracies for four classifiers. Both of 9 and 79 feature scenarios are considered.

TABLE I
PERFORMANCE OF DIFFERENT CLASSIFIERS: 10 AND 79 FEATURES

	Classifier Name	KNN (%)	SVM Linear (%)	SVM RBF (%)	RF (%)
10 features	Sensitivity	97.57	94.93	97.97	97.78
	Specificity	73.56	90.35	91.26	90.38
	Accuracy	87.89	93.09	**95.27**	94.79
	PPV	84.92	93.68	94.41	93.83
	NPV	95.41	92.46	96.87	96.6
79 features	Sensitivity	84.01	98.18	98.99	99.4
	Specificity	64.55	93.98	95.18	93.68
	Accuracy	76.16	96.48	**97.45**	97.09
	PPV	78.01	96.08	96.87	95.93
	NPV	74.16	97.3	98.53	99.07

patterns and features can vary significantly depending on the cardiovascular health state of subjects. For the 20 subjects, leave-one-subject-out cross-validation is performed to account for subject variability. Every time the algorithm is trained with 19 subjects and tested with the remaining one subject; the entire process is repeated 20 times and the average results are reported. Similar to the segment-wise scenario, different classifiers are implemented with 9 automatically-selected features and all of 79 features.

KNN classifier achieved 77.81% (K = 3) and 75.51% (K = 11) average accuracy across 20 subjects when 9 and 79 features, respectively, were used. The "Manhattan" distance was selected for both models. Linear SVM obtained 84.03% accuracy for 9 features while the performance improved to 89.05% for all 79 features. The RBF SVM achieved 85.16% accuracy when 9 features were used; hyper parameter sigma = 1.15. The accuracy increased to 90.68% when all 79 features were provided with sigma = 10.50.

Random forest achieved an overall accuracy of 89.56% for 9 selected features; 200 trees were used while 4 predictors were used to sample in the model. When all the 79 features were used, the accuracy was 91.93% with 100 trees. Fig. 12 shows the bar-chart showing the overall accuracies obtained by different classifiers for both scenarios: 9 and 79 features.

The accuracy for each subject when all 79 features are used, is reported in Table II. The 10 AF and 10 PAC/PVC subjects

TABLE II
SUBJECT-WISE CLASSIFICATION ACCURACY WITH 79 FEATURES

Subject ID	Segment number	KNN (%)	Linear SVM (%)	RBF SVM (%)	RF (%)
AF #1	50	98	100	100	100
AF #2	50	100	100	100	94
AF #3	50	86	98	100	100
AF #4	50	20	82	94	100
AF #5	50	96	94	100	100
AF #6	50	98	90	92	94
AF #7	50	98	100	100	100
AF #8	47	97.87	100	59.57	93.62
AF #9	49	83.67	75.51	71.43	59.18
AF #10	47	38.30	87.23	93.62	97.87
PAC #1	3	100	100	100	100
PAC #2	9	88.89	88.89	88.89	100
PAC #3	50	70	80	74	66
PAC #4	34	55.88	100	100	100
PAC #5	63	30.16	71.43	100	100
PAC #6	50	92	78	100	98
PAC #7	50	40	36	40	36
PAC #8	19	47.37	100	100	100
PAC #9	5	100	100	100	100
PAC #10	50	70	100	100	100
Mean	Total =826	75.51	89.05	90.68	91.93

are denoted with "AF #" and "PAC #" followed by a segment number. "Segment number" denotes the number of 2 minute segments for each subject. Although for the PAC/PVC subjects, the number of segments are uneven, the overall mean accuracy is satisfactory.

C. Comparison With Existing Methods

In Table III, the performance of our proposed method is compared with four previously published methods by examining AF accuracy, PAC accuracy, and overall accuracy. For all these methods reported, the leave-one-subject-out cross-validation was used. The root mean square of successive differences, Shannon entropy, and turning point ratio calculated from RR intervals were used in the statistical method [6] to determine AF. The Lorentz plot generated from an RR series was described in [7] to

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

TABLE III
PERFORMANCE COMPARISON WITH EXISTING METHODS

Method	AF Accuracy (%)	PAC Accuracy (%)	Overall Accuracy (%)
Statistical method [6]	95.60	26.96	61.28
Lorentz plot [7]	78.20	87.20	82.70
MCCE [11]	77.08	74.18	75.63
Predictive power of AR [11]	88.52	75.60	82.06
COSEn [8]	80.06	80.16	80.11
Our method with SVM	91.06	90.29	90.68
Our method with RF	93.87	90	91.93

TABLE IV
CLASSIFICATION PERFORMANCE ON ARMBAND DATA

Subject ID	No of seg.	Accuracy (%)				
		SVM 79 feat.	*SVM 9 feat.*	*RF 79 feat.*	*RF 9 feat.*	*Stat. Method [6]*
AF #1	50	76	88	86	86	100
AF #2	14	100	92.86	64.29	100	100
PAC #1	55	100	100	100	100	85.45
PAC #2	55	100	100	100	100	96.36
PAC #3	10	100	100	100	100	100
PAC #4	30	100	100	100	100	96.67
PAC #5	50	100	100	100	100	100
Mean	Total =264	96.57	**97.27**	92.9	**98**	96.93

detect AF. Dynamics of RR intervals were analyzed to determine AF episodes in [11]; auto regressive (AR) model order-based percentage of predicted power was shown to have the best performance in [11] while the minimum of corrected conditional entropy (MCCE) obtained from RR series also achieved good performance in [11], [60]. Finally, the coefficient of sample entropy (COSEn) calculated from RR interval series [8] was also compared in Table III. It is evident from Table III that our density Poincaré image-based features obtained better results for both RBF SVM and RF when compared to all the reported methods.

D. Results on Armband Data

Next, we tested our proposed density Poincaré plot method using the wearable armband ECG data (independent test set). From this data set, we have two AF subjects and five PAC/PVC subjects, resulting in 64 AF and 200 PAC/PVC segments of 2-minute duration. To classify AF vs. PAC/PVC segments from the armband ECG data, the SVM and RF models were used for both 79 and 9 selected features. It is to be noted that the RF and SVM models were trained with the MIMIC III data (20 subjects) and those trained models were directly applied to this armband data without any further tuning, resulting in a true test scenario. Table IV shows the performance of the SVM and RF on the armband data. From the table it is evident that both SVM and RF achieved better performance when 9 selected features were used rather than all 79 features. This shows the importance of feature selection for unknown/blind test data. When compared with statistical method [6], both SVM and RF achieved better overall accuracy (for 9 selected features); for PAC subjects, the accuracy was 100%.

E. Results on AFPDB

Finally, we tested the proposed method using AFPDB, which is another independent test set. From the AFPDB, we have 124 2-minute PAC segments from 13 different subjects. Since this data set does not include any AF, we used it to evaluate the specificity of our algorithm (i.e., how the method will work on a different PAC/PVC data set). The same SVM and RF models which were trained using only the MIMIC data, were directly applied to AFPDB without any further tuning. For both of the selected 9 features and all of the 79 features, the SVM and RF detected all of the 124 PAC segments correctly, resulting in 100% PAC detection accuracy (i.e., specificity).

V. DISCUSSION

We presented a novel density Poincaré plot-based approach to classify AF from PAC/PVCs using three different data sets: a subset of MIMIC III database, the wearable armband data and AFPDB. Most common PAC/PVCs are known to have repeated triangular kite shaped patterns in the Poincaré trajectory [7], [13], [32]. However, binary Poincaré plot only provides the triangular kite shape information and nothing about how many times the triangular patterns were formed in the Poincaré phase space. Moreover, other irregular heartbeats can also distort the repeated triangular pattern. AF rhythms due to their inherent random dynamics, do not exhibit repeatable random kite patterns. Our proposed density Poincaré plot was used to extract the overlapped information in the kite shape to better identify PAC/PVCs since the binary Poincaré plot itself only provides the kite shape information.

To detect AF from PAC/PVCs, a total of 79 features were extracted from five different image processing approaches- statistical moments, template-based correlation, Zernike moments, DWT and the Hough transform. Among these different types of features, template-based correlation and DWT features were most effective, which was also evident from the ILFS algorithm. One of the underlying reasons is that for PAC/PVCs and AF, distinct triangular kite shaped and random patterns are generated, respectively. As a result, high correlation values were obtained when templates from two different categories were used. For the DWT, when energy and entropy were calculated from different levels of the coefficient matrices, the patterns became separable as different resolutions provided more discriminatory information.

During the density Poincaré plot formation, the resolution was empirically set to 1000, although other resolutions can be used. To study the effect of image resolution on the algorithm performance, we have varied the resolution from 256 to 1500 and for each setting, the accuracy of the SVM classifier is reported with leave-one-subject-out scenario. Fig. 13 (a) represents the bar chart showing the overall accuracy for four

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

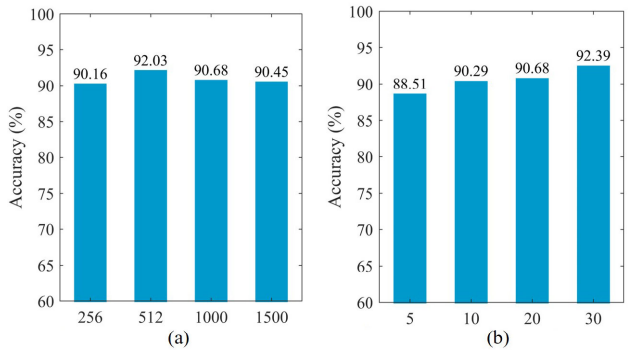

Fig. 13. Bar charts of overall accuracies for (a) different image resolutions and (b) different line widths.

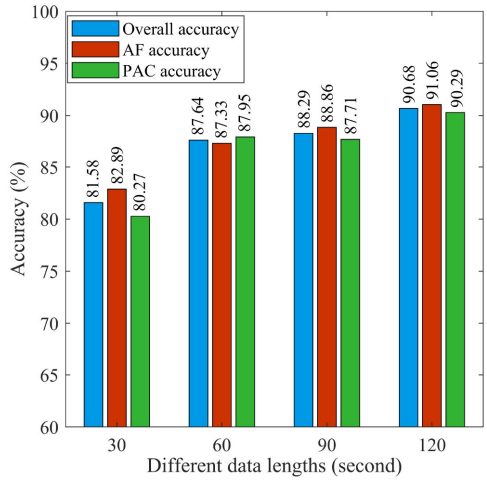

Fig. 14. Bar chart showing accuracy for different input data lengths.

different resolutions. From the figure, it can be seen that for all four resolutions, the algorithm has performance over 90% and the maximum accuracy is 92.03%. This shows that the density Poincaré resolution is flexible and can be chosen depending on the computational requirements, as higher resolution demands higher computation.

Similar to the density Poincaré resolution, the width of each new pixel line is set to 20 pixels to increase the probability that it can overlap with nearby trajectories. To study how the algorithm's performance changes with line width, we have varied the line width from 5 to 30 pixels, at a fixed resolution of 1000. Fig. 13 (b) shows the bar charts of overall accuracies for different line widths obtained from the SVM classifier during leave-one-subject-out cross-validation. With the increase of line width, the overall accuracies increase gradually as the overlap with nearby trajectories increases; the highest accuracy of 92.39% is obtained for a line width of 30, although others also have good performance.

SVM with radial basis function achieved the best performance with segment-wise 10-fold cross validation while for leave-one-subject-out cross validation, RF had the best result. During the segment-wise 10-fold classification, when all 79 features were used, the accuracies increased by a few percent for all the reported classifiers except for KNN. Similarly, for the subject-wise scenario, using all the 79 features improved the performance when compared to using only the selected 9 features, except for KNN. This indicates that optimized feature selection has higher impact on KNN. Moreover, when all 79 features were used, the accuracy of SVM RBF and RF did not differ much for both segment-wise and subject-wise validations.

For RF, 79 features needed nearly twice computational time than using 9 features. Moreover, the computational complexity was much higher for the 79 feature model when compared to 9 feature model. The $\sim$1% increase in overall accuracy is not that significant if the additional computational cost and model complexity are considered.

For generating the density Poincaré plots, ECG segments of 2-minute length were used. Since Poincaré density depends on the number of repetitions, to study the effect of input data length on the algorithm, we have varied the ECG segment length from 30 seconds to 2 minutes. For different ECG segment lengths, the leave-one-subject-out accuracy is calculated using the SVM classifier. Fig. 14 shows the bar chart of overall accuracy, AF accuracy, and PAC accuracy for different input data lengths. From the figure, it can be seen that with the increase of data (segment) length, the accuracy increases. For ECG segments of 60 seconds, the accuracy becomes 87.64% which gradually goes up to 90.68% for 2-minute segments; the algorithm achieved 81.58% overall accuracy for 30-second segments. This shows that the algorithm is robust to data lengths above 60 seconds. It is to be noted that even for the 30 second segments, the accuracy of our algorithm is comparable to the existing methods with 2-minute ECG segments, as shown in Table III, thus showing the efficacy of the proposed density Poincaré features.

We compared our proposed method with four existing AF detection methods. The Das *et al.* [6] method did not explicitly focus on PAC detection, but rather tried to filter ectopic beats as a preprocessing step to have better AF detection accuracy. That is why that method had the highest AF detection accuracy (sensitivity) but low specificity (i.e., low PAC detection accuracy) which is evident from Table III. On the other hand, the Lorentz plot-based AF detection method had better PAC detection accuracy than the statistical method. This is because [7] used the $2D$ histogram-based technique to look for PAC's kite shaped patterns. However, it achieved lower AF detection accuracy. As described in Cerutti *et al.* [11], the individual performance of MCCE and predictive power of AR models were implemented, with results reported in Table III. Despite being relatively simple, the AR model-based percentage of predicted power achieved similar performance to Lorentz plots; MCCE resulted in lower accuracy as reported in [11]. The COSEn-based method achieved 80.11% overall accuracy. Our proposed density Poincaré image domain-based method achieved the highest overall accuracy among all the compared methods (for either the SVM or RF classifier).

We analyzed the overall mean accuracy during leave-one-subject-out validation. For most of the subjects, the resulting

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

accuracy was good which showed the efficacy of our proposed method. However, for only a few subjects of the MIMIC data, particularly PAC #3 and PAC #7, the subject-wise accuracy was low. One possible reason could be that their PAC/PVC patterns differ from other PAC/PVC subjects which may also be due to noise contamination in the data. Another explanation can be that since not all PAC/PVC subjects had the same number of segments, our PAC/PVC data set may not have captured all the pattern variations. However, the difference in the segment number for PAC/PVC subjects is expected because PAC/PVCs can be paroxysmal and difficult to find. Given the challenge of capturing diverse PAC/PVC patterns, our proposed method achieved high PAC/PVC detection accuracy which is evident from the 100% specificity for the APFDB.

Leave-one-subject-out cross-validation was performed to prevent subject-wise dependency. Moreover, once we trained the model with the MIMIC data, the trained models were applied to the armband and AFPDB data sets without any further tuning. The high accuracy achieved on these two independent test data sets showed that our model was not over-fitted to the MIMIC data set and the proposed density Poincaré features captured the different patterns of AF and PAC.

The results on the wearable armband data are particularly interesting and show two important points. First, the selected 9 features achieved better results than all 79 features, which shows the importance of feature selection. One possible explanation for this can be that both models (with 79 and 9 features, respectively) were tuned using the MIMIC III ICU data; thus, the 79 feature based model achieved better results on MIMIC than that with 9 features. However, when tested on a different data set (armband) with pre-determined hyper-parameters, only the important/significant features resulted in better performance. Second, the proposed density Poincaré image-based method used only R-R interval information. For wearable devices like the armband, due to the characteristics of this device and extreme muscle artifact, the P-waves are not always visible. Consequently, the R-R interval-based PAC/PVC vs. AF classification approach is necessary for continuous AF monitoring with the wearable armband device.

There are some limitations of this study. We studied AF only based on the surface ECG recordings where QRS complex-based randomness is measured. However, based on the atrial electrogram recordings, AF can be divided into four types as shown in [61], [62]. This is beyond the current study as we do not have atrial electrogram recordings in the MIMIC III database. Based on only the surface ECG recordings, we cannot discriminate among different types of AF. Further experiments to study how different types of AF impact the algorithm are certainly warranted. Moreover, future studies can incorporate larger data sets and other approaches to cover different patterns so that the overall accuracy during the subject-wise scenario can be improved. With larger data sets consisting of PAC/PVC, deep learning approaches involving for example, convolutional neural networks or long short-term memory can be used. Certainly, given the limited PAC/PVC data sets, we did not perform deep learning to discriminate between AF and PAC/PVC in this work.

VI. Conclusion

In this study, we introduced a novel machine learning approach using the density Poincaré plot to perform PAC/PVC detection from AF. After the density Poincaré plots are generated from the difference of the heart rate, several image domain-based approaches are applied to extract suitable features. Moreover, ILFS based feature ranking was performed. The extracted features are fed into several machine learning-based classifiers to distinguish PAC/PVC from AF. High accuracy is obtained not only for 10-fold cross validation, but also for leave-one-subject-out cross validation. The subject-wise cross validation shows the effectiveness of the proposed method for unseen subjects which is most important for real-life clinical applications. When our proposed method is compared against four existing algorithms, better classification performance is obtained which demonstrates the efficacy of our method. Our density Poincaré image-based method also achieved high accuracy on the wearable armband ECG data and Physionet AFPDB databases, further validating the effectiveness of the proposed method. Future research can incorporate other image domain-based feature extraction methods to further improve performance.

References

[1] A. Murphy *et al.*, "The world heart federation roadmap for nonvalvular atrial fibrillation," *Global Heart*, vol. 12, no. 4, pp. 273–284, 2017.
[2] S. S. Chugh *et al.*, "Worldwide epidemiology of atrial fibrillation: A global burden of disease 2010 study," *Circulation*, vol. 129, no. 8, pp. 837–847, 2014.
[3] D. Jabaudon *et al.*, "Usefulness of ambulatory 7-day ECG monitoring for the detection of atrial fibrillation and flutter after acute stroke and transient ischemic attack," *Stroke*, vol. 35, no. 7, pp. 1647–1651, 2004.
[4] O. Lahdenoja *et al.*, "Atrial fibrillation detection via accelerometer and gyroscope of a smartphone," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 1, pp. 108–118, Jan. 2017.
[5] G. Moody and R. Mark, "A new method for detecting atrial fibrillation using RR intervals," in *Proc. Comput. Cardiol.*, 1983, pp. 227–230.
[6] S. Dash, K. Chon, S. Lu, and E. Raeder, "Automatic real time detection of atrial fibrillation," *Ann. Biomed. Eng.*, vol. 37, no. 9, pp. 1701–1709, 2009.
[7] S. Sarkar, D. Ritscher, and R. Mehra, "A detector for a chronic implantable atrial tachyarrhythmia monitor," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1219–1224, Mar. 2008.
[8] D. E. Lake and J. R. Moorman, "Accurate estimation of entropy in very short physiological time series: The problem of atrial fibrillation detection in implanted ventricular devices," *Amer. J. Physiol.-Heart Circulatory Physiol.*, vol. 300, no. 1, pp. H319–H325, 2011.
[9] C. Huang *et al.*, "A novel method for detection of the transition between atrial fibrillation and sinus rhythm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 4, pp. 1113–1119, Apr. 2010.
[10] A. Kennedy *et al.*, "Automated detection of atrial fibrillation using RR intervals and multivariate-based classification," *J. Electrocardiol.*, vol. 49, no. 6, pp. 871–876, 2016.
[11] S. Cerutti *et al.*, "Analysis of the dynamics of RR interval series for the detection of atrial fibrillation episodes," in *Proc. IEEE Comput. Cardiol.*, 1997, pp. 77–80.
[12] P. Langley *et al.*, "Accuracy of algorithms for detection of atrial fibrillation from short duration beat interval recordings," *Med. Eng. Phys.*, vol. 34, no. 10, pp. 1441–1447, 2012.
[13] J. W. Chong *et al.*, "Arrhythmia discrimination using a smart phone," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 3, pp. 815–824, May 2015.
[14] M. Henriksson *et al.*, "Model-based assessment of f-wave signal quality in patients with atrial fibrillation," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 11, pp. 2600–2611, Nov. 2018.
[15] S. Bashar *et al.*, "Atrial fibrillation detection during sepsis: Study on MIMIC III ICU data," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 11, pp. 3124–3135, Nov. 2020.

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.

[16] S. Sahoo *et al.*, "Multiresolution wavelet transform based feature extraction and ECG classification to detect cardiac abnormalities," *Measurement*, vol. 108, pp. 55–66, 2017.
[17] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, 2016, Art. no. 285.
[18] Y.-C. Yeh, W.-J. Wang, and C. W. Chiou, "Cardiac arrhythmia diagnosis method using linear discriminant analysis on ECG signals," *Measurement*, vol. 42, no. 5, pp. 778–789, 2009.
[19] C. Ye, B. V. Kumar, and M. T. Coimbra, "Heartbeat classification using morphological and dynamic features of ECG signals," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2930–2941, Oct. 2012.
[20] S. Karimifard and A. Ahmadian, "A robust method for diagnosis of morphological arrhythmias based on Hermitian model of higher-order statistics," *Biomed. Eng. Online*, vol. 10, no. 1, 2011, Art. no. 22.
[21] R. J. Martis *et al.*, "Application of principal component analysis to ECG signals for automated diagnosis of cardiac health," *Expert Syst. Appl.*, vol. 39, no. 14, pp. 11 792–11 800, 2012.
[22] R. J. Martis *et al.*, "Cardiac decision making using higher order spectra," *Biomed. Signal Process. Control*, vol. 8, no. 2, pp. 193–203, 2013.
[23] F. A. Elhaj *et al.*, "Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals," *Comput. Methods Programs Biomedicine*, vol. 127, pp. 52–63, 2016.
[24] Y. Kutlu and D. Kuntalp, "A multi-stage automatic arrhythmia recognition and classification system," *Comput. Biol. Med.*, vol. 41, no. 1, pp. 37–45, 2011.
[25] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
[26] A. E. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," *Scientific Data*, vol. 3, 2016, Art. no. 160035.
[27] E. Y. Ding *et al.*, "Novel method of atrial fibrillation case identification and burden estimation using the MIMIC-III electronic health data set," *J. Intensive Care Med.*, vol. 34, no. 10, pp. 851–857, 2019.
[28] J. Lázaro *et al.*, "Wearable armband device for daily life electrocardiogram monitoring," *IEEE Trans. Biomed. Eng.*, to be published, doi: 10.1109/TBME.2020.3004730.
[29] G. Moody *et al.*, "Predicting the onset of paroxysmal atrial fibrillation: The computers in cardiology challenge 2001," in *Proc. Computers Cardiol.*, 2001, vol. 28, pp. 113–116.
[30] S. K. Bashar *et al.*, "Noise detection in electrocardiogram signals for intensive care unit patients," *IEEE Access*, vol. 7, pp. 88 357–88 368, 2019.
[31] S. K. Bashar *et al.*, "VERB: VFCDM-based electrocardiogram reconstruction and beat detection algorithm," *IEEE Access*, vol. 7, pp. 13 856–13 866, 2019.
[32] D. Han *et al.*, "Smartwatch PPG peak detection method for sinus rhythm and cardiac arrhythmia," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Medicine Biol. Soc.*, 2019, pp. 4310–4313.
[33] S. Dhahbi, W. Barhoumi, and E. Zagrouba, "Breast cancer diagnosis in digitized mammograms using curvelet moments," *Comput. Biol. Med.*, vol. 64, pp. 79–90, 2015.
[34] M. Schmidt *et al.*, "ECG derived respiration: Comparison of time-domain approaches and application to altered breathing patterns of patients with schizophrenia," *Physiological Meas.*, vol. 38, no. 4, pp. 601–615, 2017.
[35] H. Esponda *et al.*, "A setting-free differential protection for power transformers based on second central moment," *IEEE Trans. Power Del.*, vol. 34, no. 2, pp. 750–759, Apr. 2019.
[36] S. Kaneko, Y. Satoh, and S. Igarashi, "Using selective correlation coefficient for robust image registration," *Pattern Recognit.*, vol. 36, no. 5, pp. 1165–1173, 2003.
[37] M. Tong, Y. Pan, Z. Li, and W. Lin, "Valid data based normalized cross-correlation (vdncc) for topography identification," *Neurocomputing*, vol. 308, pp. 184–193, 2018.
[38] A. Mahmood and S. Khan, "Correlation-coefficient-based fast template matching through partial elimination," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2099–2108, Apr. 2011.
[39] S.-K. Hwang and W.-Y. Kim, "A novel approach to the fast computation of zernike moments," *Pattern Recognit.*, vol. 39, no. 11, pp. 2065–2076, 2006.
[40] S. Li, M.-C. Lee, and C.-M. Pun, "Complex zernike moments features for shape-based image retrieval," *IEEE Trans. Syst., Man, Cybern.-Part A: Syst. Humans*, vol. 39, no. 1, pp. 227–237, 2008.
[41] C. Singh, N. Mittal, and E. Walia, "Face recognition using Zernike and complex Zernike moment features," *Pattern Recognit. Image Anal.*, vol. 21, no. 1, pp. 71–81, 2011.
[42] A. Tahmasbi, F. Saki, and S. B. Shokouhi, "Classification of benign and malignant masses based on zernike moments," *Comput. Biol. Medicine*, vol. 41, no. 8, pp. 726–735, 2011.
[43] D. R. Nayak, R. Dash, and B. Majhi, "Brain MR image classification using two-dimensional discrete wavelet transform and adaboost with random forests," *Neurocomputing*, vol. 177, pp. 188–197, 2016.
[44] A. Singh *et al.*, "Image processing based automatic diagnosis of glaucoma using wavelet features of segmented optic disc from fundus image," *Comput. Methods Programs Biomed.*, vol. 124, pp. 108–120, 2016.
[45] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design," *IEEE Trans. Signal Process.*, vol. 40, no. 9, pp. 2207–2232, Sep. 1992.
[46] C.-L. Chang and B. Girod, "Direction-adaptive discrete wavelet transform for image compression," *IEEE Trans. Image Process.*, vol. 16, no. 5, pp. 1289–1302, May 2007.
[47] T. T. Nguyen, X. Dai Pham, and J. W. Jeon, "An improvement of the standard hough transform to detect line segments," in *Proc. IEEE Int. Conf. Ind. Technol.*, 2008, pp. 1–6.
[48] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, no. 1, pp. 11–15, 1972.
[49] Z. Xu, B.-S. Shin, and R. Klette, "Accurate and robust line segment extraction using minimum entropy with hough transform," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 813–822, Mar. 2014.
[50] S. Maji and J. Malik, "Object detection using a Max-Margin Hough transform," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit.*, 2009, pp. 1038–1045.
[51] D. Duan *et al.*, "An improved hough transform for line detection," in *Proc. Int. Conf. Comput. Appl. Syst. Model.*, 2010, vol. 2, pp. V2–354.
[52] "Mathworks - makers of matlab and simulink," 2019. [Online]. Available: https://www.mathworks.com/, Accessed: Feb. 17, 2019.
[53] G. Roffo *et al.*, "Infinite latent feature selection: A probabilistic latent graph-based ranking approach," in *Proc. IEEE Int. Conf. Comput. Vision*, 2017, pp. 1398–1406.
[54] M. R. Homaeinezhad *et al.*, "ECG arrhythmia recognition via a neuro-SVM–KNN hybrid classifier with virtual qrs image-based geometrical features," *Expert Syst. Appl.*, vol. 39, no. 2, pp. 2047–2058, 2012.
[55] S. S. Mullick, S. Datta, and S. Das, "Adaptive learning-based k-nearest neighbor classifiers with resilience to class imbalance," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5713–5725, Nov. 2018.
[56] S.-N. Yu and K.-T. Chou, "Selection of significant independent components for ECG beat classification," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2088–2096, 2009.
[57] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Germany: Springer, 2006.
[58] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[59] A. Elola *et al.*, "ECG-based pulse detection during cardiac arrest using random forest classifier," *Med. Biol. Eng. Comput.*, vol. 57, no. 2, pp. 453–462, 2019.
[60] A. Porta *et al.*, "Measuring regularity by means of a corrected conditional entropy in sympathetic outflow," *Biol. Cybern.*, vol. 78, no. 1, pp. 71–78, 1998.
[61] L. Mainardi *et al.*, "Linear and non-linear analysis of atrial signals and local activation period series during atrial-fibrillation episodes," *Med. Biol. Eng. Comput.*, vol. 39, no. 2, pp. 249–254, 2001.
[62] J. L. Wells Jr, R. B. Karp, N. T. Kouchoukos, W. A. Maclean, T. N. James, and A. L. Waldo, "Characterization of atrial fibrillation in man: Studies following open heart surgery," *Pacing Clin. Electrophysiol.*, vol. 1, no. 4, pp. 426–438, 1978.

Authorized licensed use limited to: UNIVERSITY OF CONNECTICUT. Downloaded on October 18,2021 at 19:30:32 UTC from IEEE Xplore. Restrictions apply.