

# Convolutional neural networks predict the onset of paroxysmal atrial fibrillation: Theory and applications

Cite as: Chaos 31, 113119 (2021); doi: 10.1063/5.0069272

Submitted: 30 August 2021 · Accepted: 4 October 2021 ·

Published Online: 9 November 2021



View Online



Export Citation



CrossMark

M. Surucu,<sup>1</sup>  Y. Isler,<sup>2,3,a)</sup>  M. Perc,<sup>4,5,6,7</sup>  and R. Kara<sup>1</sup> 

## AFFILIATIONS

<sup>1</sup>Department of Computer Engineering, Duzce University, 81620 Duzce, Turkey

<sup>2</sup>Department of Biomedical Engineering, Izmir Katip Celebi University, Cigli, 35620 Izmir, Turkey

<sup>3</sup>Islerya Medical and Information Technologies Company, Bornova, 35030 Izmir, Turkey

<sup>4</sup>Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia

<sup>5</sup>Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404332, Taiwan

<sup>6</sup>Alma Mater Europaea, Slovenska ulica 17, 2000 Maribor, Slovenia

<sup>7</sup>Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria

<sup>a)</sup> Author to whom correspondence should be addressed: [islerya@yahoo.com](mailto:islerya@yahoo.com)

## ABSTRACT

In this study, we aimed to detect paroxysmal atrial fibrillation episodes before they occur so that patients can take precautions before putting their and others' lives in potentially life-threatening danger. We used the atrial fibrillation prediction database, open data from PhysioNet, and assembled our process based on convolutional neural networks. Conventional heart rate variability features are calculated from time-domain measures, frequency-domain measures using power spectral density estimations, time-frequency-domain measures using wavelet transform, and nonlinear Poincaré plot measures. In addition, we also applied an alternative heart rate normalization, which gave promising results only in a few studies, before calculating these heart rate variability features. We used these features directly and their normalized versions using min–max normalization and z-score normalization methods. Thus, heart rate variability features extracted from six different combinations of these normalizations, in addition to no normalization cases, were applied to the convolutional neural network classifier. We tuned the classifiers' hyperparameters using 90% of feature sets and tested the classifiers' performances using 10% of feature sets. The proposed approach resulted in 87.76% accuracy, 91.30% precision, 80.04% recall, and 87.50% f1-score in heart rate variability with z-score feature normalization. When the heart rate normalization was also utilized, the suggested method gave 100% accuracy, 100% precision, 100% recall, and 100% f1-score in heart rate variability with z-score feature normalization. The proposed method with heart rate normalization and z-score normalization methods resulted in better classification performance than similar studies in the literature. By comparing the existing studies, we conclude that our approach provides a much better tool to determine a near-future paroxysmal atrial fibrillation episode. However, although the achieved benchmarks are impressive, we note that the approach needs to be supported by other studies and on other datasets before clinical trials.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0069272>

Atrial fibrillation is one of the most common types of arrhythmia.<sup>1,2</sup> It can lead to hospitalization, impaired life quality, depression, vascular dementia, heart failure, stroke, and even death.<sup>4</sup> Atrial fibrillation generally starts in a paroxysmal (i.e., auto-terminating) pattern. It progressively gets worse and transforms into a steady condition.<sup>5</sup> Paroxysmal atrial fibrillation starts suddenly with no indication and terminates automatically within a maximum of seven days of onset. During the

episode, the patient may show one or more symptoms, including excessive fatigue, dizziness, fast and irregular heartbeat, palpitations in the chest, shortness of breath, anxiety, and weakness.<sup>4</sup> Thus, if the patient is driving or performing a vital task, life-threatening situations may arise for both themselves and others. Therefore, it is crucial to predict these attacks earlier to enhance the patient's comfort and reduce possible risks.<sup>6</sup> Many machine learning-based approaches have been proposed to detect these

episodes; nonetheless, none has achieved a perfect classifier performance yet. Recently, deep learning algorithms have promised better results in various applications. In this study, we calculated conventional features from 30-min heart rate variability data and utilized two preprocessing methods to data, although deep learning can calculate features from raw data itself. After removing the patient's usual mean heart rate and extracted features' size effects, we observed that the convolutional neural network can detect patients who faced a near-future attack of paroxysmal atrial fibrillation with a perfect accuracy of 100%.

## I. INTRODUCTION

Medical doctors generally rely on analyzing Electrocardiography (ECG) to diagnose several heart diseases first.<sup>7</sup> Since examining ECG graphs is time-consuming, a computer program may help to analyze them. The literature contains many methods to detect and follow heart diseases using digital ECG data.<sup>8,9</sup> In the last half-century, heart rate variability (HRV), derived from ECG, has become popular.<sup>10</sup> HRV analysis has included various (i) time-domain, (ii) frequency-domain, (iii) frequency-time-domain, and (iv) non-linear features.<sup>11–13</sup>

A general pattern recognition study reveals various classification methods after the features were calculated. Many studies, based on disease diagnosis using HRV, have employed different well-known classifier algorithms such as k-nearest neighbors (kNNs), linear discriminant analysis (LDA), decision tree (DT), fuzzy logic (FL), multi-layer perceptron (MLP), support vector machines (SVM), stochastic gradient descent (SGD), radial basis function (RBF), etc.<sup>6,14–16</sup> Similarly, various studies evaluated different classifiers to discriminate paroxysmal atrial fibrillation (PAF) patients from normal subjects.<sup>17–27</sup> Also, several studies evaluated different classifiers to predict recent PAF episodes. For example, Lynn and Chiang<sup>28</sup> achieved 78.00% accuracy using the kNN classifier. Another study resulted in 81.63% accuracy with heart rate and min-max normalizations and 83.67% accuracy with min-max normalized heart rate variability measures using the RBF classifier.<sup>29</sup> Chesnokov reached 82.05% accuracy using the MLP classifier.<sup>30</sup> Boon and colleagues obtained 83.90% accuracy by utilizing genetic algorithms and SVM.<sup>31</sup> The same authors achieved 86.80% using the same configuration in 2019.<sup>32</sup> Schreier *et al.* reached 84.00% accuracy using FL.<sup>18</sup> Three papers use DT as a classifier. Among them, Zong *et al.* reached 88.00% accuracy, Thong *et al.* achieved 89.29%, and Costin *et al.* obtained 89.40% accuracy, interestingly. Maier *et al.* determined 92.00% accuracy using LDA.<sup>17</sup> Recently, Mohebbi and Ghassemian can determine 94.50% accuracy using the SVM classifier. In recent studies, Surucu *et al.*<sup>20,29</sup> highlighted the significance of both feature and heart rate normalizations in HRV-related studies.

Lately, deep learning algorithms have been displayed very famously, notably in image labeling studies.<sup>33,34</sup> After thriving outcomes in image identification, deep learning has also become popular in signal classification studies.<sup>26,35,36</sup> Due to its assuring higher achievement in many classification problems, we decided to utilize the convolutional neural network (CNN) to predict near PAF attacks.

Shortly, this study examines the impacts of both heart rate and feature normalizations in predicting the PAF episodes using CNN from 30-min HRV data. For this purpose, six feature sets evaluated the CNN classifier.

## II. MATERIALS AND METHODS

We summarized our study in Fig. 1 by determining blocks from the HRV-based literature and our previous studies. The dashed line indicates the alternative heart rate normalization method. Feature extraction methods were time-domain parameters, frequency-domain measures (using FFT, Lomb-Scargle, and Welch), time-frequency-domain measures (using a wavelet transform), and Poincaré plot measures. Then, one of the feature normalization methods (no normalization, min-max normalization, and Z-score normalization) was applied to the extracted HRV (or NHRV) features. Finally, a deep learning classifier decides whether the data belong to a PAF patient with the following episode or a PAF patient with no recent attack. Subsections II A–II I cover brief definitions of these blocks.

### A. ECG data

The Atrial Fibrillation Prediction Database (AFPDB, distributed for “The Computers in Cardiology Challenge 2001”) is freely available at the PhysioNet website of <https://www.physionet.org/content/afpdb/1.0.0/> and comprises two channels of 30-min ECG data with a resolution of 16 bits at a sampling frequency of 128 Hz.<sup>37</sup> ECGs are collected from 25 patients with the following PAF episodes and 24 patients with no near PAF episodes. ECG data, obtained from the subject p37, were excluded from the study because it involves many noises.

Although HRV studies have used different data durations in the literature, Seker *et al.*<sup>38</sup> showed the necessity of a minimum of 10 000 heartbeat samples to get reliable results. For this reason, we extracted all these frequently used features for 30-min HRV data from the database.

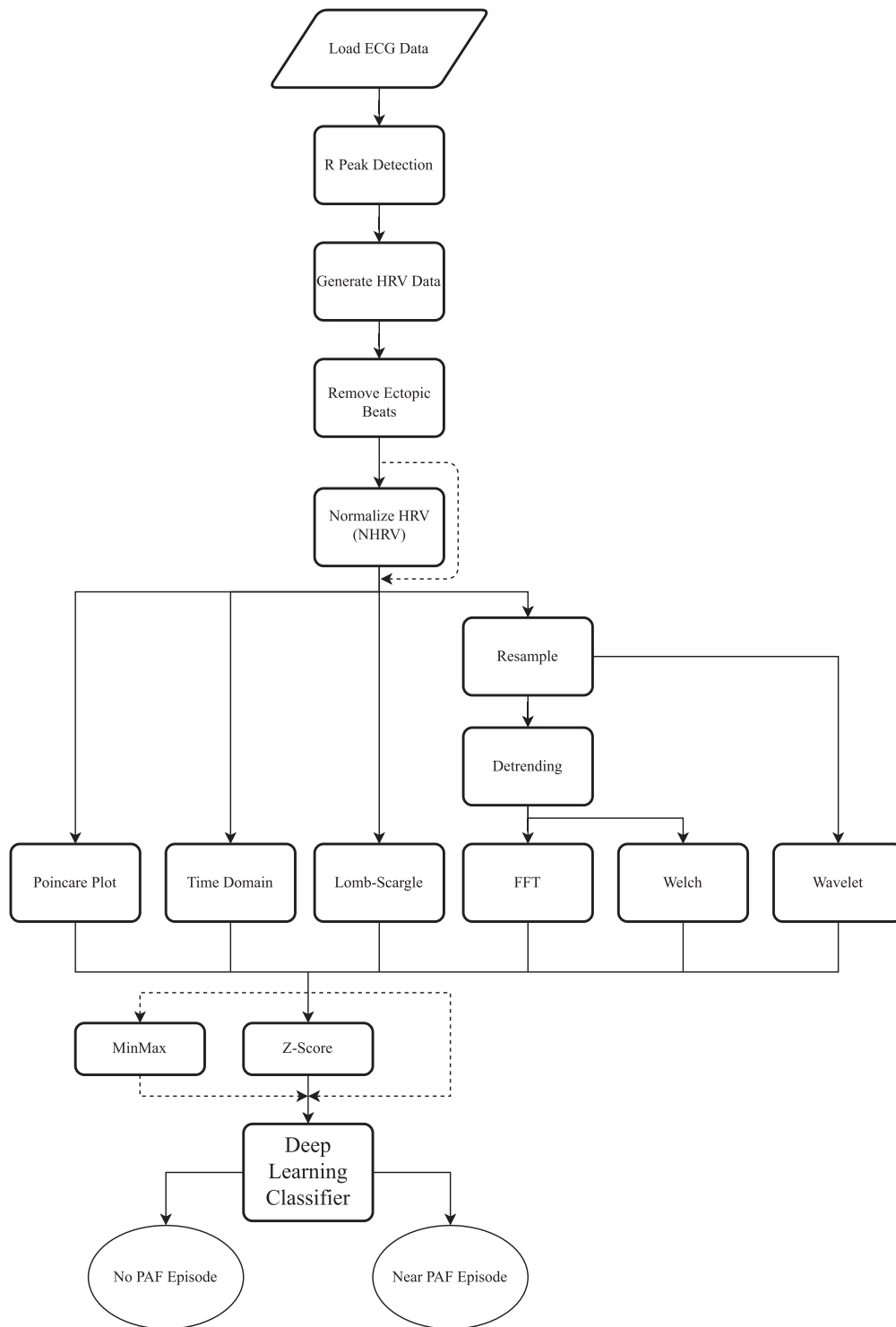
### B. Beat detection and HRV data

QRS complexes in ECG are matched by blood-pumping activity in the body. Also, since its detection is easier than the detection of other waveforms in ECG, QRS complexes are assumed as heartbeat times, in general. Hence, the interval variations among these peaks create the beat time-series data [or heart rate variability (HRV) intervals].

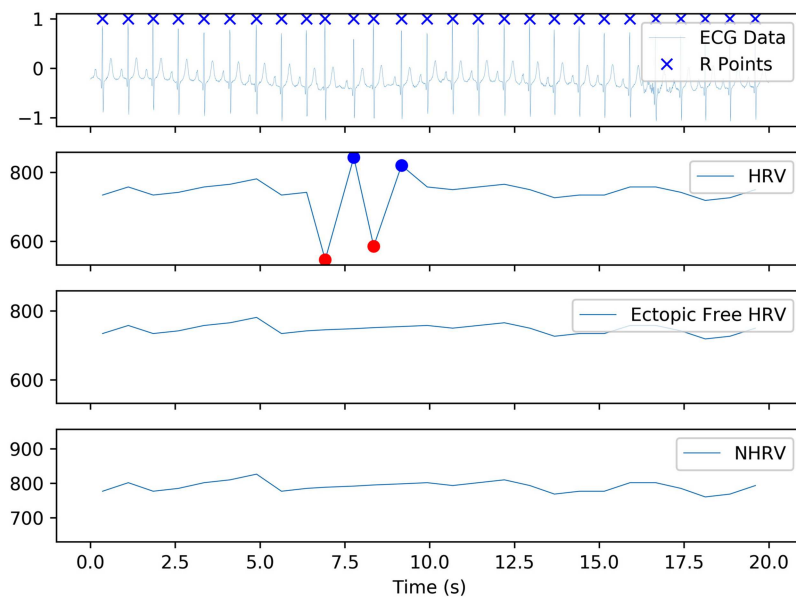
For this purpose, the ECG signal is objected to a bandpass filter, differentiation, squaring, time-averaging, and applying thresholds, in an order.<sup>39</sup> HRV data are the interval variations between successive data vs their occurrences. Since the peak intervals are not identical, the HRV data shift to an irregularly sampled nature.

### C. Removing ectopic beats

Heartbeats that are not originating from the sinoatrial node of the heart are called ectopic beats. The workgroup has offered to remove these irregular beats before the HRV analysis.<sup>10</sup> Langley *et al.*<sup>40</sup> presented an easy method to detect these premature beats. In



**FIG. 1.** Block diagram of the proposed study. The dashed line shows the optional processes. Solid-line optional processes (normalize HRV and Z-score normalization) resulted in the maximum classifier accuracy in this study.



**FIG. 2.** Removing ectopic beats from HRV data. The horizontal axis shows the time in ms where the vertical axis shows the amplitude. The top figure shows the raw ECG data and QRS complexes, the next figure shows ectopic points where data change higher than 20% of the average value, the third figure visualizes the ectopic-free data, and the last figure shows heart rate normalized ectopic-free heart rate variability data.

this method, a heartbeat and its following heartbeat shape an ectopic beat together if the occurring time of the heartbeat is 20% less than the average occurring time. Simply eliminating these beats from the data is the ectopic removal process. Figure 2 shows the detection and removal of ectopic beats. From top to bottom, this figure plots an example ECG record, heart rate data derived from the ECG, and ectopic-free heart rate data. The red points emphasize the earlier heartbeats, and blue dots are succeeding heartbeats.

#### D. Heart rate normalization (optional)

Heart rate normalization, pioneered by Hallstrom *et al.*,<sup>41</sup> is a process to remove the mean from HRV data by

$$data_{NHRV} = \frac{60}{new\_bpm} \times \frac{data_{HRV}}{mean(data_{HRV})}, \quad (1)$$

$$data_{NHRV} = \frac{1000}{new\_mean\_period} \times \frac{data_{HRV}}{mean(data_{HRV})}, \quad (2)$$

where  $data_{NHRV}$  is heart rate normalized HRV data,  $data_{HRV}$  is HRV data,  $new\_bpm$  is the new beat per minutes,  $new\_mean\_period$  is the mean of new heartbeats, and  $mean(data_{HRV})$  is the mean of HRV data given.

HRV features, calculated from HRV data after this heart rate normalization, are called normalized HRV (NHRV) features. NHRV analysis increased classifier performances in CHF diagnosis,<sup>16</sup> systolic dysfunction diagnosis,<sup>14</sup> and PAF diagnosis.<sup>20,29</sup> Therefore, we used mean values as 75 beats/min (or 800 ms in period) in this study as offered in the original article<sup>41</sup> and some similar research.<sup>14,16</sup>

#### E. Re-sampling and de-trending

HRV data are unevenly sampled and contain non-stationary components. Some feature extraction methods (FFT, Welch, and

Wavelet) require the data sampled at equal time intervals.<sup>42</sup> Re-sampling (or interpolation) is the solution to cope with this issue. Although there are several interpolation methods in the literature, Clifford and Tarassenko<sup>43</sup> have offered a robust interpolation method, called cubic-spline, to re-sample HRV data. The various sampling frequencies (number of sample points in a second) of 1–10 Hz have been used.<sup>44</sup> We preferred 4 Hz, similar to our previous studies, in this study.

All linear feature extraction methods use linear data, naturally. The possible nonlinear components of the data disturb the results. Slowly changing polynomials or sinusoidal trends are common non-stationarity origins.<sup>45</sup> In recent years, Tarvainen *et al.*<sup>46</sup> pioneered the smoothness priors method to make the data non-stationary by the following equation:

$$R_{detrended} = (I - (I + \lambda D_2^T D_2)^{-1})R, \quad (3)$$

where  $R$  is HRV data,  $R_{detrended}$  is the stationary HRV data,  $I$  is the unity matrix,  $\lambda$  is the regularity parameter,  $D_2$  is the second-order derivative operator, and  $T$  is the transpose of a matrix. We used  $\lambda = 1000$  as offered in Ref. 46.

#### F. Feature extraction

Feature extraction discovers the data to calculate features.<sup>47</sup> We decided to extract 16 time-domain features, 24 frequency-domain features from 4 different transforms, 4 time-frequency wavelet entropy measures, and 4 non-linear features, which equals 48 features in total. Since these methods are outlined here only, detailed information can be discovered in Ref. 6.

##### 1. Time-domain features

The time-domain features are calculated from the raw time-series data, in general. These features hold statistical measures such

**TABLE I.** Commonly used frequency-domain heart rate variability measures.

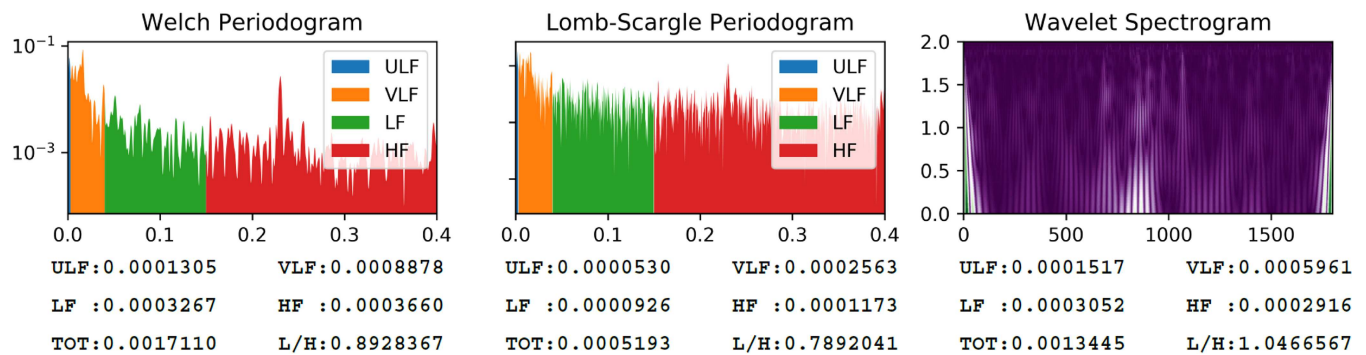
Features	Spectral power	Frequency range (Hz)
$P_{ULF}$	In the ultra-low-frequency (ULF) band	0.000–0.003
$P_{VLF}$	In the very-low-frequency (VLF) band	0.003–0.040
$P_{LF}$	In the low-frequency (LF) band	0.040–0.150
$P_{HF}$	In the high-frequency (HF) band	0.150–0.400
$P_{Total}$	In total	0.000–0.400
$\frac{P_{LF}}{P_{HF}}$	The ratio of $P_{LF}$ to $P_{HF}$	0.040–0.400

as mean, minimum, maximum, standard deviation (SDNN), root means square of successive differences (RMSSD), standard deviation of successive differences (SDSD), NN50 (the number of successive differences greater than 50 ms), NN20 (the number of successive differences greater than 20 ms), PNN50 (the ratio of NN50), PNN20 (the ratio of NN20), etc., as offered in Ref. 10.

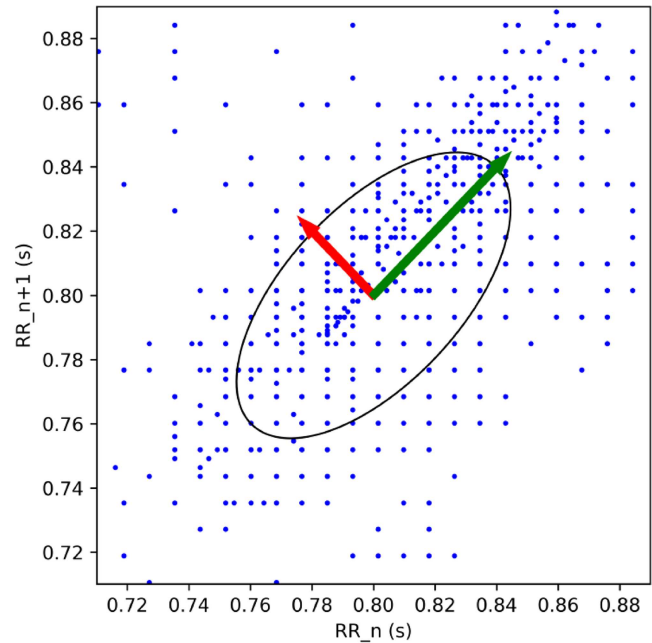
### 2. Frequency-domain features

These features are estimated from the power spectral density (PSD) estimation. There are some methods to estimate PSD that require some preprocessing steps. Among them, the Lomb–Scargle algorithm does not require re-sampling and de-trending sub-steps, but it is a computationally expensive method compared to other methods.<sup>48,49</sup> On the other hand, the Fast Fourier Transform (FFT) and Welch Periodogram algorithms require both re-sampling and de-trending steps since these methods can only work on evenly sampled stationary data.<sup>43,50</sup>

The spectrum in the HRV analysis contains four frequency-band components: ultra-low-frequency (ULF), very-low-frequency (VLF), low-frequency (LF), and high-frequency (HF) bands.<sup>10</sup> The total power, powers of each frequency band, and the ratio of LF band to HF band are frequency-domain features (Table I).<sup>6,10</sup> These features can be calculated from the FFT-based periodogram, the Welch periodogram, and the Lomb–Scargle periodogram similar to previous studies (Fig. 3).<sup>6,15,16</sup>



**FIG. 3.** An example for the Welch periodogram, the Lomb–Scargle periodogram, and the wavelet spectrogram, with six calculated spectral HRV features for each.



**FIG. 4.** Poincaré plot example. The red arrow shows the SD1 feature, and the green arrow shows the SD2 feature from the fitted ellipse.

### 3. Nonlinear features

The Poincaré plot (Fig. 4) reflects the nonlinearity. Since heartbeats show nonlinear nature, the Poincaré plot becomes popular in HRV studies.<sup>51,52</sup> There are two commonly used measures<sup>53</sup> derived from the fitted ellipse method<sup>54</sup> on this plot,

$$SD_1 = \sqrt{\frac{1}{2}(SDSD)^2}, \tag{4}$$

$$SD_2 = \sqrt{2(SDNN)^2 - \frac{1}{2}(SDSD)^2}, \tag{5}$$

where  $SDSD$  and  $SDNN$  are standard time-domain features. In addition, there are two measures,<sup>6</sup> derived from  $SD_1$  and  $SD_2$ , that are also calculated as

$$SD_1SD_2 = SD_1 \times SD_2, \tag{6}$$

$$RATIO = \frac{SD_1}{SD_2}. \tag{7}$$

#### 4. Time-frequency-domain features

The wavelet transform can examine signals in both the time and frequency domains. It also eliminates polynomial non-stationarity.<sup>55</sup> Various features extracted from the wavelet transform have been used in HRV studies.<sup>6,15</sup>

Although choosing an appropriate mother wavelet is an important issue,<sup>56,57</sup> many HRV-related studies preferred Daubechies-4 as a mother wavelet.<sup>14-16</sup> This study applied Daubechies-4 with the level of 7 to the re-sampled data as reported enough to group wavelet packets into specific HRV frequency bands.<sup>6</sup>

The energy of each coefficient was calculated using the following equation:

$$E_j = C_j^2, \tag{8}$$

where  $C_j$  is the wavelet coefficients. The total energy of an HRV band,  $E_f$ , was calculated separately,

$$E_f = \sum_{j \in f} E_j, \tag{9}$$

where  $f$  represents the HRV frequency band.<sup>6</sup> The wavelet entropy features ( $ENT_f$ ) were calculated as follows:

$$ENT_f = - \sum_{j \in f} (p_j \log_2(p_j)), \tag{10}$$

where the probability of energies of all frequency ( $f$ ) values in the frequency band of interest is calculated as  $p_j$ ,<sup>6,55</sup> and  $p_j$  is the value

TABLE II. Time-frequency-domain heart rate variability measures.

Features	Description
$ENT_{ULF}$	Wavelet entropy of the ultra-low-frequency (ULF) band
$ENT_{VLF}$	Wavelet entropy of the very-low-frequency (VLF) band
$ENT_{LF}$	Wavelet entropy of the low-frequency (LF) band
$ENT_{HF}$	Wavelet entropy of the high-frequency (HF) band

obtained by dividing the energy of the frequency of interest by the total band energy ( $\frac{E_j}{E_f}$ ). The entropy features were calculated for the standard frequency bands (Table II).

#### G. Feature normalization

Since ranges of features are very different, high-value features affect the classifier performances more than low-value features.<sup>47</sup> Eliminating this negative effect is very important in many pattern recognition applications.<sup>20,29</sup> There are two commonly used feature normalization methods: Min-Max normalization (11) and Z-score normalization (12). The min-max normalized ( $f_i^{min-max}$ ) and z-score ( $f_i^{z-score}$ ) normalized samples can be calculated using

$$f_i^{min-max} = \frac{f_i - \min(f)}{\max(f) - \min(f)}, \tag{11}$$

$$f_i^{z-score} = \frac{f_i - \mu_f}{\sigma_f}, \tag{12}$$

where  $f_i$  is the  $i$ th sample,  $\min(f)$  is the minimum value,  $\max(f)$  is the maximum value,  $\mu_f$  is the average value, and  $\sigma_f$  is the standard deviation of the feature  $f$ .

This study iterates HRV and NHRV features using no normalization, min-max normalization, and z-score normalization and reports corresponding classifier performances to examine the effect of feature normalization methods.

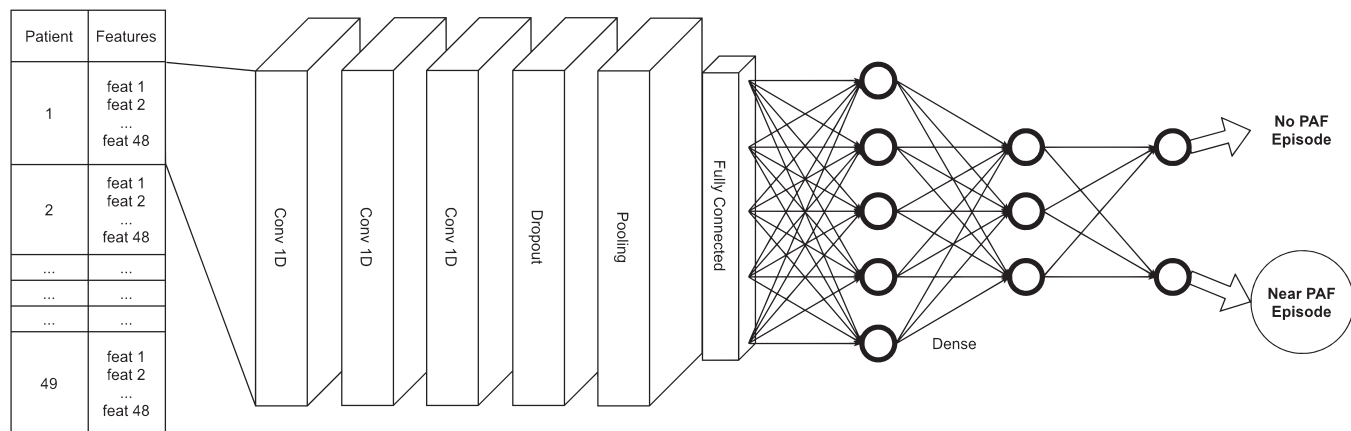


FIG. 5. Block diagram of the proposed convolutional neural network classifier. It accepts extracted features to its inputs, and it gives a decision whether there is a following paroxysmal atrial fibrillation (PAF) episode or not.

## H. Classifier: Convolutional neural network

The Convolutional Neural Network (CNN) is the most famous Deep learning (DL) model. DL seems more complex than commonly accepted artificial neural network models by having more layers.<sup>58-60</sup> Since CNN exhibits outstanding achievement in analyzing two-dimensional (2D) pictures and three-dimensional (3D) videos, it becomes popular in one-dimensional (1D) time-series data classification recently.<sup>35,36</sup>

A generic CNN classifier has an input layer, one or more convolution blocks, a dropout block, a pooling block, one (or more) fully connected classifier block(s), a dense layer, and an output layer. The input layer takes 1D data from the physical system, and the output layer exhibits the judgment given by the classifier. Other blocks reveal some arithmetical formulas. For instance, Conv1D layers calculate convolutions with varied dimensions and acting filters to evoke cryptic knowledge from the data. The dropout layer stops overfitting by disengaging random links from input to output, which is described as regularization.<sup>61</sup> The pooling layer diminishes the number of parameters by calculating averages of small-size input boxes. The fully connected layer runs as a hidden layer of a traditional Multi-Layer Perceptron (MLP), as usual. In some papers, different algorithms (such as k-nearest neighbors) are favored alternatively.<sup>26</sup> The dense layer involves activation functions (ReLU, tanh, or sigmoid) to assess natural outputs. Figure 5 reflects the classification stage of our model. No PAF Episode is the decision of No-Near Attack PAF Patient (or negative), and Near PAF Episode is the decision of PAF Patient with the Following Episode (or positive) in this study.

All layers between the input and output layers should be determined by trials or by experience<sup>62</sup> since the performance of a CNN network is strictly dependent on its hyper-parameters.<sup>63</sup> We proposed this model by tuning hyper-parameters in 1000 iterations.

## I. Evaluating classifier performances

In pattern recognition applications, the performance of a classifier is calculated by its responses against unseen inputs before.<sup>47</sup> All samples were divided into two parts by 90% for training and 10% for testing the classifier in this study. The performance of the classifier is calculated using test samples, while hyperparameters of the classifier are adjusted using the training data.

The confusion matrix is calculated for evaluating classifier performances. It is generated by comparing the responses of the classification algorithm to the test set with the actual values in the dataset.

TABLE III. Python libraries with versions used in this study.

Library	Version	Description
hyperopt	0.2.5	Hyper-parameter optimization module
Keras	2.4.3	ANNs for the TensorFlow library
Matplotlib	3.1.2	To show image, animation, and views
Numpy	1.18.1	Scientific calculations
Pandas	1.2.3	Data analysis and manipulation tool
PyWavelets	1.1.1	To calculate the wavelet transform
Scipy	1.4.1	Engineering, scientific, and mathematical tools
Sklearn	0.24.1	Predictive data analysis
Tensorflow	2.4.0	To develop and train machine learning models

In the case of two-class problems, one class is positive, and the other is negative. True Positive (TP) is the number of patients classified correctly, and True Negative (TN) is the number of healthy subjects classified correctly. On the other hand, False Negative (FN) is the number of patients misclassified as healthy ones, and False Positive (FP) is the number of healthy subjects misclassified as patients.<sup>47,64</sup> In this study, we assigned the negative class for PAF patients with no-near attack and the positive class for PAF patients with near-future attack.

Four commonly used performance measures (Accuracy, Recall, Precision, and F1-Score) were used to evaluate classifiers in this study.<sup>47,65</sup>

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}, \quad (13)$$

$$Recall = \frac{TP}{TP + FN}, \quad (14)$$

$$Precision = \frac{TP}{TP + FP}, \quad (15)$$

$$F1 - Score = \frac{Precision \times Recall}{Precision + Recall} \quad (16)$$

TABLE IV. Classifier performances achieved in this study in predicting whether PAF patients will have the following episode or not. The bold-indicated row emphasizes the highest classifier accuracy among them.

Features	Feature normalization	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
HRV	No normalization	51.02	52.00	52.00	52.00
	Min-max	83.67	86.96	80.00	83.33
	Z-score	87.76	91.30	84.00	87.50
NHRV	No normalization	53.06	53.57	60.00	56.60
	Min-max	93.88	95.83	92.00	93.88
	<b>Z-score</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>	<b>100.0</b>

### III. RESULTS AND DISCUSSION

This study examined the impacts of both heart rate and feature normalizations in the onset prediction of PAF attacks using CNN. HRV features were calculated from both 30 min HRV data and heart rate normalized 30 min HRV data (NHRV). In addition, two feature normalization methods were applied. The CNN classifier experimented with six distinct feature sets. [Table III](#) listed the libraries used in the programming language of Python 3.8 to discriminate PAF patients as to whether they have a near attack or not.

The first part of our study repeated the classifiers from the HRV features for no normalization, min–max normalization, and z-score normalization cases. The achieved classifier accuracies are 51.02%, 83.67%, and 87.76% ([Table IV](#)), respectively. Since both the min–max and z-score normalizations give higher accuracies, the CNN classifier using one of these two normalization methods seems enough in predicting the near PAF episodes based on HRV data.

The second part of our study repeated the classifiers from the NHRV features for the same normalization cases. The achieved classifier accuracies become 53.06%, 93.88%, and 100.0% ([Table IV](#)), respectively. Although the min–max normalization gives satisfactorily noticeable accuracy again, the z-score normalization method results in excellent accuracy. As a result, the CNN classifier using both the heart rate normalization and z-score feature normalization methods together is enough to predict PAF episodes from 30 min NHRV data.

[Table V](#) summarizes PAF diagnosis studies using 30 min HRV data given in the literature. The classifier accuracies from the literature varied from 78.00% to 94.50%. Our proposed method, using the CNN classifier with the heart rate normalized and z-score normalized heart rate variability features, gives the highest classifier accuracy of 100%. By comparing the accuracy of the proposed classifier to other studies, our method has superior accuracy among them, to our knowledge.

**TABLE V.** Studies to determine recent PAF episodes using 30 min heart rate variability measures in the literature. The bold-indicated row emphasizes the highest classifier accuracy.

Study	Features	Classifier	Description	Accuracy (%)
Ref. 28	HRV	kNN	...	78.00
Ref. 29	NHRV	RBF	Min–max	81.63
Ref. 30	HRV	MLP	...	82.05
Ref. 29	HRV	RBF	Min–max	83.67
Ref. 19	HRV	SVM	GAs	83.90
Ref. 18	ECG	FL	...	84.00
Ref. 32	HRV	SVM	GAs	86.80
This study	HRV	CNN	Z-score	87.76
Ref. 66	ECG	DT	...	88.00
Ref. 21	ECG	DT	...	89.29
Ref. 67	ECG + HRV	DT	...	89.40
Ref. 17	HRV	LDA	...	92.00
Ref. 5	HRV	SVM	...	94.50
<b>This study</b>	<b>NHRV</b>	<b>CNN</b>	<b>Z-score</b>	<b>100.0</b>

Consequently, the proposed method in this study results in better classification performance than similar studies in the literature. We may propose that our approach provides a better tool to determine a PAF episode among PAF patients. On the other hand, CNN works well with big data. The database used in this study consists of ECG data obtained from only 49 patients with PAF only, which is a weakness of this study. The findings obtained in our study, which are preliminary research in character, need to be supported by other studies on larger datasets.

### ACKNOWLEDGMENTS

M. Perc was supported by the Slovenian Research Agency (Grant Nos. P1-0403 and J1-2457).

### AUTHOR DECLARATIONS

#### Conflict of Interest

The authors have no conflicts to disclose.

#### Ethics Approval

Ethics approval is not required.

#### Author Contributions

The subject of this paper is part of M.S.'s Ph.D. thesis. R.K. is the supervisor and Y.I. is the co-supervisor of the thesis. M.P. offers suggestions. All authors contributed equally to this work.

### DATA AVAILABILITY

The data that support the findings of this study are openly available in PhysioNet at <https://www.physionet.org/content/afpdb/1.0.0/>, Ref. 3.

### REFERENCES

- <sup>1</sup>K. M. Ryder and E. J. Benjamin, "Epidemiology and significance of atrial fibrillation," *Am. J. Cardiol.* **84**, 131–138 (1999).
- <sup>2</sup>X. Du, L. Guo, S. Xia, J. Du, C. Anderson, H. Arima, M. Huffman, Y. Yuan, Y. Zheng, S. Wu, X. Guang, X. Zhou, H. Lin, X. Cheng, J. Dong, and C. Ma, "Atrial fibrillation prevalence, awareness and management in a nationwide survey of adults in China," *Heart* **107**, 535–541 (2021).
- <sup>3</sup>G. B. Moody, 2001, PAF Prediction Challenge Database, published in PhysioNet, Version: 1.0.0, <https://doi.org/10.13026/C2H59W>
- <sup>4</sup>G. Hindricks, T. Potpara, N. Dagres, E. Arbelo, J. J. Bax, C. Blomstrom-Lundqvist, G. Boriani, M. Castella, G.-A. Dan, P. E. Dilaveris, L. Fauchier, G. Filippatos, J. M. Kalman, M. La Meir, D. A. Lane, J.-P. Lebeau, M. Lettino, G. Y. H. Lip, F. J. Pinto, G. N. Thomas, M. Valgimigli, I. C. Van Gelder, B. P. Van Putte, C. L. Watkins, and ESC Scientific Document Group, "2020 ESC guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS)," *Eur. Heart J.* **42**, 373–498 (2020).
- <sup>5</sup>M. Mohebbi and 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 Programs Biomed.* **105**, 40–49 (2012).
- <sup>6</sup>Y. Isler and M. Kuntalp, "Combining classical HRV indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure," *Comput. Biol. Med.* **37**, 1502–1510 (2007).



- <sup>7</sup>A. Soudani and M. Almusallam, "Atrial fibrillation detection based on ECG-features extraction in WBSN," *Procedia Comput. Sci.* **130**, 472–479 (2018).
- <sup>8</sup>U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network," *Knowl. Based Syst.* **132**, 62–71 (2017).
- <sup>9</sup>H. Guruler, M. Sahin, and A. Ferikoglu, "Feature selection on single-lead ECG for obstructive sleep apnea diagnosis," *Turk. J. Electr. Eng. Comput. Sci.* **22**, 465–478 (2014).
- <sup>10</sup>Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, "Heart rate variability," *Circulation* **93**, 1043–1065 (1996).
- <sup>11</sup>M. G. Tsiouras and D. I. Fotiadis, "Automatic arrhythmia detection based on time and time–frequency analysis of heart rate variability," *Comput. Methods Programs Biomed.* **74**, 95–108 (2004).
- <sup>12</sup>M.-Y. Lee and S.-N. Yu, "Multiscale sample entropy based on discrete wavelet transform for clinical heart rate variability recognition," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (IEEE, 2012).
- <sup>13</sup>J. Park, S. Lee, and M. Jeon, "Atrial fibrillation detection by heart rate variability in Poincaré plot," *Biomed. Eng. Online* **8**, 38 (2009).
- <sup>14</sup>Y. Isler, "Discrimination of systolic and diastolic dysfunctions using multi-layer perceptron in heart rate variability analysis," *Comput. Biol. Med.* **76**, 113–119 (2016).
- <sup>15</sup>A. Narin, Y. Isler, M. Ozer, and M. Perc, "Early prediction of paroxysmal atrial fibrillation based on short-term heart rate variability," *Physica A* **509**, 56–65 (2018).
- <sup>16</sup>Y. Isler and M. Kuntalp, "Heart rate normalization in the analysis of heart rate variability in congestive heart failure," *Proc. Inst. Mech. Eng., Part H: J. Eng. Med.* **224**, 453–463 (2009).
- <sup>17</sup>C. Maier, M. Bauch, and H. Dickhaus, "Screening and prediction of paroxysmal atrial fibrillation by analysis of heart rate variability parameters," in *Computers in Cardiology 2001. Vol. 28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>18</sup>G. Schreier, P. Kastner, and W. Marko, "An automatic ECG processing algorithm to identify patients prone to paroxysmal atrial fibrillation," in *Computers in Cardiology 2001. Vol. 28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>19</sup>K. H. Boon, M. Khalil-Hani, M. B. Malarvili, and C. W. Sia, "Paroxysmal atrial fibrillation prediction method with shorter HRV sequences," *Comput. Methods Programs Biomed.* **134**, 187–196 (2016).
- <sup>20</sup>M. Surucu, Y. Isler, and R. Kara, "Investigation of the effect of normalization techniques on discriminating patients with paroxysmal atrial fibrillation," in *2nd International Conference of Applied Sciences, Engineering and Mathematics* (IBU-ICASEM 2020, Skopje/North Macedonia, 2020).
- <sup>21</sup>T. Thong, J. McNames, M. Aboy, and B. Goldstein, "Prediction of paroxysmal atrial fibrillation by analysis of atrial premature complexes," *IEEE Trans. Biomed. Eng.* **51**, 561–569 (2004).
- <sup>22</sup>E. Ros, S. Mota, F. J. Fernández, F. J. Toro, and J. L. Bernier, "ECG characterization of paroxysmal atrial fibrillation: Parameter extraction and automatic diagnosis algorithm," *Comput. Biol. Med.* **34**, 679–696 (2004).
- <sup>23</sup>N. Ozcan and M. Kuntalp, "Determining best hrv indices for paf screening using genetic algorithm," in *10th International Conference on Electrical and Electronics Engineering (ELECO)* (ELECO, 2017), pp. 1385–1388.
- <sup>24</sup>Y. V. Chesnokov, A. V. Holden, and H. Zhang, "Screening patients with paroxysmal atrial fibrillation (PAF) from non-PAF heart rhythm using HRV data analysis," in *2007 Computers in Cardiology* (IEEE, 2007).
- <sup>25</sup>A. Martínez, R. Alcaraz, and J. J. Rieta, "Study on the p-wave feature time course as early predictors of paroxysmal atrial fibrillation," *Physiol. Meas.* **33**, 1959–1974 (2012).
- <sup>26</sup>B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani, "Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients," *IEEE Trans. Syst., Man, Cybern.: Syst.* **48**, 2095–2104 (2018).
- <sup>27</sup>M. Surucu, Y. Isler, and R. Kara, "Diagnosis of paroxysmal atrial fibrillation from thirty-minute heart rate variability data using convolutional neural networks," *Turk. J. Electr. Eng. Comput. Sci.* **29**, 2986–2990 (2021).
- <sup>28</sup>K. S. Lynn and H. D. Chiang, "A two-stage solution algorithm for paroxysmal atrial fibrillation prediction," in *Computers in Cardiology 2001. Vol. 28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>29</sup>M. Surucu, Y. Isler, and R. Kara, "Heart rate normalization on determining a paroxysmal atrial fibrillation attack," in *3rd International Conference of Applied Sciences, Engineering and Mathematics, IBU-ICASEM 2021, Skopje/North Macedonia* (IBU, 2021).
- <sup>30</sup>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**, 151–165 (2008).
- <sup>31</sup>K. H. Boon, M. Khalil-Hani, and M. B. Malarvili, "Paroxysmal atrial fibrillation prediction based on HRV analysis and non-dominated sorting genetic algorithm III," *Comput. Methods Programs Biomed.* **153**, 171–184 (2018).
- <sup>32</sup>K. H. Boon, M. Khalil-Hani, and C. W. Sia, "Paroxysmal atrial fibrillation onset prediction using heart rate variability analysis and genetic algorithm for optimization," in *Proceedings of the International Conference on Data Engineering 2015 (DaEng-2015)* (Springer, Singapore, 2019), pp. 609–617.
- <sup>33</sup>O. Balli and Y. Kutlu, "Effect of deep learning feature inference techniques on respiratory sounds," *J. Intell. Syst. Appl.* **3**, 134–137 (2020).
- <sup>34</sup>Y. Camgozlu and Y. Kutlu, "Examining the difference between image size, background color, gray picture and color picture in leave classification with deep learning," *J. Intell. Syst. Appl.* **3**, 130–133 (2020).
- <sup>35</sup>G. Altan, Y. Kutlu, A. O. Pekmezci, and S. Nural, "Deep learning with 3D-second order difference plot on respiratory sounds," *Biomed. Signal Process. Control* **45**, 58–69 (2018).
- <sup>36</sup>G. Altan, Y. Kutlu, and N. Allahverdi, "Deep learning on computerized analysis of chronic obstructive pulmonary disease," *IEEE J. Biomed. Health Inform.* **24**, 1344–1350 (2020).
- <sup>37</sup>G. Moody, A. Goldberger, S. McClennen, and S. Swiryn, "Predicting the onset of paroxysmal atrial fibrillation: The computers in cardiology challenge 2001," in *Computers in Cardiology 2001. Vol.28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>38</sup>R. Seker, S. Saliu, A. Birand, and G. Kudaiberdieva, "Validity test for a set of nonlinear measures for short data length with reference to short-term heart rate variability signal," *J. Syst. Integr.* **10**, 41–53 (2000).
- <sup>39</sup>J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.* **BME-32**, 230–236 (1985).
- <sup>40</sup>P. Langley, D. di Bernardo, J. Allen, E. Bowers, F. E. Smith, S. Vecchiotti, and A. Murray, "Can paroxysmal atrial fibrillation be predicted?," in *Computers in Cardiology 2001. Vol. 28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>41</sup>A. P. Hallstrom, P. K. Stein, R. Schneider, M. Hodges, G. Schmidt, and K. Ulm, "Structural relationships between measures based on heart beat intervals: Potential for improved risk assessment," *IEEE Trans. Biomed. Eng.* **51**, 1414–1420 (2004).
- <sup>42</sup>G. G. Berntson, J. T. Bigger, Jr., D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. van der Molen, "Heart rate variability: Origins, methods, and interpretive caveats," *Psychophysiology* **34**, 623–648 (1997).
- <sup>43</sup>G. D. Clifford and L. Tarassenko, "Quantifying errors in spectral estimates of HRV due to beat replacement and resampling," *IEEE Trans. Biomed. Eng.* **52**, 630–638 (2005).
- <sup>44</sup>G. D. Clifford, F. Azuaje, and P. E. McSharry, *Advanced Methods and Tools for ECG Data Analysis* (Artech House, Inc., Norwood, MA, 2006).
- <sup>45</sup>E. J. M. Weber, P. C. M. Molenaar, and M. W. van der Molen, "A nonstationarity test for the spectral analysis of physiological time series with an application to respiratory sinus arrhythmia," *Psychophysiology* **29**, 55–62 (1992).
- <sup>46</sup>M. P. Tarvainen, P. O. Ranta-Aho, and P. A. Karjalainen, "An advanced detrending method with application to HRV analysis," *IEEE Trans. Biomed. Eng.* **49**, 172–175 (2002).
- <sup>47</sup>R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. (Wiley, New York, 2001).
- <sup>48</sup>N. R. Lomb, "Least-squares frequency analysis of unequally spaced data," *Astrophys. Space Sci.* **39**, 447–462 (1976).
- <sup>49</sup>J. D. Scargle, "Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data," *Astrophys. J.* **263**, 835 (1982).

- <sup>50</sup>P. Laguna, G. B. Moody, and R. G. Mark, "Power spectral density of unevenly sampled data by least-square analysis: Performance and application to heart rate signals," *IEEE Trans. Biomed. Eng.* **45**, 698–715 (1998).
- <sup>51</sup>P. W. Kamen and A. M. Tonkin, "Application of the Poincaré plot to heart rate variability: A new measure of functional status in heart failure," *Aust. N. Z. J. Med.* **25**, 18–26 (1995).
- <sup>52</sup>P. W. Kamen, H. Krum, and A. M. Tonkin, "Poincaré plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans," *Clin. Sci.* **91**, 201–208 (1996).
- <sup>53</sup>M. Brennan, M. Palaniswami, and P. Kamen, "Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability?," *IEEE Trans. Biomed. Eng.* **48**, 1342–1347 (2001).
- <sup>54</sup>F. Marciano, M. L. Migaux, D. Acanfora, G. Furgi, and F. Rengo, "Quantification of Poincaré maps for the evaluation of heart rate variability," in *Computers in Cardiology 1994* (IEEE Computer Society Press, 1994).
- <sup>55</sup>R. Q. Quiroga, O. A. Rosso, E. Basar, and M. Schurmann, "Wavelet entropy in event-related potentials: A new method shows ordering of EEG oscillations," *Biol. Cybern.* **84**, 291–299 (2001).
- <sup>56</sup>E. Sayilgan, Y. Yuçe, and Y. Isler, "Evaluation of wavelet features selected via statistical evidence from steady-state visually evoked potentials to predict the stimulating frequency," *J. Fac. Eng. Archit. Gazi Univ.* **36**, 593 (2021).
- <sup>57</sup>E. Sayilgan, Y. K. Yuçe, and Y. Isler, "Evaluation of mother wavelets on steady-state visually-evoked potentials for triple-command brain-computer interfaces," *Turk. J. Electr. Eng. Comput. Sci.* **29**, 2263–2279 (2021).
- <sup>58</sup>J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.* **61**, 85–117 (2015).
- <sup>59</sup>Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
- <sup>60</sup>Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1798–1828 (2013).
- <sup>61</sup>N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
- <sup>62</sup>J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.* **13**, 281–305 (2012).
- <sup>63</sup>A. H. Abdi, C. Luong, T. Tsang, G. Allan, S. Nouranian, J. Jue, D. Hawley, S. Fleming, K. Gin, J. Swift, R. Rohling, and P. Abolmaesumi, "Automatic quality assessment of echocardiograms using convolutional neural networks: Feasibility on the apical four-chamber view," *IEEE Trans. Med. Imaging* **36**, 1221–1230 (2017).
- <sup>64</sup>Y. Isler, A. Narin, and M. Ozer, "Comparison of the effects of cross-validation methods on determining performances of classifiers used in diagnosing congestive heart failure," *Meas. Sci. Rev.* **15**, 196–201 (2015).
- <sup>65</sup>D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over f1 score and accuracy in binary classification evaluation," *BMC Genom.* **21**, 6 (2020).
- <sup>66</sup>W. Zong, R. Mukkamala, and R. G. Mark, "A methodology for predicting paroxysmal atrial fibrillation based on ECG arrhythmia feature analysis," in *Computers in Cardiology 2001. Vol. 28 (Cat. No. 01CH37287)* (IEEE, 2001).
- <sup>67</sup>H. Costin, C. Rotariu, and A. Pasarica, "Atrial fibrillation onset prediction using variability of ECG signals," in *2013 8TH International Symposium on Advanced Topics in Electrical Engineering (ATEE)* (IEEE, 2013).