Computers in Biology and Medicine 133 (2021) 104367

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: http://www.elsevier.com/locate/compbiomed

Prediction of paroxysmal atrial fibrillation using new heart rate variability features

Ashkan Parsi *, Martin Glavin, Edward Jones, Dallan Byrne

National University of Ireland (NUI) Galway, Galway, H91 TK33, Ireland

ARTICLE INFO

Keywords:
Biomedical signal processing
Heart rate variability (HRV)
Paroxysmal atrial fibrillation (PAF)
Machine learning

ABSTRACT

Paroxysmal atrial fibrillation (PAF) is a cardiac arrhythmia that can eventually lead to heart failure or stroke if left untreated. Early detection of PAF is therefore crucial to prevent any further complications and avoid fatalities. An implantable defibrillator device could be used to both detect and treat the condition though such devices have limited computational capability. With this constraint in mind, this paper presents a novel set of features to accurately predict the presence of PAF. The method is evaluated using ECG signals from the widely used atrial fibrillation prediction database (AFPDB) from PhysioNet. We analysed 106 signals from 53 pairs of ECG recordings. Each pair of signals contains one 5-min ECG segment that ends just before the onset of a PAF event and another 5-min ECG segment at least 45 min distant from the PAF event, to represent a non-PAF event. Seven novel features are extracted through the Poincaré representation of R-R interval signals, and are prioritised through feature ranking schemes. The features are used with four standard classification techniques for PAF prediction and compared to the existing state of the art from the literature. Using only the seven proposed features, classification performance outperforms those of the classical state-of-the-art feature set, registering sensitivity and specificity measurements of over 96%. The results further improve when the features are combined with several of the classical features, with an accuracy increasing to 98% using a linear kernel SVM. The results show that the proposed features provide a useful representation of the PAF condition and achieve good prediction with off-the-shelf classification techniques that would be suitable for ICU deployment.

1. Introduction

Atrial fibrillation (AF) is one of the most common cardiac arrhythmias affecting adults of any age [1]. AF occurs when the heart beats in a disorganized and irregular way, and if persistent and left untreated, can lead to various heart-related complications [2]. Although it is not immediately life-threatening in the same way as some other arrhythmias, AF can lead to heart failure or stroke and thromboembolic events which increase overall mortality [3]. One in four people over 50 are at risk of AF which may severely impact on the quality of their life [4]. Approximately 0.7 million (13%) of the ≈5.3 million cases of atrial fibrillation in the United States are undiagnosed [5], and approximately 15% of stroke patients present with AF, a figure that is projected to double by 2030 [5].

Atrial Fibrillation can be divided into three categories: paroxysmal atrial fibrillation (PAF), persistent atrial fibrillation and chronic atrial fibrillation. PAF presents as short duration episodes of AF, lasting from several minutes to days and is self-terminating. Persistent AF occurs similarly to PAF, but it cannot self-terminate without external treatment such as medication or electrical shock. Finally, chronic atrial fibrillation has the most significant effect on the body, lasting more than 7 days, and can prevent the heart rhythm returning to normal behaviour even with treatment. AF patients often start with episodes of PAF before their condition escalates to a chronic stage. Furthermore, about 18% of PAF evolves to permanent AF (persistent or chronic) over 4 years [6].

AF can be treated by medication or electrical shock issued by the implantable defibrillator device (ICD) [7]. Therefore, having an accurate predictor of the onset of PAF is clinically important because it increases the possibility to prevent the onset of atrial arrhythmias either electrically or using pharmacological treatments, and can enable more efficient and cost-effective screening protocols [8]. An accurate predictor would allow for a time-efficient and cost-effective screening procedure during a clinical visit which may decrease the risk of strokes and thromboembolic events [9].

* Corresponding author.
E-mail addresses: ashkan.parsi@nuigalway.ie (A. Parsi), martin.glavin@nuigalway.ie (M. Glavin), edward.jones@nuigalway.ie (E. Jones), dallan.byrne@nuigalway.ie (D. Byrne).

https://doi.org/10.1016/j.compbiomed.2021.104367
Received 10 November 2020; Received in revised form 9 March 2021; Accepted 29 March 2021
Available online 2 April 2021

Table 1
Performance comparison between the previous works using PAC analysis on AFPDB database.

Previous Works	Method, Classifier and Signal Length	VM	SN (%)	SP (%)
Zong et al., 2001 [11]	Number and timing of PACs, 30 min segment	Single hold	79	–
Hickey et al., 2002 [12]	PACs analysis and spectral based HRV features along with proposed classier, 30 min segment	5-fold CV	79	72
Thong et al., 2004 [13]	PAC analysis with proposed three criteria classification method, 30 min segment	Single hold	89	91
Erdenebayar et al., 2019 [14]	CNN, 30 s segment	Single hold	98.7	98.6
Jalali et *al.*, 2020 [15]	PACs analysis and resampling with CNN, 30 min segment	3-fold CV	99.6	99.4
Bashar et al., 2021 [27]	PACs detection using SVM and RF along with 9 selected features	Single hold	–	100

CNN = convolutional neural network; VM = validation method; SN = sensitivity; SP = specificity; CV = cross validation.

For more than four decades, researchers have attempted to predict PAF using electrocardiogram (ECG) signals [10]. PAF prediction work can be categorised into:

a) Premature atrial complex (PAC) detection [11–15];
b) Heart rate variability (HRV) analysis [12,16–24].

PAC detection assumes approximately 93% of PAF have been initiated by PAC [25]. The majority of PAC detection methods in the literature analyse a 30-min ECG signal. The highest results, 99.6% in sensitivity and 99.4%, was obtained using a CNN by Jalali et al. [15]. PAC-based methods based on data obtained from the well-known PhysioNet atrial fibrillation prediction database (AFPDB) [26] are summarized in Table 1. Furthermore, Bashar et al. [27] has used the AFPDB database to test their proposed algorithm only specificity where they have reached 100% in specificity while only using selected number of patients from the database. They have obtained 124 segments of PACs from 13 patients, to evaluate proposed PAC detection system using tuned support vector machine and random forest along with 9 selected features. In more recent works, it has been shown better performance (as much as 10%) could be achieved using deep learning algorithm such as long short-term memory (LSTM) and recurrent neural network (RNN) as compared to machine learning classifiers, such as support vectors, logistic regression, etc while ECG signals have been process in the proposed PAC detection methods [28–30].

Other approaches use heart rate variability (HRV) analysis to directly detect a PAF event. Previous works have used different HRV metrics such as time, frequency, bispectrum and nonlinear feature extracted from 5, 10, 15 and/or even 30 min HRV signal to predict PAF as shown in Table 2. In contrast to PAC analysis, these methods directly detect a PAF episode, require less computational power and can be implemented on an implantable device like ICD or pacemaker. As an ICD is one of the most common methods to restore the normal rhythm, it could be used effectively in PAF patients, especially with the recent advances in their power and battery life [7,21,31]. A PAF predictor could control the ICD's antitachycardia pacing treatment method possibly allowing it to immediately restore normal sinus rhythm once an arrhythmia is detected. Generally, the ICD device is expected to operate for more than 5 years after it is implanted in the human body [32]. To maintain the device lifespan, on-board computation must be minimised, which, in the context of a PAF prediction method means there must be a reduction in the:

a) length of the signals analysed (reduce storage)
b) the number of features calculated
c) the complexity of the detection method

In order to contribute to the goal of reduced computational load of onset of PAF detection in this context, this study presents a set of novel HRV features for the prediction of a PAF event. The paper describes the extraction of seven novel features that, when tested using four well established classification methods, yield accuracies of over 96.7%, almost 4% greater than the classical state-of-the-art feature set. These features are extracted from a Poincaré representation of 5-min segments of R-R interval signals extracted from patient ECG data. The features are compared to the state-of-the-art features within the literature and analysed in terms of importance using feature ranking schemes. A 10-fold cross-validation method is used to demonstrate the PAF prediction performances with different features sets. The novel features are tested by themselves, and when combined with existing features from previous research.

The rest of the paper is organized as follows. Section II describes the proposed feature set. The experimental procedure is outlined in Section III. Results are presented and discussed in Section IV while conclusions

Table 2
Performance comparison between the previous works using HRV analysis on AFPDB database.

Previous Works	Method, Classifier and Signal Length	VM	SN (%)	SP (%)
Lynn et al., 2001 [16]	HRV based return and difference map proposed features with k-NN classifier, 30 min segment	Single hold	57	–
Yang et al., 2001 [17]	HRV based proposed method: Footprint analysis, 10 min segment	Single hold	64	–
Hickey et al., 2002 [12]	Spectral based HRV features along with proposed classier, 10 min segment	5-fold CV	53	80
Chesnokov et al., 2008 [18]	HRV based spectral and complexity analysis with neural network classifier, 30 min segment	Single hold	72.7	88.2
Mohebbi et al., 2012 [19]	Time, frequency and nonlinear HRV features (12 metrics) with SVM classifier, 30 min segment	Single hold	96.3	93.1
Costin et al., 2013 [20]	HRV features and morphological variability of QRS with proposed classifier, 30 min segment	Single hold	89.3	89.4
Boon et al., 2016 [21]	Bispectrum and nonlinear HRV features (9 selected metrics) with SVM classifier, 15 min segment	10-fold CV	77.4	81.1
Boon et al., 2018 [23]	Time, frequency, bispectrum and nonlinear HRV features (7 selected metrics) with SVM classifier, 5 min segment	10-fold CV	86.8	88.7
Ebrahimzadeh et al., 2018 [24]	Time, frequency and nonlinear HRV features (12 selected metrics) using mixture of experts classifiers, 5 min segment	10-fold CV	100	95.5

k-NN = k-nearest neighbour; VM = validation method; SN = sensitivity; SP = specificity; SVM = support vector machine; CV = cross validation.

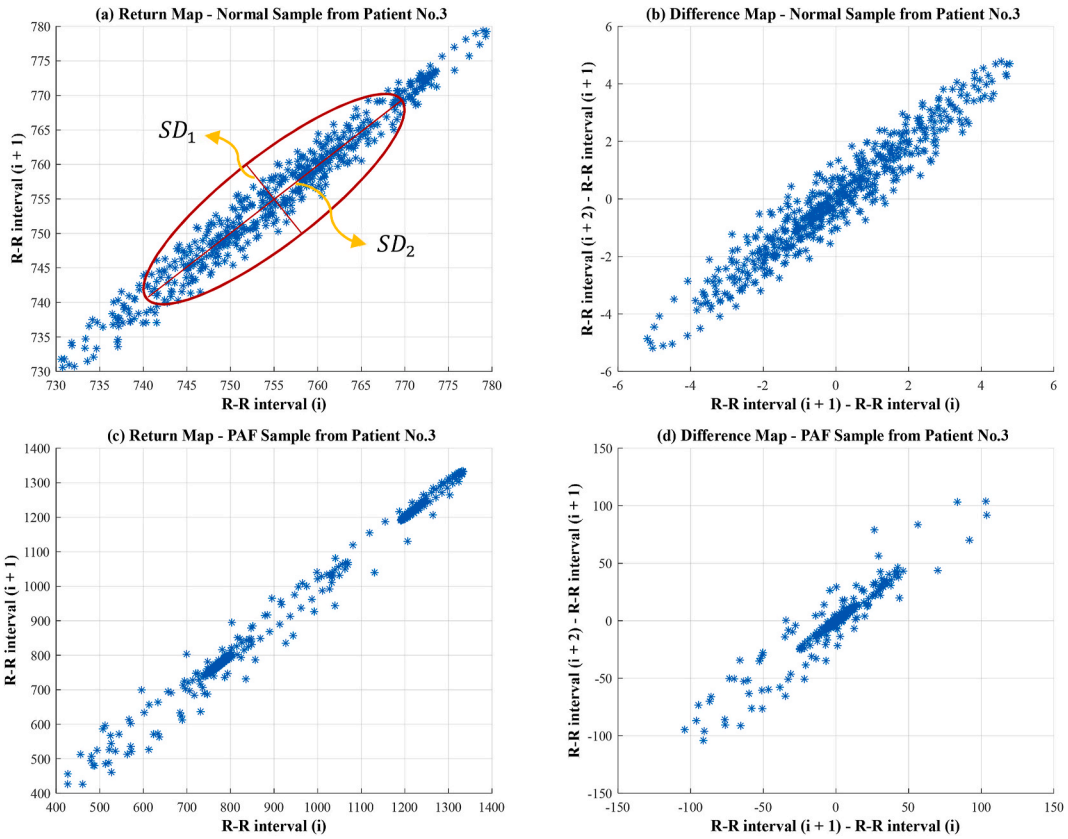

Fig. 1. The return map (left) and the difference map (right) of R-R intervals for normal and PAF ECG signals from the AFPDB (Patient No. 3.) SD_1 and SD_2 represents the dispersion along minor and major axis of the fitted ellipse.

are detailed in the final section.

2. Proposed heart rate variability metrics

The instantaneous heart rate is calculated from the R-R intervals detected from a recorded ECG signal. Considering each heartbeat has a maximum value, an R peak, R-R intervals are defined as a the time between instantaneous heartbeats, also known as the normal-to-normal intervals [33]. There are many classic HRV measures in time, frequency, bispectrum and nonlinear domain which can capture both the sympathetic and the parasympathetic components of the autonomic nervous system [34,35] and have been used for arrythmia detection and prediction including PAF [19,21,22,36].

a) Poincaré Plot Representation

The Poincaré plot is a visual representation of R-R intervals [16, 37–40] and is constructed as follows. Let us denote the R-R interval signal by: $RR_1, RR_2, RR_3, \ldots, RR_t$, where RR_i represent each R-R interval in millisecond and t is the number of R-R intervals in the signal under analysis. The return map is a plot of $(RR_i, RR_{i+1}),\ i \in \{1, 2, \ldots, t-1\}$ i.e., a plot of the points: $(RR_1,RR_2),(RR_2,RR_3),\ldots,(RR_{t-1},RR_t)$. Two standard deviations can be derived, namely SD_1 and SD_2, which, if the return map is considered as an ellipse, capture the minor and major semi-axes (Fig. 1 (a)). SD_1 and SD_2 are defined as:

$$SD_1 = \sqrt{\frac{1}{2}SDSD^2} \tag{1}$$

$$SD_2 = \sqrt{2SDRR^2 - \frac{1}{2}SDSD^2} \tag{2}$$

where $SDSD$ is the standard deviation (SD) of differences between adjacent R-R intervals, and $SDRR$ is the standard deviation of all R-R intervals in the whole signal:

$$SDSD = SD(RR_i - RR_{i+1}),\ i \in \{1, 2, \ldots, t-1\} \tag{3}$$

$$SDRR = SD(RR_i),\ i \in \{1, 2, \ldots, t\} \tag{4}$$

The width (SD_1) of this ellipse, which is related to the fast beat-to-beat variability in the HRV, and the length (SD_2) of the ellipse related to the longer-term variability of that data [41] are not the only features from return map that have been used before. Recent works also have proposed different angle of points or even number of clusters in the Poincaré plot as effective features in PAF prediction applications [38, 42].

b) Difference Map

Another Poincaré-based representation for PAF prediction is the difference map of R-R intervals. Difference maps are constructed from the difference between consecutive R-R intervals $(RR_{n+1} - RR_n, RR_{n+2} - RR_{n+1})$. Fig. 1 (b) and (d) shows the difference map plot of a healthy and onset of PAF R-R interval signals, respectively. If the return map is considered to capture the velocity of R-R intervals, the difference map is a visual representation of the rate-of-change of the velocity, i.e., the acceleration.

In both Fig. 1 (a) and (b) which present return and difference map of normal cardiac signal respectively, changes are more gradual. In contrast, in Fig. 1 (c) and (d), changes are more sudden and chaotic showing disordered cardiac activity which may lead to atrial fibrillation. By showing the magnitude of velocity in the return map, we can study changes in the heart rate which is a relevant factor to monitor in the event of cardiac arrhythmias such as atrial fibrillation. Furthermore, the

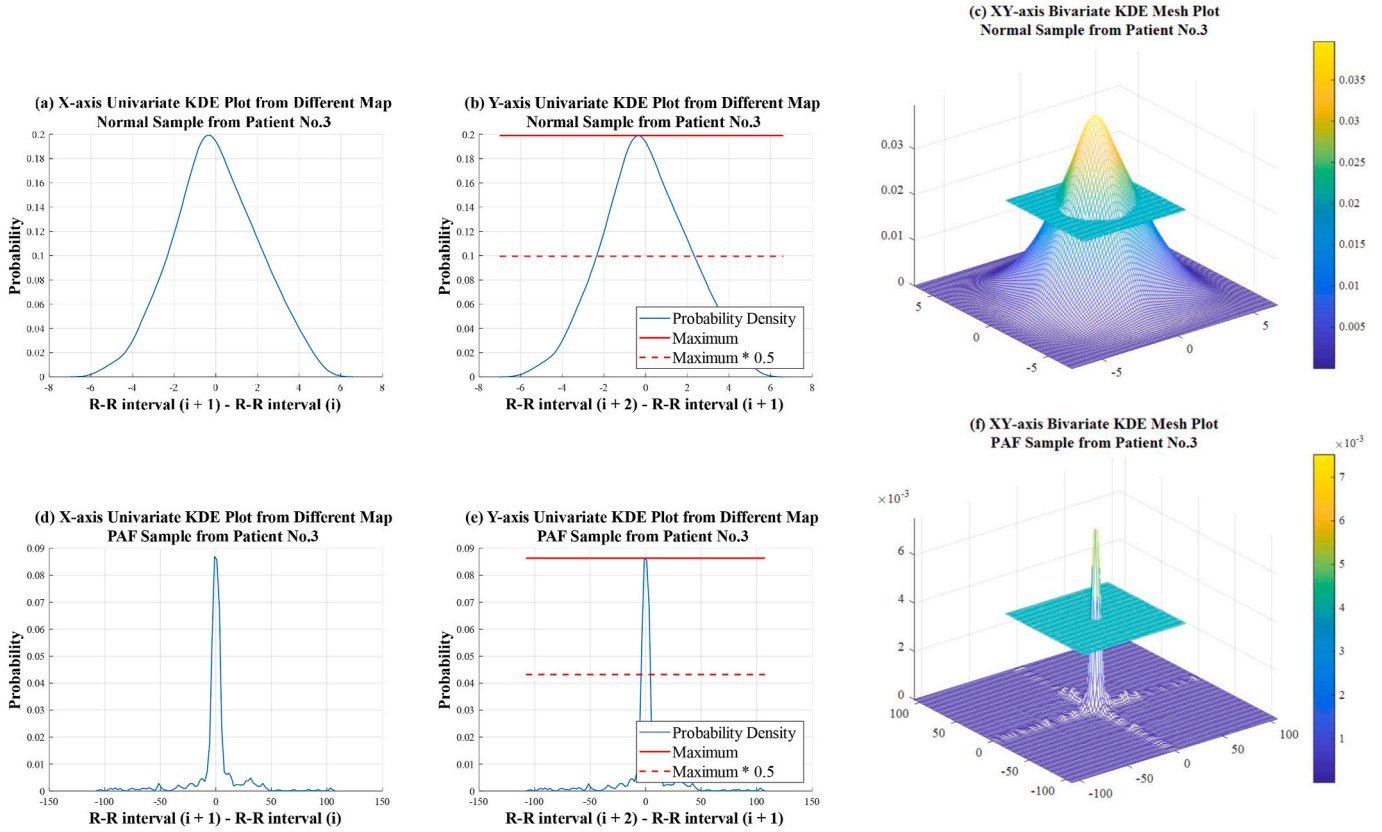

Fig. 2. The difference map probability density of one axis (left and centre columns) and the difference map KDE mesh plot (right) of the subject No. 3. The top row shows the normal event and bottom shows the onset of PAF for this patient.

difference map allows us to monitor the acceleration. This could help to monitor anomalies in the sinus node that control the heartbeat rate.

The first feature proposed here is the covariance of X and Y axes of the difference map, calculated as follows:

$$Cov(X,Y)=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu_x)(Y_i-\mu_y) \tag{5}$$

where $X_i = RR_{i+1} - RR_i$ and $Y_i = RR_{i+2} - RR_{i+1}$ and μ represent the mean of each axis. Slope and angle differences between normal and PAF difference maps which show the changes in heartbeat acceleration can be represented using the change in covariance value over time.

c) Kernel Density Estimation

A bivariate kernel density estimation (KDE) can also be applied to difference map. KDE is a nonparametric density estimator in statistical data analysis [43]. To calculate the final bivariate KDE of difference map, first consider a one-dimensional distribution function of each axis, where $RR_{n+1} - RR_n$ is the x-axis and $RR_{n+2} - RR_{n+1}$ is the y-axis. Taking the x-axis first, let $X_i = \{x_1, x_2, \ldots, x_n\} \in \mathbb{R}^d$ be a random sample from distribution with unknown univariate probability density $f(x)$. The standard kernel density estimator for $f(x)$ is calculated as follows [44]:

$$\widehat{f}(x)=\frac{1}{nh}\sum_{i=1}^{n}k\left(\frac{x-X_i}{h}\right) \tag{6}$$

where n is the number of observations, h is positive number called the smoothing parameter ($h \to 0$ with $nh \to \infty$ as $n \to \infty$) and $k(x)$ is the kernel function satisfying the following conditions:

$$0 \le k(x) < \infty \text{ for all } x, \text{ and } \int_{-\infty}^{\infty} k(x)dx = 1 \tag{7}$$

where the most common kernel function is a Gaussian:

$$k(x,\ X_i)=\left(2\pi\sigma^2\right)^{-d/2}exp\left\{-\frac{x-X_i^2}{2\sigma^2}\right\} \tag{8}$$

After calculating KDE for each dimension, the univariate form can be simply extended to bivariate or even multivariate form. In the most popular form, the standard bivariate kernel density estimator is written as follows [45]:

$$\widehat{f}(x,y)=\frac{1}{nh_xh_y}\sum_{i=1}^{n}k\left(\frac{x-X_i}{h_x}\right)k\left(\frac{x-X_i}{h_y}\right) \tag{9}$$

Fig. 2 shows both univariate and bivariate KDE from difference map for both normal ECG, and ECG at onset of PAF event for patient No. 3. There are some similarities between the univariate KDEs for x- and y-axis for each condition (normal and PAF), as can be seen Fig. 2. There are also visible differences in the bivariate plots Fig. 2 (c) and Fig 2 (f). Six novel features have been calculated to better highlight these differences between normal and PAF distributions.

1) Univariate KDE Features

There are two features extracted from one axis of a univariate KDE. It is sufficient to calculate these from one axis, given the similarity between the axes, which differ by one shift value.

The first feature is the area between the peak and half peak values of the univariate KDE, indicated by the solid red line and dashed red line respectively on Fig. 2 (b) and (e). Then the area between is calculated as

Table 3
Proposed features extracted from difference map.

Features Category	Features Name	Descriptions
Difference Map Covariance (1 Metrics)	$Cov(X,Y)$	The covariance of X and Y axes of the difference map.
Univariate KDE Features (2 Metrics)	$Area(y)$ $Energy(y)$	The area and energy of the peak to half peak in univariate KDE.
Bivariate KDE Features (4 Metrics)	$SurfMax$ $SurfMin$	the minimum and maximum spans of the bottom plan (XY) in bivariate KDE
	$Volume(x,y)$ $Energy(x,\ y)$	The volume and energy of the area between that half-max plane and the peak value in bivariate KDE.

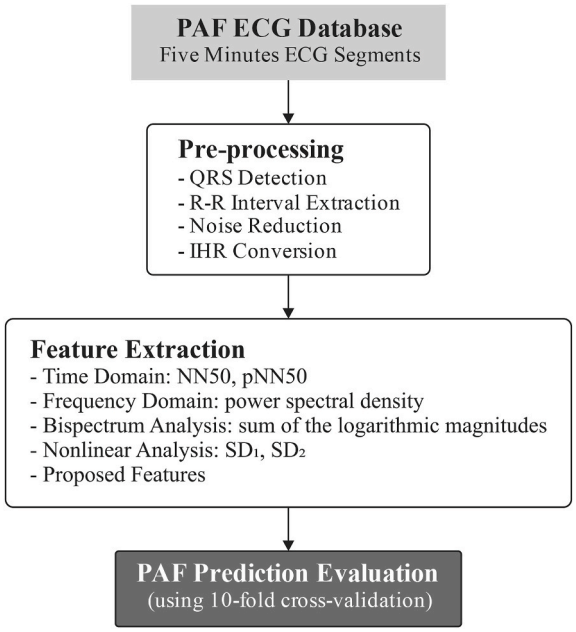

Fig. 3. PAF prediction proposed algorithm block diagram.

follows:

$$Area(y) = \sum_i \widehat{f}(y)_i,\ where\ \widehat{f}(y)_i > \max\left(\widehat{f}(y)\right)*0.5 \tag{10}$$

where $\widehat{f}(y)$ is the univariate KDE function of y-axis and i counts all probability values which are higher than the half of the $\widehat{f}(y)$ maximum probability.

As a second feature, the energy of the peak to half peak of the KDE is calculated as follows:

$$Energy(y) = \sqrt{\sum_i \widehat{f}(y)_i^2},\ where\ \widehat{f}(y)_i > \max\left(\widehat{f}(y)\right)*0.5 \tag{11}$$

2) Bivariate KDE Features

Four metrics are extracted from the bivariate KDE. The first two metrics are the minimum and maximum spans of the bottom plan (XY) in the 3-D KDE plot which we call *SurfMax* and *SurfMin* respectively. As an example, in Fig. 2 (c) *SurfMax* is about 5.5 and *SurfMin* is −5.5.

The other two features are similar to the area and energy features univariate derived from the univariate distribution. In this case, we define a half-maximum plane in the distribution, and calculate the volume and energy of the area between that plane and the peak value:

$$Volume(x,y) = \sum_i \widehat{f}(x,\ y)_i \tag{12}$$

$$Energy(x,\ y) = \sqrt{\sum_i \widehat{f}(x,\ y)_i^2} \tag{13}$$

where $\widehat{f}(x,\ y)_i > \max(\widehat{f}(x,\ y))*0.5$. The seven proposed features are summarized in Table 3.

3. Experimental methodology

This Section discusses the use of the measures proposed in Section II to predict PAF based on 5-min segments of R-R interval signals derived from ECG patient data. An overview of the entire method is given in Fig. 3. The data are first pre-processed to extract R-R information and reduce noise. Features are then extracted and finally the system determines whether the patient data is normal or abnormal using a range of classification techniques. The data are obtained from the PhysioNet atrial fibrillation database [26,46].

a) PAF ECG Database

The ECG data that have been used for this study were taken from the PhysioNet AFPDB database [26]. This is an annotated database and consists of 3 types of record sets. The first record set, starting with the letter "n", comes from 50 subjects who did not experience atrial fibrillation at all. This set is usually used as "normal" ECG signal for tuning detectors [11,13,16,20]. The second record set was taken from 25 subjects staring with letter "p". Each subject in this set has 30 min of ECG signal (odd number) during a period that is distant from PAF, labelled as a normal ECG signal, and 30 min of ECG signal (even number) immediately precedes a PAF episode. The third record set contains 100 annotated 30 min ECG signal recordings from 50 subjects. In this record set, there are subjects with signals that are all normal, in PAF onset, or one in each category.

Based on previous works [12,18–21,23,24], the selection of subjects from the database can vary. Chesnokov [18] only extracted HRV signal from 16 patients. The selection criteria was the availability of at least 60 min of signal length before PAF event to investigate the possibility of long-term prediction of PAF onset. Mohebbi et al. [19], extracted 106 ECG segments immediately prior to an episode of PAF to evaluate their proposed algorithm. In common with the approach take in Refs. [21,24], in this research 106 events from 53 patients have been processed. Each patient contributes a pair of signals which consists of one 30-min ECG segment that ends just prior to the onset of a PAF event, and another 30-min ECG segment at least 45 min distant from any PAF event, and therefore represents normal heart behaviour. With a number of classifiers, 10-fold cross-validation has been used for evaluation. Each ECG segment contains two-channel traces from Holter recordings with sampling rate of 128 Hz and 12-bit resolution. In this study we consider a 5-min ECG segment based on proposed signals length in previous studies [23,24]. The 5-min ECG segment occurring at least 45 min from the PAF event is assigned a class label of "Normal", while the ECG segment that immediately precedes the PAF event is given a class label of "Abnormal". This method has been chosen as the algorithm is designed for use in implantable devices where the decision is based on each patient's

Table 4
Features from previous studies along with their descriptions [33,64].

Features Category	Features Name	Descriptions
Time Domain (4 Metrics)	MeanNN	Mean of normal-to-normal intervals or instantaneous heartbeats
	SDNN	Standard deviation
	NN50	Number of adjacent R-R intervals differing by more than 50 ms
	pNN50	The number obtained by dividing NN50 by the total number of NN intervals
Frequency Domain (4 Metric)	PSD	Power spectral density (PSD) in LL, LH, HH and ROI
Bispectrum Analysis (15 Metrics)	M_{avg}	Magnitude average of bispectrum in LL, LH, HH and ROI: $M_{avg} = \frac{1}{N}\sum_{\Gamma}\lvert B(f_1,f_2)\rvert$
	P_{avg}	Power average of bispectrum in LL, LH, HH and ROI: $P_{avg} = \frac{1}{N}\sum_{\Gamma}\lvert B(f_1,f_2)\rvert^2$
	L_m	Sum of the logarithmic magnitude of the bispectrum in LL, LH, HH and ROI: $L_m = \sum_{\Gamma}\log(\lvert B(f_1,f_2)\rvert)$
	L_{dm}	Sum of the logarithmic magnitudes of the diagonal elements of the bispectrum in LL, HH and ROI: $L_{dm} = \sum_{D}\log(\lvert B(f_1,f_2)\rvert)$
Nonlinear Analysis (6 Metrics)	SD_1	The width (SD_1) and the length (SD_2) of the of generated ellipse in Poincaré plot (Fig. 1)
	SD_2	
	SD_1/SD_2	
	Entropy	Sample entropy, Rényi entropy, and Tsallis entropy

healthy and PAF onset signals.

b) Pre-processing and Noise Reduction

HRV signals are extracted from the ECG signals using the annotation files provided with the database [26,46]. The annotation files are unaudited and contain some errors in QRS detection, which are generally unavoidable in such data collection [26]. The presence of these errors presents a more challenging scenario for PAF prediction and will further test the robustness of the proposed algorithm.

After using the annotations to find the R peaks, the R-R interval signals have been calculated and then resampled to 7 Hz by using the cubic spline interpolation. Then the instantaneous heart rate (*IHR*) signal is calculated from the R-R interval signals using the following equation:

$$IHR(beats/min) = \frac{60,000}{RRI\ (msec)} \tag{14}$$

where *RRI* (R-R intervals) is the time in milliseconds between instantaneous heartbeats.

A two-step noise reduction method has been implemented to remove unwanted signal noise caused by e.g. muscular activity [47]. In the first step, to deal with spikes which present as noise or as ectopic beats, the signal has been corrected by McNames's algorithm [48]. To apply the algorithm, the following statistic is first calculated on the *IHR* signals:

$$D(n) = \frac{|IHR(n) - IHR_m|}{1.483 med\{|IHR(n) - IHR_m|\}} \tag{15}$$

where $med\{\ldots\}$ is the median operator, $IHR(n)$ is instantaneous heart rate for beat n, IHR_m is the median value of the heart rate over a given window. The filtered instantaneous heart rate, $\widehat{IHR}(n)$, is then calculated as a follow:

$$\widehat{IHR}(n) = \begin{cases} IHR(n) & D(n) < \tau \\ med\left\{IHR(n+m) : |m| < \frac{w_m - 1}{2}\right\} & D(n) \geq \tau \end{cases} \tag{16}$$

where τ is application specific threshold value. If $D(n)$ exceeds this threshold, the instantaneous heart rate is considered as abnormal, otherwise it considered as a normal value. In the abnormal case the heart rate is corrected as shown in equation (16), where w_m is the window length of the median filter. In this work, τ and w_m are set to 4 and 11 respectively as suggested in Ref. [48]. The second step of the noise reduction scheme is based on the wavelet transform (WT) adapted to remove any artefact beats, similar to Refs. [33,49,50]. Using high and low pass filters, the WT can remove the higher frequencies of the background noise from the signals. In this research, the first approximation level of WT is taken, and the sym8 wavelet is used. The pre-processed signal is now ready for feature extraction.

c) Feature Extraction

In addition to the seven proposed features described in Section II, other features are also extracted from the pre-processed signals. These features have been used in previous PAF prediction studies [19–24]. The features have been selected from four categories: time domain, frequency domain, bispectrum and nonlinear analysis and they have been summarized in Table 4 [33]. Using the formula in equation (14), the pre-processed signal can be converted to *IHR* or R-R interval signals for different feature extraction techniques. Four widely-used time-domain features are extracted from HRV analysis [51]. The time-domain features used are MeanNN, SDNN, NN50, and pNN50 [50–57]. Time domain methods have low computational cost but cannot easily discriminate between the sympathetic (at low-frequency range (0.04–0.15 Hz)) and the parasympathetic contributions (at higher frequency range (0.15–0.4 Hz)) in the HRV signals [58]. This distinction facilitates preventive intervention at an early stage when it is most beneficial [57]. In the frequency domain category, four features based on average power of HRV signals across the very low frequency (VLF) band (0.0–0.04 Hz), low frequency (LF) band (0.04–0.15 Hz), high frequency (HF) band (0.15–0.4 Hz), and ratios like LF/HF have been computed by integrating the power spectral density (PSD) estimated using Welch's method in each frequency range of interest [19,23,56,59–61].

$B(f_1, f_2)$ = bispectrum of HRV signal; LF = low frequency band (0.04–0.15 Hz); HF = high frequency band (0.15–0.4 Hz); LL = low frequency sub band region (LF-LF); HH = high frequency sub band region (HF–HF).

Higher order spectral (HOS) features capture any quadratic phase-coupled harmonics which can be present due to non-linearities within the HRV signal. HOS reveals phase relations between the frequency components which cannot be observed through the regular power spectrum [57,62]. HOS features up to the third-order cumulant were used to estimate the bispectrum from HRV data in different HRV studies [19,21,23,36,56]. Based on [63] the bispectrum, as a symmetric function, has twelve symmetric regions. Within the region of interest (ROI) there are three sub-bands which can highlight the sympathetic and parasympathetic attributes of a HRV signal [36]. The magnitude average (M_{avg}), power average (P_{avg}), and logarithmic bispectrum features, are the most commonly used features from the bispectrum which have been extracted from the ROI; they are summarized in Table 4 [33,64].

By considering a HRV signal as an indirect representation of permanent interplay between the two branches of the ANS, the signal can be decomposed using nonlinear and dynamic analysis [65,66]. On top of the of the Poincaré plot features.

Table 5
The Lasso, mRMR, and ILFS features ranking and selection processing on 5 min signals from AFPDB database.

Features Category	Features Name	Lasso	mRMR	ILFS
Time Domain (4 Metrics)	MeanNN	0	0	0
	SDNN	1	0	0
	NN50	1	4	14
	pNN50	1	2	2
Frequency Domain (4 Metric)	PSD in VLF	0	0	0
	PSD in LF	1	0	0
	PSD in HF	0	0	0
	PSD in LF/HF	1	1	3
Bispectrum Analysis (15 Metrics)	M_{avg} in LL	0	0	0
	M_{avg} in LH	1	0	0
	M_{avg} in HH	0	0	0
	M_{avg} in ROI	0	0	0
	P_{avg} in LL	0	0	0
	P_{avg} in LH	0	0	0
	P_{avg} in HH	0	0	0
	P_{avg} in ROI	0	0	0
	L_m in LL	1	0	0
	L_m in LH	1	0	0
	L_m in HH	1	0	0
	L_m in ROI	1	0	0
	L_{dm} in LL	1	12	4
	L_{dm} in HH	1	14	11
	L_{dm} in ROI	1	0	0
Nonlinear Analysis (6 Metrics)	SD_1	1	7	5
	SD_2	1	13	12
	SD_1/SD_2	0	0	0
	SampEn	0	0	0
	Rényi Entropy	0	0	0
	Tsallis Entropy	0	0	0
Difference Map Covariance (1 Metrics)	$Cov(X, Y)$	1	11	6
Univariate KDE Features (2 Metrics)	$Area(y)$	1	6	8
	$Energy(y)$	1	3	1
Difference Map Covariance (4 Metrics)	*SurfMin*	1	9	6
	SurfMax	1	8	7
	$Volume(x,y)$	1	10	10
	$Energy(x, y)$	1	5	13

SD_1, SD_2, and SD_1/SD_2 from HRV data entropy features, as a measure of signal complexity are also used. Sample entropy [67], Rényi entropy, Tsallis entropy [68] have been extracted to highlight the complexity of HRV and aim to capture any regularity changes during the onset of a VT-VF event. Adding these 29 features from previous studies to the 7 proposed features makes an initial feature set consisting of thirty-six features. Feature ranking and analysis is then used in order to prioritise features; this is described in more detail in Section IV.

Table 6
Summaries of experimental results using different feature set on AFPDB database using 10-fold cross-validation method.

Features Category	Classifier	SN (%)	SP (%)	ACC (%)
Classic Features	SVM	97.8	88.9	93.3
	MLP	95.6	91.1	93.3
	RF	**98.8**	87.8	93.3
	k-NN	91.1	82.2	86.7
Proposed Features	**SVM**	96.7	**96.7**	**96.7**
	MLP	90	96.7	93.3
	RF	93.3	96.7	95
	k-NN	90	83.3	86.7
Combined	SVM	**98.8**	**96.7**	**97.7**
	MLP	96.7	97.8	97.2
	RF	97.8	91.1	95
	k-NN	87.8	95.6	91.7

SN = sensitivity; SP = specificity; ACC = accuracy.

d) PAF Prediction

In this work, four different classifiers have been applied to predict onset of PAF episodes by classifying each sample in one of two classes, either normal or abnormal, based on extracted features presented as a feature vector. A k-fold cross-validation technique is used to split the data evenly and record performance metrics for $k = 10$ folds in the 53-patient database (106 Normal and Abnormal events). 10 folds equates to having 48 patients in the training set and 5 patients in the test set. The same patient is never present in both the training and test set.

The four classifiers are: support vector machine (SVM) [19,21,23], k-nearest neighbours (k-NN) [16,33,64], random forest (RF) [69] and multilayer perceptron (MLP) [18,24]. SVM has proven to be effective in previous studies on HRV classification [40,70]. The HRV pre-processing, feature extraction, optimization, and classification have been implemented and developed in MATLAB R2018b, and feature analysis and selection used the MATLAB Statistics Toolbox along with the feature selection toolbox in Ref. [71]. For the SVM classifier, using Gaussian function as a kernel, we obtained the same results. In this paper we adopt a linear SVM to reduce the computational load and reduce the risk of overfitting. The kernel scale was set to 1 and box constraint set to 1.6 [64]. Optimization of k-NN was carried out using different numbers of neighbours (1–50), along with different distance metrics such as Chebyshev distance, cosine distance, Minkowski distance along with Euclidean distance. For k-NN the best number of neighbours was 5 and the Chebyshev distance has been applied in this study [16,33,50]. The

Table 7
Comparison of experimental results in the leading literature using different classifiers on 5-minue of AFPDB ECG signals using 10-fold cross-validation method.

Features Category	Classifier	SN (%)	SP (%)	ACC (%)
Boon et al., 2018 [23] using time, frequency, bispectrum and nonlinear HRV features (7 selected metrics)	SVM	86.8	88.7	87.7
Ebrahimzadeh et al., 2018 [24] using time, frequency and nonlinear HRV features (12 selected metrics)	SVM	96.3	93.1	94.6
	k-NN	92.3	86.7	89.3
	MLP	92.6	89.7	91.1
	ME	100	95.5	98.2
Proposed Method using time, frequency, bispectrum and nonlinear HRV features (7 metrics) along with 7 proposed metrics	SVM	98.8	96.7	97.7
	k-NN	87.8	95.6	91.7
	MLP	96.7	97.8	97.2
	RF	97.8	91.1	95

SN = sensitivity; SP = specificity; ACC = accuracy.

method of random forests (RF) comprises an ensemble of decision trees, and has previously been used to distinguish various cardiac arrhythmias [50,72,73]. RF uses 64 trees with leaf size equal to 4 in this study [64, 69]. Finally, the MLP has 5 hidden layers with 32, 32, 16, 16 and 4 nodes on each layer, respectively. The maximum epochs are set to 4096 using mean squared error with regularization function, and implemented in MATLAB R2018b.

4. Experimental results

In this section we test seven proposed HRV features along with a subset of the state-of-the-art features from previous studies to evaluate PAF prediction using 5-min R-R interval signals.

a) Results of Feature Selection and Ranking

Feature selection and ranking is used to firstly, make the model easier to interpret by removing the variables that are redundant; secondly, to reduce the size of the problem which enables algorithms to work faster; and thirdly, to guard against model overfitting. A range of methods for feature selection were evaluated in this study, with results shown in Table 5. LASSO (least absolute shrinkage and selection operator) was first formulated by Tibshirani [74]. It uses an L1 norm and tends to force individual coefficient values completely towards zero. It performs two main tasks: regularization and feature selection. LASSO puts a constraint on the sum of the absolute values of the model parameters where the sum has to be less than a fixed value which is called the upper bound. To do that Lasso applies a shrinking (regularization) process where it penalizes the coefficients of the regression variables, shrinking some of them to zero. The variables that still have a non-zero coefficient after the shrinking process are selected. As indicated in Table 5, twenty-two features have been selected with LASSO where all the newly proposed features are included. LASSO feature analysis and selection helps to deal with multicollinearity and redundant predictors by quickly identifying the key variables [75]. This set of features provides a useful baseline for further prioritization.

To further analyse the original set of features, two other feature selection and ranking methods are used. The minimal redundancy-maximal Relevance (mRMR) and the infinite latent feature selection (ILFS) are used for this. In the mRMR feature ranking method, the aim is to maximize the mutual information between the feature distribution and classes and, at the same time, minimize the redundancy between features [76]. The ILFS feature ranking method, on the other hand, is a robust probabilistic latent graph-based method which performs the feature ranking using all possible feature subsets [77]. Table 5 summarises the results of ranking of the top 14 features using both approaches. These 14 features selected by mRMR and ILFS are a subset of the features identified by LASSO. It was found that incrementally adding more features does not improve the cross-validation result. Based on the results of feature selection and ranking, at least five out of the seven proposed features are included in the top 10 rankings using both ranking methods. Both methods include one of the proposed features at the highest rank.

b) PAF Prediction: Results of Feature Comparison

Results from different studies for PAF prediction have mainly been presented in terms of sensitivity (true positive rate), specificity (true negative rate), and accuracy:

$$Sensitivity = \frac{TP}{TP + FN} \tag{17}$$

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{19}$$

where TP, TN, FP and FN stand for true positive, true negative, false positive and false negative respectively. The results for different feature sets are presented in Table 6. The proposed features by themselves yield higher specificity and accuracy results than the classic feature set, though the classical features result in better sensitivity. By combining the proposed and state-of-the-art feature sets, a linear SVM achieves 98.8% in sensitivity and 96.7% in specificity. The combination of proposed and classical features combines the best performance features of each set of features, i.e., better overall sensitivity in the classical features mixed with high specificity provided by the proposed features.

c) Discussion and Comparison with Previous Work

The highest result using HRV analysis on the AFPDB, 100% in sensitivity and 95.5% specificity, was obtained in Ref. [24]. Although the sensitivity reported in that study cannot be improved upon, better specificity has been achieved in the present study while both methods, our study here and proposed method in Ref. [24] resulted in the same overall accuracy (97.7%). However, the study in Ref. [24] is based on mixture of experts (ME) classifier that is a much more complex approach in comparison with the linear SVM that has been used on this study. Other classifiers were also used in Ref. [24] as shown in Table 7. Compared to the results reported in Ref. [24], the SVM-based method in this research has shown over 3% improvement in overall accuracy, k-NN over 2% improvement, and finally MLP over 6% improvement. These results suggest that the proposed features have potential for application in implantable device, where computational resources are limited. In particular, the proposed features improve upon specificity compared to previous research.

5. Conclusion

Predicting the onset of PAF can dramatically improve quality of life

in cardiac patients and decrease the risk of mortality. Accurate prediction remains a significant challenge, considering the noise and interference present in recorded ECG measurements, as well as the computational limitations of wearable and implantable devices increasingly used for cardiac rhythm management. In this study, seven novel HRV features have been presented to address the PAF prediction problem. The features are shown to represent the problem space well compared to a selection of popular state-of-the-art features from the literature, using two feature ranking methods. Four classification algorithms were used to compare the classic and proposed feature sets for PAF prediction, using data from the PhysioNet AFPDB database. Test and training data were split in a 10-fold cross-validation manner and both sets remained distinct. Using a linear SVM kernel, the proposed feature set provides an improvement of over 3% in accuracy over the leading published results in literature using the same classifier (97.7% vs 94.6%). The proposed method could be used to trigger antitachycardia pacing (ATP) which is a less aggressive therapy that controls heartbeats on ICDs or other implanted devices, with more invasive electrotherapy reserved only for cases that absolutely require it such as ventricular tachycardia and ventricular fibrillation [78–80].

In the context of ATP, the occurrence of false positives is not associated with a mortality [81]. On the other hand, ATP that is only applied on detection even on false negative has no effect on mortality [78, 82–84]. However, there is a trade-off between early prediction and the buffering overhead available on a computationally limited ICD. The computational complexity has not been considered in the present study, however, examining the combinations of features along with computational requirements is planned as a future study. Investigating the KDE resolution could be the key to implement proposed method on ICDs as computational complexity of KDE decreases dramatically by lowering the resolution or using estimation methods for less complex calculation of KDE [85,86]. Furthermore, as the maps are almost Gaussian in nature, future works will investigate using Gaussian approximations instead of KDE [87] to classify between Normal and PAF events on verity of atrial fibrillation databases.

Acknowledgment

This work was supported by the Irish Research Council (Grant number GOIPG/2016/1604) and Enterprise Ireland Marie Skłodowska-Curie Career-FIT Fellowship (Grant number MF-2018-0202).

References

[1] Y. Iwasaki, K. Nishida, T. Kato, S. Nattel, Atrial fibrillation pathophysiology, Circulation 124 (2011) 2264–2274, https://doi.org/10.1161/CIRCULATIONAHA.111.019893.

[2] A. Brandes, M.D. Smit, B.O. Nguyen, M. Rienstra, I.C. Van Gelder, Risk factor management in atrial fibrillation, Arrhythmia Electrophysiol. Rev. 7 (2018) 118, https://doi.org/10.15420/aer.2018.18.2.

[3] A. Alonso, F.L. Norby, Predicting atrial fibrillation and its complications, Circ. J. 80 (2016) 1061–1066, https://doi.org/10.1253/circj.CJ-16-0239.

[4] Irish heart atrial fibrillation - Irish heart, n.d. https://irishheart.ie/heart-and-stroke-conditions-a-z/atrial-fibrillation/. accessed February 5, 2020

[5] E.J. Benjamin, P. Muntner, A. Alonso, M.S. Bittencourt, C.W. Callaway, A. P. Carson, A.M. Chamberlain, A.R. Chang, S. Cheng, S.R. Das, F.N. Delling, L. Djousse, M.S.V. Elkind, J.F. Ferguson, M. Fornage, L.C. Jordan, S.S. Khan, B. M. Kissela, K.L. Knutson, T.W. Kwan, D.T. Lackland, T.T. Lewis, J.H. Lichtman, C. T. Longenecker, M.S. Loop, P.L. Lutsey, S.S. Martin, K. Matsushita, A.E. Moran, M. E. Mussolino, M. O'Flaherty, A. Pandey, A.M. Perak, W.D. Rosamond, G.A. Roth, U. K.A. Sampson, G.M. Satou, E.B. Schroeder, S.H. Shah, N.L. Spartano, A. Stokes, D. L. Tirschwell, C.W. Tsao, M.P. Turakhia, L.B. VanWagner, J.T. Wilkins, S.S. Wong, S.S. Virani, Heart disease and stroke statistics—2019 update: a report from the American heart association. https://doi.org/10.1161/CIR.0000000000000659, 2019.

[6] S.M. Al-Khatib, W.E. Wilkinson, L.L. Sanders, E.A. McCarthy, E.L.C. Pritchett, Observations on the transition from intermittent to permanent atrial fibrillation, Am. Heart J. 140 (2000) 142–145, https://doi.org/10.1067/mhj.2000.107547.

[7] E.N. Prystowsky, Management of atrial fibrillation: therapeutic options and clinical decisions, Am. J. Cardiol. 85 (2000) 3–11, https://doi.org/10.1016/S0002-9149(00)00908-5.

[8] A. Prakash, S. Saksena, M. Hill, PhD, R.B. Krol, A.N. Munsif, I. Giorgberidze, P. Mathew, R. Mehra, Acute effects of dual-site right atrial pacing in patients with spontaneous and inducible atrial flutter and fibrillation, J. Am. Coll. Cardiol. 29 (1997) 1007–1014, https://doi.org/10.1016/S0735-1097(97)00043-0.

[9] Y.V. Chesnokov, A.V. Holden, H. Zhang, Screening patients with paroxysmal atrial fibrillation (PAF) from non-PAF heart rhythm using HRV data analysis, in: 2007 Comput. Cardiol, IEEE, 2007, pp. 459–462, https://doi.org/10.1109/CIC.2007.4745521.

[10] S. Poli, V. Barbaro, P. Bartolini, G. Calcagnini, F. Censi, Prediction of atrial fibrillation from surface ECG: review of methods and algorithms, Ann. Ist. Super Sanita 39 (2003) 195–203. http://www.ncbi.nlm.nih.gov/pubmed/14587218.

[11] W. Zong, R. Mukkamala, R.G. Mark, A methodology for predicting paroxysmal atrial fibrillation based on ECG arrhythmia feature analysis, in: Comput. Cardiol vol. 28, IEEE, 2001, pp. 125–128, https://doi.org/10.1109/CIC.2001.977607. No.01CH37287.

[12] B. Hickey, C. Heneghan, Screening for paroxysmal atrial fibrillation using atrial premature contractions and spectral measures, in: Comput. Cardiol, IEEE, 2002, pp. 217–220, https://doi.org/10.1109/CIC.2002.1166746.

[13] T. Thong, J. McNames, M. Aboy, B. Goldstein, Prediction of paroxysmal atrial fibrillation by analysis of atrial premature complexes, IEEE Trans. Biomed. Eng. 51 (2004) 561–569, https://doi.org/10.1109/TBME.2003.821030.

[14] U. Erdenebayar, H. Kim, J.-U. Park, D. Kang, K.-J. Lee, Automatic prediction of atrial fibrillation based on convolutional neural network using a short-term normal electrocardiogram signal, J. Kor. Med. Sci. 34 (2019) 1–10, https://doi.org/10.3346/jkms.2019.34.e64.

[15] A. Jalali, M. Lee, Atrial fibrillation prediction with residual network using sensitivity and orthogonality constraints, IEEE J. Biomed. Heal. Informatics. 24 (2020) 407–413, https://doi.org/10.1109/JBHI.2019.2957809.

[16] K.S. Lynn, H.D. Chiang, A two-stage solution algorithm for paroxysmal atrial fibrillation prediction, in: Comput. Cardiol, vol. 28, IEEE, 2001, pp. 405–407, https://doi.org/10.1109/CIC.2001.977678. Cat. No.01CH37287.

[17] A.C.C. Yang, H.W. Yin, Prediction of paroxysmal atrial fibrillation by footprint analysis, in: Comput. Cardiol, vol. 28, IEEE, 2001, pp. 401–404, https://doi.org/10.1109/CIC.2001.977677. Cat. No.01CH37287.

[18] Y.V. Chesnokov, Complexity and spectral analysis of the heart rate variability dynamics for distant prediction of paroxysmal atrial fibrillation with artificial intelligence methods, Artif. Intell. Med. 43 (2008) 151–165, https://doi.org/10.1016/j.artmed.2008.03.009.

[19] M. Mohebbi, H. Ghassemian, Prediction of paroxysmal atrial fibrillation based on non-linear analysis and spectrum and bispectrum features of the heart rate variability signal, Comput. Methods Progr. Biomed. 105 (2012) 40–49, https://doi.org/10.1016/j.cmpb.2010.07.011.

[20] H. Costin, C. Rotariu, A. Pasarica, Atrial fibrillation onset prediction using variability of ECG signals, in: 8th Int. Symp. Adv. Top. Electr. Eng, 2013, https://doi.org/10.1109/ATEE.2013.6563419. ATEE 2013. (2013) 0–3.

[21] K.H. Boon, M. Khalil-Hani, M.B. Malarvili, C.W. Sia, Paroxysmal atrial fibrillation prediction method with shorter HRV sequences, Comput. Methods Progr. Biomed. 134 (2016) 187–196, https://doi.org/10.1016/j.cmpb.2016.07.016.

[22] T. Hurnanen, E. Lehtonen, M.J. Tadi, T. Kuusela, T. Kiviniemi, A. Saraste, T. Vasankari, J. Airaksinen, T. Koivisto, M. Pankaala, Automated detection of atrial fibrillation based on time–frequency analysis of seismocardiograms, IEEE J. Biomed. Heal. Informatics. 21 (2017) 1233–1241, https://doi.org/10.1109/JBHI.2016.2621887.

[23] K.H. Boon, M. Khalil-Hani, M. Malarvili, Paroxysmal atrial fibrillation prediction based on HRV analysis and non-dominated sorting genetic algorithm III, Comput. Methods Progr. Biomed. 153 (2018) 171–184, https://doi.org/10.1016/j.cmpb.2017.10.012.

[24] E. Ebrahimzadeh, M. Kalantari, M. Joulani, R.S. Shahraki, F. Fayaz, F. Ahmadi, Prediction of paroxysmal Atrial Fibrillation: a machine learning based approach using combined feature vector and mixture of expert classification on HRV signal, Comput. Methods Progr. Biomed. 165 (2018) 53–67, https://doi.org/10.1016/j.cmpb.2018.07.014.

[25] C. Kolb, S. Nürnberger, G. Ndrepepa, B. Zrenner, A. Schömig, C. Schmitt, Modes of initiation of paroxysmal atrial fibrillation from analysis of spontaneously occurring episodes using a 12-lead Holter monitoring system, Am. J. Cardiol. 88 (2001) 853–857, https://doi.org/10.1016/S0002-9149(01)01891-4.

[26] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J. E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioToolkit PhysioBank, PhysioNet, Circulation 101 (2000) e215–e220, https://doi.org/10.1161/01.CIR.101.23.e215.

[27] S.K. Bashar, D. Han, F. Zieneddin, E. Ding, T.P. Fitzgibbons, A.J. Walkey, D. D. McManus, B. Javidi, K.H. Chon, Novel density Poincaré plot based machine learning method to detect atrial fibrillation from premature atrial/ventricular contractions, IEEE Trans. Biomed. Eng. 68 (2021) 448–460, https://doi.org/10.1109/TBME.2020.3004310.

[28] S. Liaqat, K. Dashtipour, A. Zahid, K. Assaleh, K. Arshad, N. Ramzan, Detection of atrial fibrillation using a machine learning approach, Information 11 (2020) 549, https://doi.org/10.3390/info11120549.

[29] Y. Baek, S. Lee, W. Choi, D. Kim, Prediction of atrial fibrillation from normal ECG using artificial intelligence in patients with unexplained stroke, Eur. Heart J. 41 (2020) 2020, https://doi.org/10.1093/ehjci/ehaa946.0348.

[30] S. Raghunath, J.M. Pfeifer, A.E. Ulloa-Cerna, A. Nemani, T. Carbonati, L. Jing, D. P. VanMaanen, D.N. Hartzel, J.A. Ruhl, B.F. Lagerman, D.B. Rocha, N.J. Stoudt, G. Schneider, K.W. Johnson, N. Zimmerman, J.B. Leader, H.L. Kirchner, C. J. Griessenauer, A. Hafez, C.W. Good, B.K. Fornwalt, C.M. Haggerty, Deep neural networks can predict new-onset atrial fibrillation from the 12-lead

electrocardiogram and help identify those at risk of AF-related stroke, Circulation 120 (2021), 047829, https://doi.org/10.1161/CIRCULATIONAHA.120.047829.
[31] Kilhwan Kim, Unsun Cho, Yunho Jung, Jaeseok Kim, Design and implementation of biomedical SoC for implantable cardioverter defibrillators, in: 2007 IEEE Asian Solid-State Circuits Conf, IEEE, 2007, pp. 248–251, https://doi.org/10.1109/ASSCC.2007.4425777.
[32] R.X. Stroobandt, S.S. Barold, A.F. Sinnaeve, Implantable Cardioverter-Defibrillators Step by Step, Wiley-Blackwell, Oxford, UK, 2009, https://doi.org/10.1002/9781444303377.
[33] A. Parsi, D. O'Loughlin, M. Glavin, E. Jones, Prediction of sudden cardiac death in implantable cardioverter defibrillators: a review and comparative study of heart rate variability features, IEEE Rev. Biomed. Eng. 13 (2020) 5–16, https://doi.org/10.1109/RBME.2019.2912313.
[34] N. Singh, K.J. Moneghetti, J.W. Christle, D. Hadley, D. Plews, V. Froelicher, Heart rate variability: an old metric with new meaning in the era of using mHealth technologies for health and exercise training guidance. Part One: physiology and methods, Arrhythmia Electrophysiol. Rev. 7 (2018) 193, https://doi.org/10.15420/aer.2018.27.2.
[35] Y. İşler, M. Kuntalp, Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure, Comput. Biol. Med. 37 (2007) 1502–1510, https://doi.org/10.1016/j.compbiomed.2007.01.012.
[36] S.-N. Yu, M.-Y. Lee, Bispectral analysis and genetic algorithm for congestive heart failure recognition based on heart rate variability, Comput. Biol. Med. 42 (2012) 816–825, https://doi.org/10.1016/j.compbiomed.2012.06.005.
[37] L. Bergfeldt, Y. Haga, Power spectral and Poincaré plot characteristics in sinus node dysfunction, J. Appl. Physiol. 94 (2003) 2217–2224, https://doi.org/10.1152/japplphysiol.01037.2002.
[38] J. Park, S. Lee, M. Jeon, Atrial fibrillation detection by heart rate variability in Poincare plot, Biomed. Eng. Online 8 (2009) 38, https://doi.org/10.1186/1475-925X-8-38.
[39] M.M. Platiša, T. Bojic, S.U. Pavlovic, N.N. Radovanovic, A. Kalauzi, Generalized Poincaré plots-A new method for evaluation of regimes in cardiac neural control in atrial fibrillation and healthy subjects, Front. Neurosci. 10 (2016) 1–9, https://doi.org/10.3389/fnins.2016.00038.
[40] J.P. Sepulveda-Suescun, J. Murillo-Escobar, R.D. Urda-Benitez, D.A. Orrego-Metaute, A. Orozco-Duque, Atrial fibrillation detection through heart rate variability using a machine learning approach and Poincare plot features, in: I. Torres, J. Bustamante, D.A. Sierra (Eds.), IFMBE Proc., Springer Singapore, Singapore, 2017, pp. 565–568, https://doi.org/10.1007/978-981-10-4086-3_142.
[41] P.W. Kamen, H. Krum, A.M. Tonkin, Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans, Clin. Sci. 91 (1996) 201–208, https://doi.org/10.1042/cs0910201.
[42] S. Moharreri, N. Jafarnia Dabanloo, S. Rezaei, S. Parvaneh, New feature set for better representation of dynamic of RR intervals in poincare plot, in: Comput. Cardiol, 2010, pp. 1–4, https://doi.org/10.22489/CinC.2017.321-395, 2017.
[43] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, second ed., Springer, New York, New York, NY, 2009 https://doi.org/10.1007/978-0-387-84858-7.
[44] B.W. Silverman, Density Estimation for Statistics and Data Analysis, Routledge, 2018, https://doi.org/10.1201/9781315140919.
[45] D.W. Scott, Multivariate Density Estimation, Wiley, 2015, https://doi.org/10.1002/9781118575574.
[46] G. Moody, A. Goldberger, S. McClennen, S. Swiryn, Predicting the onset of paroxysmal atrial fibrillation: the Computers in Cardiology Challenge 2001, in: Comput. Cardiol. 2001, vol. 28, IEEE, 2001, pp. 113–116, https://doi.org/10.1109/CIC.2001.977604. Cat. No.01CH37287.
[47] W. Irnich, Muscle noise and interference behavior in pacemakers: a comparative study, Pacing Clin. Electrophysiol. 10 (1987) 125–132, https://doi.org/10.1111/j.1540-8159.1987.tb05932.x.
[48] J. McNames, T. Thong, M. Aboy, Impulse rejection filter for artifact removal in spectral analysis of biomedical signals, in: 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, 2004, pp. 145–148, https://doi.org/10.1109/IEMBS.2004.1403112.
[49] J. Zhuang, X. Ning, S. Du, Z. Wang, C. Huo, X. Yang, A. Fan, Nonlinear short-term heart rate variability prediction of spontaneous ventricular tachyarrhythmia, Sci. Bull. 53 (2008) 2446–2453, https://doi.org/10.1007/s11434-008-0345-y.
[50] A. Parsi, D. Byrne, M. Glavin, E. Jones, Heart rate variability feature selection method for automated prediction of sudden cardiac death, Biomed. Signal Process Contr. 65 (2021) 102310, https://doi.org/10.1016/j.bspc.2020.102310.
[51] M. Malik, A.J. Camm, J.T. Bigger, G. Breithardt, S. Cerutti, R.J. Cohen, P. Coumel, E.L. Fallen, H.L. Kennedy, R.E. Kleiger, F. Lombardi, A. Malliani, A.J. Moss, J. N. Rottman, G. Schmidt, P.J. Schwartz, D.H. Singer, A.J. Camm, R.E. Kleiger, A. Malliani, A.J. Moss, P.J. Schwartz, Heart rate variability: standards of measurement, physiological interpretation, and clinical use, Eur. Heart J. 17 (1996) 354–381, https://doi.org/10.1093/oxfordjournals.eurheartj.a014868.
[52] P. Melillo, N. De Luca, M. Bracale, L. Pecchia, Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability, IEEE J. Biomed. Heal. Informatics. 17 (2013) 727–733, https://doi.org/10.1109/JBHI.2013.2244902.
[53] S. Joo, K. Choi, S. Huh, Prediction of ventricular tachycardia by a neural network using parameters of heart rate variability, in: Comput. Cardiol. Conf. (CinC), IEEE, 2010, pp. 585–588. https://ieeexplore.ieee.org/document/5738040/.
[54] S. Joo, S.-J. Huh, K.-J. Choi, A predictor for ventricular tachycardia based on heart rate variability analysis, in: 2011 IEEE Biomed. Circuits Syst. Conf. BioCAS 2011, 2011, pp. 409–411, https://doi.org/10.1109/BioCAS.2011.6107814.
[55] S. Joo, K.-J. Choi, S.-J. Huh, Prediction of spontaneous ventricular tachyarrhythmia by an artificial neural network using parameters gleaned from short-term heart rate variability, Expert Syst. Appl. 39 (2012) 3862–3866, https://doi.org/10.1016/j.eswa.2011.09.097.
[56] K.H. Boon, M. Bala Krishnan, M. Khalil-Hani, Ventricular tachyarrhythmia prediction based on heart rate variability and genetic algorithm, TELKOMNIKA (Telecommunication Comput. Electron. Control. 14 (2016) 999–1008, https://doi.org/10.12928/telkomnika.v14i3.3665.
[57] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C.M. Lim, J.S. Suri, Heart rate variability: a review, Med. Biol. Eng. Comput. 44 (2006) 1031–1051, https://doi.org/10.1007/s11517-006-0119-0.
[58] L. Chen, C. Song, X. Zhang, Statistical modeling of electrocardiography signal for subject monitoring and diagnosis, in: Healthc. Anal. From Data to Knowl. To Healthc. Improv., John Wiley & Sons, Inc, Hoboken, New Jersey, 2016, pp. 95–126, https://doi.org/10.1002/9781118919408.ch4.
[59] H. Lee, M. Seo, S. Joo, Early prediction of ventricular tachyarrhythmias based on heart rate variability analysis, in: 2015 Comput. Cardiol. Conf, IEEE, 2015, pp. 1041–1044, https://doi.org/10.1109/CIC.2015.7411092.
[60] C.G. Wollmann, R. Gradaus, D. Böcker, T. Fetsch, F. Hintringer, G. Hoh, R. Hatala, A. Podczeck-Schweighofer, U. Kreutzer, P. Kamaryt, T. Hauser, J.F. Kersten, K. Wegscheider, G. Breithardt, Variations of heart rate variability parameters prior to the onset of ventricular tachyarrhythmia and sinus tachycardia in ICD patients. Results from the heart rate variability analysis with automated ICDs (HAWAI) registry, Physiol. Meas. 36 (2015) 1047–1061, https://doi.org/10.1088/0967-3334/36/5/1047.
[61] S. Bilgin, O.H. Çolak, O. Polat, E. Koklukaya, Estimation and evaluation of sub-bands on LF and HF base-bands in HRV for ventricular tachyarrhythmia patients, Expert Syst. Appl. 36 (2009) 10078–10084, https://doi.org/10.1016/j.eswa.2009.01.014.
[62] I. Pinhas, E. Toledo, D. Aravot, S. Akselrod, Bicoherence analysis of new cardiovascular spectral components observed in heart-transplant patients: statistical approach for bicoherence thresholding, IEEE Trans. Biomed. Eng. 51 (2004) 1774–1783, https://doi.org/10.1109/TBME.2004.831519.
[63] C.L. Nikias, M.R. Raghuveer, Bispectrum estimation: a digital signal processing framework, Proc. IEEE 75 (1987) 869–891, https://doi.org/10.1109/PROC.1987.13824.
[64] A. Parsi, D. O'Loughlin, M. Glavin, E. Jones, Heart rate variability analysis to predict onset of ventricular tachyarrhythmias in implantable cardioverter defibrillators*, in: 2019 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, 2019, pp. 6770–6775, https://doi.org/10.1109/EMBC.2019.8857911.
[65] J. Piskorski, P. Guzik, Filtering Poincaré plots, Comput. Methods Sci. Technol. 11 (2005) 39–48, https://doi.org/10.12921/cmst.2005.11.01.39-48.
[66] M.P. Tulppo, T.H. Mäkikallio, T.E. Takala, T. Seppänen, H. V Huikuri, Quantitative beat-to-beat analysis of heart rate dynamics during exercise, Am. J. Physiol. 271 (1996) H244–H252. http://ajpheart.physiology.org/content/ajpheart/271/1/H244.full.pdf. accessed September 24, 2017.
[67] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. Cell Physiol. 278 (2000) H2039–H2049, https://doi.org/10.1152/ajpheart.2000.278.6.H2039.
[68] H. Fujita, U.R. Acharya, V.K. Sudarshan, D.N. Ghista, S.V. Sree, L.W.J. Eugene, J.E. W. Koh, Sudden cardiac death (SCD) prediction based on nonlinear heart rate variability features and SCD index, Appl. Soft Comput. 43 (2016) 510–519, https://doi.org/10.1016/j.asoc.2016.02.049.
[69] Z. Mei, X. Gu, H. Chen, W. Chen, Automatic atrial fibrillation detection based on heart rate variability and spectral features, IEEE Access 6 (2018) 53566–53575, https://doi.org/10.1109/ACCESS.2018.2871220.
[70] A. Kampouraki, G. Manis, C. Nikou, Heartbeat time series classification with support vector machines, IEEE Trans. Inf. Technol. Biomed. 13 (2009) 512–518, https://doi.org/10.1109/TITB.2008.2003323.
[71] G. Roffo, S. Melzi, Ranking to learn:, in: A. Appice, M. Ceci, C. Loglisci, E. Masciari, Z.W. Raś (Eds.), New Front. Min. Complex Patterns, Springer International Publishing, Cham, 2017, pp. 19–35, https://doi.org/10.1007/978-3-319-61461-8_2.
[72] A. Batra, V. Jawa, Classification of arrhythmia using conjunction of machine learning algorithms and ECG diagnostic criteria, Int. J. Biol. Biomed. 1 (2016) 1–7. http://www.iaras.org/iaras/filedownloads/ijbb/2016/021-0001.pdf. accessed April 4, 2017.
[73] C. González, E.W. Jensen, P.L. Gambús, M. Vallverdú, Poincaré plot analysis of cerebral blood flow signals: feature extraction and classification methods for apnea detection, PloS One 13 (2018), e0208642, https://doi.org/10.1371/journal.pone.0208642.
[74] R. Tibshirani, Regression shrinkage and selection via the lasso, J. R. Stat. Soc. Ser. B. 58 (1996) 267–288, https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.
[75] R. Tibshirani, Regression shrinkage and selection via the lasso: a retrospective, J. R. Stat. Soc. Ser. B (Statistical Methodol. 73 (2011) 273–282, https://doi.org/10.1111/j.1467-9868.2011.00771.x.
[76] Hanchuan Peng, Fuhui Long, C. Ding, Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2005) 1226–1238, https://doi.org/10.1109/TPAMI.2005.159.
[77] G. Roffo, S. Melzi, U. Castellani, A. Vinciarelli, Infinite latent feature selection: a probabilistic latent graph-based ranking approach, in: 2017 IEEE Int. Conf. Comput. Vis., IEEE, 2017, pp. 1407–1415, https://doi.org/10.1109/ICCV.2017.156.
[78] M.S. Wathen, M.O. Sweeney, P.J. DeGroot, A.J. Stark, J.L. Koehler, M.B. Chisner, C. Machado, W.O. Adkisson, Shock reduction using antitachycardia pacing for

spontaneous rapid ventricular tachycardia in patients with coronary artery disease, Circulation 104 (2001) 796–801, https://doi.org/10.1161/hc3101.093906.
[79] M.S. Wathen, P.J. DeGroot, M.O. Sweeney, A.J. Stark, M.F. Otterness, W. O. Adkisson, R.C. Canby, K. Khalighi, C. Machado, D.S. Rubenstein, K.J. Volosin, Prospective randomized multicenter trial of empirical antitachycardia pacing versus shocks for spontaneous rapid ventricular tachycardia in patients with implantable cardioverter-defibrillators, Circulation 110 (2004) 2591–2596, https://doi.org/10.1161/01.CIR.0000145610.64014.E4.
[80] F.W. Horlbeck, J.O. Schwab, Programming implantable cardioverter/defibrillators and outcomes, F1000Prime Rep. 7 (2015) 1–7, https://doi.org/10.12703/P7-10.
[81] J.P. Daubert, W. Zareba, D.S. Cannom, S. McNitt, S.Z. Rosero, P. Wang, C. Schuger, J.S. Steinberg, S.L. Higgins, D.J. Wilber, H. Klein, M.L. Andrews, W.J. Hall, A. J. Moss, Inappropriate implantable cardioverter-defibrillator shocks in MADIT II, J. Am. Coll. Cardiol. 51 (2008) 1357–1365, https://doi.org/10.1016/j.jacc.2007.09.073.
[82] C. Kouakam, B. Lauwerier, D. Klug, M. Jarwe, C. Marquié, D. Lacroix, S. Kacet, Effect of elevated heart rate preceding the onset of ventricular tachycardia on antitachycardia pacing effectiveness in patients with implantable cardioverter defibrillators, Am. J. Cardiol. 92 (2003) 26–32, https://doi.org/10.1016/S0002-9149(03)00459-4.
[83] N. Nasir, A. Pacifico, T.K. Doyle, N.R. Earle, M.L. Hardage, P.D. Henry, Spontaneous ventricular tachycardia treated by antitachycardia pacing, Am. J. Cardiol. 79 (1997) 820–822, https://doi.org/10.1016/S0002-9149(96)00881-8.
[84] A. Schaumann, F.v. z. Muhlen, B. Herse, B.-D. Gonska, H. Kreuzer, Empirical versus tested antitachycardia pacing in implantable cardioverter Defibrillators : a prospective study including 200 patients, Circulation 97 (1998) 66–74, https://doi.org/10.1161/01.CIR.97.1.66.
[85] V.C. Raykar, R. Duraiswami, L.H. Zhao, Fast computation of kernel estimators, J. Comput. Graph Stat. 19 (2010) 205–220, https://doi.org/10.1198/jcgs.2010.09046.
[86] T.A. O'Brien, K. Kashinath, N.R. Cavanaugh, W.D. Collins, J.P. O'Brien, A fast and objective multidimensional kernel density estimation method: fastKDE, Comput. Stat. Data Anal. 101 (2016) 148–160, https://doi.org/10.1016/j.csda.2016.02.014.
[87] M. Bullmann, T. Fetzer, F. Ebner, F. Deinzer, M. Grzegorzek, Fast kernel density estimation using Gaussian filter approximation, in: 2018 21st Int. Conf. Inf. Fusion, IEEE, 2018, pp. 1233–1240, https://doi.org/10.23919/ICIF.2018.8455686.